Multivariate methods and the search for single top-quark production in association with a W boson in ATLAS

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

von
Péter Kövesárki
aus
Budapest

Bonn, 10.9.2012

Acknowledgements

Contents

Introduction

1.1 Introduction

In the real world we recognise objects and action without paying attention to the process of recognition itself. It is an unconscious part of the observation, but a very difficult part indeed. Imagine looking at a picture, seeing an animal, a bear for example. In that moment the brain receives a huge amount of information, a few million bits representing colours and pixels from the eye and it can reduce this by orders of magnitude to short sentences which can be communicated later on more easily. The very same thing is happening in modern particle physics experiments. Physicists would like to interpret the millions of bits of information from every particle collision recorded in the detectors by gradually decreasing this size in order to extract only a few bits of information representing if, for example a given event was either a $t\bar{t}$ or a W+jets production. What happens with the original information content in the meantime and why were a number of bits eliminated in the process? How much information do we lose with this?

We can assume that we lose information if the original picture can not be reconstructed, and in many cases this is what is happening. To reconstruct a picture, we naturally just read a text and imagine the mentioned objects and environment. This obviously must be based on previous experience on the subject, and the capability to associate the current and previous pictures with simple symbols. What helps here most, is the abstract idea of objects. In this sense an object is a specialised entity of a class, that has only a finite number of parameters. In case we assume that the given picture can be described by using a sum of objects, then it follows that the picture is a specialised entity of another representation of the class. There might not be one-to-one correspondence between a series of pictures and a class, as some internal parameters may remain hidden, but if the picture can be classified in this way, the problem of information loss more or less disappears. This is because everyday objects have only a small number of parameters, individual solid objects in three dimensions have actually only six parameters. These fully describe the position of the object and its setting, meaning that series of pictures of that given object can be described with only this six parameters, hence giving us a possibility to reduce the size of data with the ability to reconstruct the original.

Non-solid objects have more than six parameters. Getting back to the picture of a bear, we can see if the bear is aggressive, hungry, if it is approaching us or going away and we can even imagine a parameter that represents the likelihood of the bear being actually a teddy bear. We can image just as many independent parameters as the pixels the picture has, so in that case we could describe every possible

picture with our new parameter space, but then we lose the possibility of lossless data compression. Obviously, some parameters are more important to us, because it is necessary to act on them. In extreme cases this can save our life, like in the case of an aggressive bear, so it is evolutionary beneficial to give higher ranks to such information. Other parameters, like the state of every hair in the fur of the bear seems absolutely unimportant. Ranking the parameters, and keeping only the important ones still makes it possible to reconstruct the original image accurately, though definitely losing some information.

Important to mention, usually there is no need to define all the possible parameters. Having all the parameters would allow us to describe any possible picture, including white noise as well, but obviously it is very rare that a common object becomes white noise, so there is no reason to include such a parameter. We may call it common sense, but the real reason is that the parameters of everyday objects have a very specific distribution. We may say that every brown bear is brown, grizzlies are black and polar bears are white, the distribution of their colour is very narrow, so we may be inaccurate a bit in describing it. Although we may forge a parameter that describes the resemblance of a bear to a teddy bear, we would normally never use all values of this parameter as it is not represented in reality. We would only use the extremes of it, saying if the observed object is either a real bear or a teddy bear and so such parameter would effectively become a discrete number or symbol. The simplest thing to do is just to forget such parameters and use two classes to describe reality. One for the real bears and one for the teddy bears.

This leads to the ability to describe data without serious information loss and gain a good level of compression. One can set a convenient accuracy for the reconstruction and determine what number of discrete classes is beneficial. The better compression level can be interpreted as a better understanding of the data, or in other words a better and simpler theoretical description, which is just a modern evaluation of Occam's Razor.

1.2 The structure of this thesis

How does one actually find classes? If we define classes in a way that they can be separated by discrete parameters, while the continuous parameters lead only to other members of the same class, then it is equivalent to saying that a class is represented by a hypersurface in the data. Such a hypersurface should be visible as a correlation of input parameters.

This thesis describe methods that help to investigate physical phenomena. Chapter 2 shows how the Standard Model of particle physics can be constructed via the knowledge of the observed constraints. Its first section reflects the author's point of view, a conjecture that quantum field theory is the simplest of the possible probabilistic theories that fulfil the space-time symmetries, accompanied with a general introduction to the properties of the Standard Model. Chapters 3 and 4 are about two newly developed probabilistic methods on classifying data sets. Additionally, Chapter 5 describes a newly developed regression method that was used for comparison in some of the chapters. The former one utilises a pre-defined description of known classes via a density estimation technique while the latter one finds clusters via a method searching for hyper-surfaces. The further chapters describe the theoretical and experimental methods that are relevant when performing an analysis of the data from the ATLAS detector. Chapter 6 deals with the methodology of simulating the various aspects of physical collisions. The next chapter in order, chapter 7 introduces the target of interest, the singe top-quark production and the challenges in its detection. The following three chapters are about the experimental setup of LHC and ATLAS, the architecture of data acquisition and the description of the definitions of the common objects. In chapter 11 the density estimator is applied to the single top-quark Wt-channel.

The author of this thesis contributed to three articles related to the thesis. The theory of a Green's

function based density estimator was published in the Journal of Physics: Conference Series as a proceedings to the Advanced Computing and Analysis Techniques Workshop in 2011 [1]. A pilot study prior to the data taking period was published as an ATLAS internal note about the jet selection for top physics [2]. Finally, the regression method described in Chapter 5 was sent for publication to the Journal of Machine Learning [3].

Standard Model of particle physics

The Standard Model of particle physics comes with extraordinary claims, saying that with the precise knowledge of only nineteen variables one can describe physical phenomena up to the energies reachable with current accelerators. The nineteen continuous parameters are masses of particles and interaction probabilities, accompanied with a few discrete parameters which describe the structure of the theory. The strong claim of its descriptive power lies within quantum field theory which postulates that space-time symmetries force constraints on the existence and possible behaviour of physical objects. Section 2.2 is a genuine attempt to show that the existence of quantised fields is one of the simplest models that can describe probabilistic laws and correlations on a space-time continuum. The next section introduces the formalism of quantum field theory with some of the simplest models, while the last section of this chapter summarises the further known non-space-time symmetries and interactions that build up into the Standard Model.

2.1 The importance of symmetries

Galileo Galilei, the definitive person of the scientific world in the XVI century was the first to state that two identical physical experiments should yield the same results if one of them was conducted in an environment moving with a constant speed. Though it sounds simple, it contains one of the fundamental principles that helps determining the physical laws. It tries to identify a symmetry of Nature, a feature that does not change despite the fact that the environment has changed. A formulation of such an observation gives a constraint to the possible physical laws, as they should also be invariant under such a change [4].

For example, the requirement that the outcome should not depend on the relative velocity can only be fulfilled if every possible sub-process is independent in it. This results in formulas that do not contain absolute velocities, just relative ones. Consider two velocities in the system $\vec{v}_1$ and $\vec{v}_2$, and a function $f(\vec{v}_1, \vec{v}_2)$ that describes some process in the system. A small $\Delta\vec{v}_{\text{obs}}$ change in the speed with which the observer is travelling should result in no change in f :

$$\Delta f = f(\vec{v}_1 + \Delta\vec{v}_{obs}, \vec{v}_2 + \Delta\vec{v}_{obs}) - f(\vec{v}_1, \vec{v}_2) = 0 \,.$$

Expressing this with the derivatives of f

$$\Delta f = \frac{\mathrm{d} f(\vec{v}_1, \vec{v}_2)}{\mathrm{d}\vec{v}_{obs}} \Delta\vec{v}_{obs} = \frac{\partial f(\vec{v}_1, \vec{v}_2)}{\partial\vec{v}_1} \frac{\partial\vec{v}_1}{\partial\vec{v}_{\mathrm{obs}}} \Delta\vec{v}_{\mathrm{obs}} + \frac{\partial f(\vec{v}_1, \vec{v}_2)}{\partial\vec{v}_2} \frac{\partial\vec{v}_2}{\partial\vec{v}_{\mathrm{obs}}} \Delta\vec{v}_{\mathrm{obs}} = 0 \,.$$

It was implicitly assumed before that changing the observers velocity affects the other velocities linearly, like $\vec{v}_1^{\mathrm{new}} = \vec{v}_1^{\mathrm{old}} + \Delta\vec{v}_{\mathrm{obs}}$. This has the derivative $\dfrac{\partial\vec{v}_1}{\partial\vec{v}_{\mathrm{obs}}} = 1$, and results in the following:

$$\frac{\partial f(\vec{v}_1, \vec{v}_2)}{\partial\vec{v}_1} = -\frac{\partial f(\vec{v}_1, \vec{v}_2)}{\partial\vec{v}_2} \,. \tag{2.1}$$

So the velocities $\vec{v}_1$ and $\vec{v}_2$ in this process have a similar influence on the observable f, but with a negative sign. This mutual dependency means that the actual number of degrees of freedom of f is reduced, and it can be parametrized with a function of $\vec{v}_1$ and $\vec{v}_2$, $f_{\mathrm{new}}(\vec{v}_1 - \vec{v}_2)$. It can be easily shown that this new function gives the same derivative structure, hence it is a solution. We can see in this example that a simple requirement for invariance reduces the complexity of the possible processes, therefore it is a powerful tool to find them.

What kind of symmetries can be found in Nature? Galilei found two types, the invariance under translation and under boost. Lorentz modified the picture by requiring the invariance under time transformation, and taking relativistic effects into account during the boost, namely that the speed of light is the same constant in every inertial reference frame.

2.2 A probabilistic model

When the symmetries of Nature have been successfully identified, one can ask what kind of correlations are possible. Group theory helps in parameterising surfaces which are associated with invariants under the group transformations. The Galilean group preserves the distance $\Delta s = \sqrt{\sum_i \Delta x_i^2}$, while the Lorentz group preserves the distance defined in Minkowsky space-time :

$$s^2 = t^2 - x^2 - y^2 - z^2 = x_\mu x^\mu \,.$$

The full group of flat space-time is the Poincaré group, containing translations, rotations and boosts. To derive the constraints for the possible physical processes that respects these symmetries, we should allow a broad set of functions as a starting point. There are several reasons to allow probability functions:

- probability can describe deterministic processes as well;

- there might be interactions that are inherently probabilistic;

- some stochastic and chaotic processes are easier to describe using probability theory.

The most general probability measure on a specified time slice would look like the following:

$$\mathrm{P}(A_a(x_\nu)|x_0 = t_0) = f[A_a(x_\nu)] \, \mathrm{d}\mu_{A_b} \, \mathrm{d}\mu_{x_\eta}|_{x_0 = t_0} \,,$$

where $A_a(x_\nu)$ is a general multi-valued function defined on the whole space, and the associated probability is the functional of it; while μ_{A_b} and μ_{x_η} are integral measures around the x_η point in space-time and the A_b value of a measurement. A single point is not enough to express temporal evolution of a physical system, so one should look for a probability that contains at least two points. The simplest example is the scalar function of probabilities giving answers to yes-no binary questions:

$$P(x_\nu^A, x_\nu^B) = f(x_\nu^A, x_\nu^B) \, d\mu_{AB} \,, \tag{2.2}$$

in the area around x_ν^A and x_μ^B. The probability measure $f \, d\mu_{AB}$ gives 1 when integrating over a 6 dimensional subspace, varying x_ν^A and x_ν^B but keeping the time constant:

$$\iint\limits_{\mathbb{R}^3\mathbb{R}^3} f \, d\mu_{AB} \Bigg|_{\substack{x_0^A=t^A \\ x_0^B=t^B}} = 1 \,.$$

The probability in eq. (2.2) should not change when observed from a different coordinate system, so it has to be invariant under an arbitrary group transformation G:

$$G\mathrm{P}(x_\nu^A, x_\nu^B) = \mathrm{P}(x_\nu^A, x_\nu^B) \,.$$

As one does not know *a priori* which group representation Nature is using, a general G has to be used. It can be constructed from different irreducible representations G^{irr}:

$$G = R^{-1} \begin{pmatrix} G_{a_1}^{irr} & & & \\ & G_{a_2}^{irr} & & 0 \\ & & \ddots & \\ & 0 & & G_{a_n}^{irr} \end{pmatrix} R \,,$$

where the a_i index refers to the type of the representations, and R is an arbitrary unitary matrix. G is actually an operator, as it acts on functions $\psi_s(x_\mu)$, from which one has to build up the probability $f \, d\mu_{AB}$. The irreducible representations of the Poincaré group are classified via spin and mass. The spin number is quantised, it can take any positive integer or half integer value, while the mass has to be a non-negative real value. How to compose f from ψ? It has to be a functional of some ψ_is, in a way that it gives a proper probability measure:

$$f(x_\nu^A, x_\nu^B) \, d\mu_{AB} = f[\psi_i] \,.$$

Such a measure should satisfy some basic criteria, namely that the probability of a volume should be the sum of the probabilities of the disjoint volumes that it can be combined from and that the probability of the full volume has to be unity. Gleason's theorem [5] says that for a Hilbert space with dimension ≥ 3 the only possible measure for a volume V on a function Ψ is

$$\mathrm{P}(\Psi, V) = \int\limits_V \Psi^* \Psi \, dV = \mathrm{Tr}(P_V |\Psi\rangle\langle\Psi|) = |\langle V|\Psi\rangle|^2 \,,$$

providing that on a Hilbert space the only possible probability measure is the one that is familiar from quantum mechanics, the square of the wave function's amplitude. The integral in this case is 6 dimensional, over the combined phase-space of the spatial part of x_μ^A and x_μ^B, which can be denoted as x_i^A and x_i^B. As a consequence, Ψ has to have 6 spatial coordinate, and because the group acts on 4 dimensional functions, all having the their role, hence a suitable Ψ has to have 8 space-time parameters, x_μ^A and x_μ^B. The requirement in eq. (2.2) that the full 6-dimensional volume should give 1 restricts the possible Ψs. Any square-integral of Ψ on a constant $t^A \otimes t^B$ hypersurface has to be constantly unity. This means that there exists a $U(t^A, t^B)$ unitary operator that produces it from a time-slice:

$$\Psi^{8D}(x_\mu^A, x_\mu^B) = U(t^A, t^B)\Psi^{6D}(x_i^A, x_i^B).\tag{2.3}$$

The volume V also has to be a function which G acts on, hence it has to have a spinor index. In the 8-dimensional space-time a composite $G = G^A \otimes G^B$ is needed, which acts separately on x_μ^A and x_μ^B. As an external observer has only a 4-dimensional coordinate, a change in its reference frame should both change x_μ^A and x_μ^B, hence $G^A = G^B$. A G transformation on the probability in eq. (2.2) would take a form

$$G\mathrm{P}(x_\nu^A, x_\nu^B) = Gf(x_\nu^A, x_\nu^B)\,\mathrm{d}\mu_{AB} = |\langle V_{AB}|G^\dagger G|\Psi^{8D}\rangle|^2\,,$$

where $\langle V_{AB}|$ is the function with which Ψ can be projected onto the volume μ_{AB}, that can be denoted later as $\langle x_\mu^A| \otimes \langle x_\mu^B|$. The principle of relativity requires that this transformation should not affect the probability, leaving it invariant. This can be formulated as

$$|\langle V_{AB}|G^\dagger G|\Psi^{8D}\rangle|^2 = |\langle V_{AB}|\Psi^{8D}\rangle|^2\,,\tag{2.4}$$

which means G has to be a unitary representation of the group, because its hermitian conjugate has to be its inverse, $G^\dagger = G^{-1}$. There is another way how an observer could determine if there is a special reference frame, namely if it only changes the experiment in one of the points, and leaves the other untouched. This is equivalent of a group action on one of the points:

$$G_{B\text{replace}} = 1 \otimes G_B\,.$$

Supposing that the observer has the freedom to set up the experiment, this replacement should provide the same probabilistic result as replacing the point A with $G_{A\text{replace}} = G_B^{-1} \otimes 1$, because a $G = G_B^{-1} \otimes G_B$ transformation of the reference frame provides the same $\langle V_{AB}|$. Functionally this means that the following should hold:

$$|\langle V_{AB}|1 \otimes G_r^{-1}|\Psi^{8D}\rangle|^2 = |\langle V_{AB}|G_r \otimes 1|\Psi^{8D}\rangle|^2\,,$$

which has to be true for an arbitrary G_r replacement and V_{AB} volume. That can only be satisfied if the Ψ^{8D} has the following property for any G_r:

$$G_r \otimes 1|\Psi^{8D}\rangle - 1 \otimes G_r^{-1}|\Psi^{8D}\rangle = 0\,.\tag{2.5}$$

Just like in eq. (2.1), $\Psi^{8D}(x_\mu^A, x_\mu^B) = \Psi^{4D}(x_\mu^A - x_\mu^B)$ is also a solution, however it is not general. Namely Ψ^{8D} is a tensor with two spinor indices, generally $\Psi^{8D} = \sum_i c_i \Psi_{Ai}^{4D} \otimes \Psi_{Bi}^{4D}$ with c_i constants. Therefore eq. (2.5) only requires that $\sum_i c_i G\Psi_{Ai}^{4D}$ invokes $\sum_i c_i G^{-1}\Psi_{Bi}^{4D}$, hence there must be a mutual dependence between $\Psi^{4D}(x_\mu^A)$ and $\Psi^{4D}(x_\mu^B)$. In other words, there must be a linear or non-linear operator M for which $\Psi_B^{4D} = M[\Psi_A^{4D}]$, and its inverse must exist, so $\Psi_A^{4D} = M^{-1}[\Psi_B^{4D}]$ can be expressed. A possible solution to satisfy eq.(2.5) is to require M to be invariant under the transformation with G_{x_A} acting on the coordinate x_A and $G_{x_B}^{-1}$ on x_B with a translation in the opposite direction but with the same amplitude:

$$M[\Psi_A^{4D}] = G_{x_B}^{-1}M[G_{x_A}\Psi_A^{4D}]\tag{2.6}$$

or equivalently, its inverse should satisfy

$$M^{-1}[\Psi_B^{4D}] = G_{x_A}^{-1}M^{-1}[G_{x_B}\Psi_B^{4D}]$$

In case M is a linear operator, what was learned from eq. (2.1) can be applied directly, meaning that M must be of the form $M(x_\mu^A - x_\mu^B)$ and $\Psi_B^{4D}(x_B) = \int M(x_\mu^A - x_\mu^B)\Psi_A^{4D}(x_A)\,\mathrm{d}x_A$. A non-linear M could

be described by perturbation theory, for which it can be proved that it must depend only on $x_B - x_A$. A second order M would look like

$$\Psi_B^{4D}(x_B) = M_0 + \int M_1(x_\mu^A - x_\mu^B)\Psi_A^{4D}(x_A)\,\mathrm{d}x_A + \iint M_2(x_B - x_A', x_B - x_A'')\Psi_A^{4D}(x_A')\Psi_A^{4D}(x_A'')\,\mathrm{d}x_A'\,\mathrm{d}x_A''\,.$$

Using an M that satisfies eq. (2.6), the transformation of a $\Psi^{8D} = \Psi_A^{4D} \otimes \Psi_B^{4D}$ will satisfy eq.(2.5), since the following two transformations generate the same Ψ^{8D}:

$$G_{x_A} \cdot \Psi_A^{4D} \otimes \Psi_B^{4D} = G_{x_A}\Psi_A^{4D} \otimes M[G_{x_A}\Psi_A^{4D}] = G_{x_A}\Psi_A^{4D} \otimes G_{x_B}^{-1}\Psi_B^{4D}\,,$$

$$G_{x_B}^{-1} \cdot \Psi_A^{4D} \otimes \Psi_B^{4D} = M^{-1}[G_{x_B}^{-1}\Psi_B^{4D}] \otimes G_{x_B}^{-1}\Psi_B^{4D} = G_{x_A}\Psi_A^{4D} \otimes G_{x_B}^{-1}\Psi_B^{4D}\,.$$

Examining the special case when there are no c_i factors, Ψ^{8D} is a diadic tensor, the criteria that the double-time slices should provide the full probability looks like the following:

$$1 = \underbrace{\iint_{\mathbb{R}^6} |\underbrace{M[\Psi_A] \otimes \Psi_A}_{M_i[\Psi_A]\cdot\Psi_{Aj}}|^2 \,\mathrm{d}^3 x_A\,\mathrm{d}^3 x_B}\Bigg|_{\substack{x_0^A=t^A \\ x_0^B=t^B}} = \underbrace{\int_{\mathbb{R}^3} |M[\Psi_A]|^2 \,\mathrm{d}\mu_B}_{c_B(t^B)} \underbrace{\int_{\mathbb{R}^3} |\Psi_A|^2 \,\mathrm{d}\mu_A}_{c_A(t^A)}\,,$$

where result of the integrals c_A and c_B may depend on t^A and t^B, but their product should be 1 for any choice. This leaves the only possibility of time-independent constants, where $c_A = 1/c_B$. The different c_A values are not observable, hence it is an equality class. Without breaking generality, both can be chosen to be unity, with the requirement that any time-slice integral of $|\Psi_A|^2$ and $|M[\Psi_A]|^2$ must be 1. As it was observed in eq. (2.3), the time dependence of Ψ^{8D} can be explicitly written into a unitary operator. That means that $M[\Psi_A]$ might be expressed with a functional of a chosen time-slice of Ψ_A, and the time-dependent unitary operator. In the case where $\Psi^{3D}(x_j^A)|_{t^A=\text{const.}}$ this may look like the following:

$$\Psi^{8D} = \sum_i c_i U(t^A - t^B)M^{3D}[\Psi_{Ai}^{3D}] \otimes \Psi_{Ai}^{3D}\,.$$

This still has a tensor product that may mix the Ψ_A part with Ψ_B, causing that the outcome of the measurement in x_μ^A may depend on the measurement in x_μ^B. There might be arguments that an effective arrow of time may emerge even from such a system, but this is also an observation of the physical world, so the feature that measurement A should have no dependence on measurement B if $t_A < t_B$ must be present. This can be built in if the M^{3D} operator is linear, so every c_i constant can be suppressed into Ψ_A^{3D}.

For the case where the measurement device is the same for A and B, it can be required that for $t_A = t_B$ $\Psi_A^{4D} \equiv \Psi_B^{4D}$ to provide the same probabilities for the same time slices. This means that up to an unknown phase $U(0)$ and M^{3D} can be chosen to be 1. In this case if Ψ_A^{3D} was known for t_A, the time dependence of the measurement can be expressed as

$$\Psi_B^{4D}(x_\mu^B) = U(t_A - t_B)\Psi_A^{3D}(x_i^A)\,,$$

where the U unitary operator has no dependence neither on x_i^A nor on x_i^B, so it must commute with the space-like, and because of the $t_A - t_B$ structure it must also commute with the time-like group operators. Such a unitary operator can be expressed with a constant Hermetian operator H as $U(\Delta t) = \exp(iH\Delta t)$,

H has to commute with any group operation. Hence for infinitesimal time shifts

$$i\partial_t \Psi^{4D} = H\Psi^{4D} \,.$$

The result is the time evolution equation known in quantum field theory. A further requirement on *H* is that it has to be positive semi-definite, so there is a global ground state that can be called vacuum. A constraint on Ψ comes from the spin-statistics theorem that the Ψ function belonging to integer spin representations that is considered to be indistinguishable should be symmetric, while half spin wave-functions should be antisymmetric. This should be there to preserve the order of time, and prevent information travelling faster than light. Measurements on systems represented by a Ψ_{sys} should also follow the physical constraints that were derived. A measurement apparatus is also a physical system that may be represented by a wave-function. In case the measurement process disturbs the observed system only minimally, than it can be modelled as a weak interaction between the two by a given Hamiltonian operator. The result is that after the interaction the wave functions of the measurement device and the observed system becomes entangled [6][7]. This can be modelled as an action of a self-adjoint operator on Ψ_{sys}. The interaction will turn the system into one of the eigenvectors of the operator, with the probability of the overlap between the eigenvector and the system's wave-function.

2.3 Free particles and interactions

The wave-functions that belong to the irreducible representations of the group may be called particles, as these are the simplest objects that can be distinguished. When their time evolution is only determined by the irreducible group's time translation, it may be called a free particle. For the Poincaré group, the time evolution of a half-spin free particle is governed by the Dirac equation:

$$\left(-i\gamma^\mu \partial_\mu + m\right)\psi(x_\mu) = 0 \,,$$

with the spin-indexed gamma matrices. Interactions with other particles are usually introduced via gauge theories, when a known global symmetry is taken to be locally true. Generally, a symmetry between a set of particles appears as a $M_{ab}(\theta_l)$ group operation with θ_l parameters, and mixes the states $\psi_a, a \in \{1..N\}$, while leaving the Lagrangian invariant.

$$
\begin{aligned}
L_{\text{free}} &= \sum_a \bar\psi_a (i\gamma^\mu \partial_\mu - m)\psi_a \\
&= \sum_{abc} \bar\psi_c M_{cb}^{-1} (i\gamma^\mu \partial_\mu - m)M_{ab}\psi_a \,.
\end{aligned}
$$

Since a local $M_{ab}(\theta(x_\mu))$ transformation does not commute with $i\partial_\mu$, the free Lagrangian is not invariant:

$$L_{\text{free}}^{\text{tr.}} = \sum_{abc} \bar\psi_c M_{cb}^{-1} (i\gamma^\mu \partial_\mu)M_{ab}\psi_a - \sum_a m\bar\psi_a \psi_a \,.$$

Hence an interaction with an $A_{\mu ab}$ external, so called gauge field, with certain transformation properties has to be introduced to compensate the effect:

$$L_{\text{int}} = \sum_a \bar\psi_a i\gamma^\mu g A_\mu \psi_a \,,$$

with

$$A^{\text{tr.}}_{\mu ab} = M_{ac}A_{\mu cd}M^{-1}_{db} - \frac{1}{g}(\partial_\mu M_{ac})M^{-1}_{cb}\,.$$

The interaction Lagrangian transforms similarly as the free Lagrangian, and as a simplification a covariant derivative can be defined:

$$D_\mu = \partial_\mu - igA_\mu\,.$$

The introduction of the new field alone does not introduce interaction, the free particles will follow the same path. Hence it is just a re-parametrisation in this form. However, additional degrees of freedom can appear while introducing further transformation invariant terms. One such term is the kinetic energy of the new field:

$$L_{\text{gauge}} = \frac{1}{2}\text{tr}F^{\mu\nu}F_{\mu\nu}\,,$$

where

$$F = \frac{1}{ig}[D_\mu, D_\nu]\,.$$

Without this term, the new field would not propagate and would have no physical effect. It is important to note that the mass term of the A field is not gauge invariant:

$$L_{\text{gauge mass}} = \frac{1}{2}m_A^2 A^2\,.$$

Nevertheless such a term can appear in effective field theories, breaking gauge invariance at certain energy levels or phase spaces. This can happen through another type of interaction by the Yukawa coupling. Strictly speaking the gauge interaction is a procedure where a change in the current of a particle radiates a boson, while in the Yukawa interaction the annihilation of one or more particle creates new types of particles. One example is the coupling of a scalar boson ϕ to a fermionic field ψ with a coupling constant g_y:

$$L_{\text{Yukawa}} = -g_y\bar\psi\phi\psi\,.$$

As Yukawa interactions are not constrained by symmetries like the invariance under gauge transformations, a large variety of additional terms can be considered. This is often denoted with a general $V(\phi)$ potential.

These equations describe microscopic, fundamental interactions that happen at small time scales. On longer time scales and distances more of these can happen, which may not be resolvable by an observer. The procedure of summing up all these possibilities is called renormalisation. This will hide the real history of events happening on the macroscopic timescale and gives rise to an effective theory with many possible interactions and coupling constants. However, a class of the renormalisable interactions do exist, which keep their original form during the renormalisation procedure. Compared to these interactions, the non-renormalisable theories require infinitely many number of coupling constants, but these will be suppressed by the renormalisation, making them invisible at all scales [8]. Renormalisation is not without effects. The most striking one is the energy dependence of the renormalised coupling constants, or its other name the running coupling. The infinite number of parameters appearing for non-renormalisable theories makes it necessary that in practice renormalisability is required for an acceptable theory.

2.4 Symmetries of the world

Discovering the symmetries in the outcome of physical experiments helped physicists to deduce the Standard Model of particle physics. The local-space time symmetry, which is described by the Poincaré group largely contributes to this by giving the constraint of momentum and angular momentum conservation, the existence of the spin and the requirement for causality. This latter one gives rise to the spin-statistics theorem that explains the Pauli principle, the anticommuting behaviour of half- spin particles and the commutation of particles with integer spin.

It has to be noted that the Lorentz and Poincaré symmetries are not global, as they are not true on cosmic scales. Instead the Einstein-Hilbert symmetries are more accurate, explaining the gravitational interaction classically with the curvature of space time by the energy-momentum tensor. Astonishingly this symmetry is not renormalisable, giving way to speculations on how to introduce this interaction into quantum field theories consistently.

What are the known microscopic symmetries? The fundamental ones are the charge conservation laws of the electroweak and the strong interaction. The electromagnetic interaction is the only one that can be observed in the classical limit, but unifies with the weak interaction at high energies. The weak interaction, first observed in β-decays, has a low probability to occur at low energies, because of the heavy vector bosons involved. Up to now the strong interaction was the hardest to unlock. Although it is behind many symmetries visible between flavours of mesons and hadrons, but due to being strong, it is also screened. This screening makes it impossible to see the strong charges, the so-called colours on larger distances.

2.4.1 Quantum electrodynamics

The electromagnetic interaction was the first to receive a quantum field theoretical explanation. Although it is a renormalisable theory, now it is thought to be the low energetic effective theory of the electroweak interaction. It is a $U(1)$ gauge symmetry of charged fields, coming from the phase of a complex field. The requirement of local invariance under gauge transformations gives place for a single massless, spin one vector boson field: the photon. The Lagrangian of the interaction between a fermionic field ψ and the photon A is the following:

$$L_{U(1)} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} ,$$

with

$$D_\mu = \partial_\mu - ieA_\mu ,$$

$$F_{\mu\nu} = \frac{1}{ie}[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu .$$

The recipe to calculate probabilities is the following. First the operator of the action must be defined, between two time slices:

$$S = \int_{t_0}^{t_1} L \, \mathrm{d}t .$$

S is needed in order to calculate the time evolution operator $U(\Delta t)$:

$$U(t_0 - t_1) = T e^{iS(t_0, t_1)} \,,$$

where the T operator means time ordering. The probability of a transition between state A at time t_0 to state B at t_1 is

$$\mathrm{P}(A \to B) = |\langle \psi_B^{t_1} | U(t_0 - t_1) | \psi_A^{t_0} \rangle|^2 \,. \tag{2.7}$$

This looks simple, but is rather hard to calculate as the operator inside the bracket contains the information for all the possible transitions, it is called the S matrix. It is often approximated in the limit of $t_0 \to -\infty$ and $t_1 \to \infty$. To calculate the transition probability for specific processes, it is better to do it within the framework of perturbation theory. As the electron charge can be considered small, the Taylor series of eq. (2.7) in powers of e gives one simple way to calculate the scattering cross sections perturbatively. The first order only contributes to a single interaction, for example

$$\mathrm{P}(\overrightarrow{k_{e^-}} + \overrightarrow{k_{e^+}} \to \overrightarrow{k}_\gamma) = |\langle A_{k_\gamma} | \int \bar{\psi} \gamma^\mu e A_\mu \psi \, \mathrm{d}t | \psi_{k_{e^-}} \bar{\psi}_{k_{e^+}} \rangle|^2 \,, \tag{2.8}$$

shows the probability that an electron and a positron with momentum $\overrightarrow{k_{e^-}}$ and $\overrightarrow{k_{e^+}}$ annihilate into a photon with momentum $\overrightarrow{k}_\gamma$. This is now much easier to calculate, as the commutation relation and integral of single operators come from first principles. Equation (2.8) is the leading term of the unperturbed probability in eq. (2.7) and consequently it is easier to calculate. The integral corresponds to a simple Feynman diagram, the electromagnetic interaction vertex. The full cross section contains all possible graphs with the same configuration of outgoing lines. To take all these into account, the masses and the charges have to be renormalised, as the bare parameters are unmeasurable for a macroscopic observer.

2.4.2 The electroweak interactions

The electroweak interaction is more complex than the simple electromagnetic interaction as its underlying group is $SU(2)_L \times U(1)_Y$. It is parity violating, as it only connects the left-handed fermions and leaves the right-handed helicity untouched. The left- and right-handed part of the fermions are two different massless half-integer spin representations of the Lorentz group. A Dirac fermion contains both, and they can be expressed as a projection:

$$\psi_L = (1 - \gamma_5)\psi \,,$$
$$\psi_R = (1 + \gamma_5)\psi \,.$$

In this interaction, only the left-handed particles form doublets, belonging to the spinor representation of the $SU(2)_L$ group, while right-handed particles are all singlets, and they transform under gauge transformation accordingly:

$$\begin{pmatrix} \psi_L^a \\ \psi_L^b \end{pmatrix}' = T_{1/2} \begin{pmatrix} \psi_L^a \\ \psi_L^b \end{pmatrix},$$
$$\psi_R^{a\,'} = T_0 \psi_R^a \,,$$
$$\psi_R^{b\,'} = T_0 \psi_R^b \,.$$

In the meantime, all these chiral fermions fields are complex, and they belong to a $U(1)_Y$ gauge group as well. This group sets a g_{hc} hypercharge for each chiral particle, making these chiral fermions a

representation of the $SU(2)_L \times U(1)_Y$. The difference in the representation between left and right handed fermions renders a different covariant derivative for them. With the T^a generator matrices of the $SU(2)$ group and the g_{w} weak charge the covariant derivatives can be expressed as

$$D_{L\mu} = \partial_\mu - ig_{\mathrm{hc}} A_\mu - ig_{\mathrm{w}} B_{\mu a} T^a \,,$$

$$D_{R\mu} = \partial_\mu - ig_{\mathrm{hc}} A_\mu \,.$$

Although the transformation of these fields correspond to $SU(2)_L \times U(1)_Y$ transformations, these are not necessarily the ones which have a propagator. Experiments showed that some of the gauge bosons must have an effective mass, which breaks the electroweak symmetry. The gauge symmetry can still exist, if the symmetry breaking is not explicit but a consequence of some other interaction. It can be spontaneous symmetry breaking, by a Higgs-like Yukawa interaction or a dynamic one creating interacting scalar bosons, and so on. There are many scenarios, and the Standard Model incorporates one of the simplest ones, the single scalar Higgs boson model which will be discussed later.

Generally, massive bosons can be modelled by a positive semi-definite matrix M^2, mixing the vector bosons:

$$L_{\mathrm{mass}} = \frac{1}{2} V_a M_{ab}^2 V_b \,,$$

$$V_\mu = \begin{pmatrix} A_\mu \\ B_{\mu 0} \\ B_{\mu 1} \\ B_{\mu 2} \end{pmatrix} \,,$$

where only the mass-eigenstates would be able to propagate. Nature realises this in a way that the diagonal elements of the $SU(2)_L \times U(1)$ field matrix mixes up into a massless and a massive boson, the electromagnetic photon field and the heavy mediator of the weak neutral currents Z. The off-diagonal elements become the two charged, massive $W^\pm$ bosons. The mass term that represents a self-interaction between the A and B_a fields contains only three non-zero parameters instead of the possible 16:

$$W_{ij} = g_{\mathrm{hc}} A \delta_{ij} + g_{\mathrm{w}} B_a T_{ij}^a \,,$$

$$L_{\mathrm{mass}} = m_{\mathrm{EW}}^2 \sum_{j \in \{0,1\}} W_{1j} W_{j1} \,.$$

In this picture, the hypercharge and the weak charge was used to parameterise the mixing of the gauge bosons. As mixing is much simpler than what is generally possible, it is the first hint that the mass term is an effect of additional interactions. Furthermore, a theory with manifestly massive bosons can only work as a low energy effective theory, because at higher energies the higher order corrections will start to dominate and eventually break the unitarity of the scattering matrix and the probabilities will not add up to one. For the effective electroweak theory, the mass scale of the vector bosons is around 100 GeV, and the theory breaks down at TeV scales. Conversely, it is expected that electroweak symmetries restore at these energies and the low energy symmetry breaking gets an explanation. Another feature of the electroweak theory is that since it requires chiral fermions, these fermions have to be massless. An explicit mass term would break the gauge symmetry, but this as well can be got around by a Yukawa coupling with a Higgs boson.

2.4.3 Quantum chromodynamics

The strong interaction produces the large family of observed mesons and baryons. Most of them can be grouped together by the closeness of their masses, which let many scientists to theorise that those consist of different configurations of a few unobservable particles. The theory of strong interaction came as a solution, introducing several of new fermions, the quarks and a new type of charge, the colour. Each of the new quarks comes in three different strong charges, so they are triplets of the theory. The symmetry between these quarks gives place for a $SU(3)$ gauge freedom, whose eight mediating bosons are the gluons. Due to the strength of the coupling, QCD is not perturbative at low energies, and only lattice calculations are predictive at this scale. As the gluons are massless and have colour charge, the self interaction between them is also large, which screens every coloured particle until it seems colourless for the external observer. This feature has another effect, namely that quarks can not be freed, they always remain in a bound state with additional quarks. As a consequence, the hadrons formed consist of seas of quarks with hardly predictable momentum distributions, or parton distribution functions. Interestingly, a large fraction of the hadron's mass comes from the gluons themself, with a positive binding energy. As a consequence, hadrons can be understood as belonging to a different vacuum in QCD.

One of the strongest evidence for QCD being to be the correct theory of the strong interaction is its running coupling constant. Although at high energies QCD is perturbable, it is still necessary to calculate higher order corrections to scattering processes, since they can give a significant contribution. This is not an easy task to do, but fortunately the renormalisation of the parameters is still doable. The energy dependence of the coupling constant gives a unique signature of the underlying symmetry. Although QCD is independent from the electroweak part, there is still a slight influence on each. The quarks form isospin doublets as well, which results for them having an electric charge. Surprisingly, this charge is only a fraction of the electron charge, being $^1/_3$ and $^{-2}/_3$, but still summing up to integer charges within the hadrons. The fractional charge is somewhat the consequence of the full Standard Model being a consistent theory, where anomalies[1] should cancel [9].

2.4.4 The Higgs field

A simple procedure to give mass to gauge bosons is via spontaneous symmetry breaking of scalar fields. The Higgs field is a complex scalar $SU(2)$ doublet. Invariance under gauge transformations is maintained through the left handed covariant derivative in the Klein-Gordon equation:

$$L_{\text{Higgs}} = \frac{1}{2} D_\mu \begin{pmatrix} \phi^{0*} \\ \phi^{1*} \end{pmatrix} D_\mu \begin{pmatrix} \phi^0 \\ \phi^1 \end{pmatrix} + \lambda \left(\begin{pmatrix} \phi^{0*} \\ \phi^{1*} \end{pmatrix} \begin{pmatrix} \phi^0 \\ \phi^1 \end{pmatrix} - \mu^2 \right)^2 .$$

Instead of a mass term, it contains a self-coupling potential. The structure of the potential ensures a degenerate ground state for the Higgs field, changing the roles of the original four degrees of freedom. When perturbed around the ground state, the vacuum, this field has three massless degrees of freedom and one massive, leaving no question that the symmetry of the Lagrangian is broken. The original $O(4)$ symmetry – which is called custodial symmetry – breaks down to an $O(3)$, which is coupled to the $SU(2)$ gauge bosons via the covariant derivative [10][11][12]. At low energies, near to the perturbative regime of the Higgs vacuum, these three scalar fields will become indistinguishable from the longitudinal degrees of freedom, making them look massive.

---

[1] Anomalies may happen during renormalisation. As the integrals may diverge and infinities appear, a regularisation of the integral is needed. But there might be no regularisation scheme which leaves all of the assumed symmetries unbroken, making the full theory non-symmetric.

charge ⟍ gen.		I	II	III
leptons	0 -1	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$
quarks	$2/3$ $-1/3$	$\begin{pmatrix} u \\ d \end{pmatrix}$	$\begin{pmatrix} c \\ s \end{pmatrix}$	$\begin{pmatrix} b \\ t \end{pmatrix}$

Table 2.1: The three generations of fermions in the Standard Model.

Phenomenologically, the massive degree of freedom of the Higgs field can also interact with the massless, chiral fermion fields, mixing them together. This, on low energies would create an apparent mass for them. Similarly to the mass term of the electroweak gauge bosons, the fermion mass term also allows mixing between the three generations of the fermion fields, explaining why the mass eigenstates of the fermions are not the eigenstates of the weak interaction.

2.4.5 Table of particles

The above described interactions are realised by many fundamental particles in nature [13]. The fermions are summarised in table 2.1. The two main categories are the leptons and the quarks, where only the latter play a role in the strong interactions, and it can be thought that every quark belongs to an additional strong triplet. The doublets shown are the weak isospin doublets, the interaction of the partner particles produce $W^\pm$ bosons, while the interaction with their own anti-particle can produce a Z boson. There are three generations of leptons and quarks, with identical charges but with increasing masses. As it was described in the context of the Higgs mechanism, the mass eigenstates do not coincide with the weak eigenstates. The vector bosons of the Standard Model are the massless photon of the electroweak interaction, the massive weak bosons, the charged $W^\pm$ and the neutral Z and finally the eight neutral gluons, the mediating particle of the strong interaction. The reason for the electroweak symmetry breaking is unknown at the moment. The simplest explanation is the Higgs mechanism, which would add only one neutral scalar particle to the picture.

2.4.6 About gravitation

Gravitation is an interaction that is missing from the Standard Model, and there are many reasons for that. The Einstein equations that describe gravity classically, holds some problems when quantised. It is not renormalisable, invalidating it as a quantum field theory. Several approaches have been suggested which circumvent this, slightly modifying the Einstein-Hilbert action, still without experimental proof. One problem is that the energy scale of quantum gravity should be around the Planck mass, 10^{19} GeV, which suggests that most attempts would be unprovable, as their predictions can not be tested in the laboratories. A no-go theorem also exists from Coleman and Mandula [14], saying that the space-time and internal symmetries can only be mixed in the trivial way, leaving no space for interaction. Theories with spontaneously broken symmetries in special ways are allowed, and supersymmetric theories are also exceptions [15] [8]. The most popular theories are loop quantum gravity and superstring theory.

Addressing correlations

Randomness is an inherent feature of quantum mechanics, disguising the structure of Nature from the observer. To obtain information about a system, many experiments have to be repeated to reveal the distribution of the observables. How can this information be used? The simplest thing to do is interpolation in the parameter range that was already scanned, but an extrapolation is never straightforward. This is especially true when such new systems need to be explored which are more complex than the already known ones. To use data for prediction, some modelling must be done.

The two ways of modelling are estimating the underlying probability distributions or creating a simplified model with regression. Though the central limit theorem says that probability distributions can be approximated with measured data with infinitely many sample points, with finite samples there are always components which are free to be chosen, a path for convergence. The same is true for regressions. This chapter investigates a data modelling method based on unparameterised density estimation, suitable to model data with minimum assumptions.

3.1 Supervised and unsupervised machine learning

Machine learning is associated with a broad range of computational methods, where decisions must be done with the help of samples, typically classification. A possible partition is *supervised* and *unsupervised* machine learning. Supervision in these terms means that an algorithm first faces a tagged data sample or training data, then tries to tag new untagged data with it. An example of this is optical character recognition, where various pictures of the alphabet are prepared, learned and this knowledge is used to transliterate texts. In unsupervised learning however, only the criteria of what defines a class is programmed. Such an algorithm finds classes independently of what others define as a class. Typically it searches for pre-defined features like peaks in data, Gaussians, Lorentz, etc... or more complex structures such as the k-means clusters.

Particle physics usually requires classification, as it is the core for further analyses such as cross section determination or mass measurements, but unfolding is also a task for machine learning. Classification is nothing more than estimating the following Bayesian posterior probability:

$$P(A|x) = \frac{P(x|A) \cdot P(A)}{P(x|B) \cdot P(B) + P(x|A) \cdot P(A)}, \qquad (3.1)$$

which gives the probability that a measured quantity x of coming from class A instead of class B,

(a) NN regression

(b) NN output
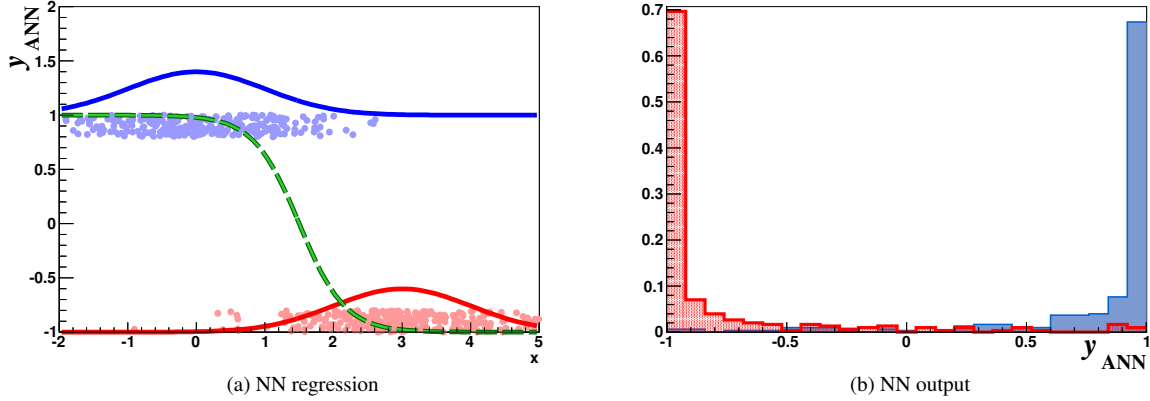
Figure 3.1: An example plot for binary regression. (a) shows two gaussian distributions (red and blue line) and samples drawn from them placed at ±1 targets. The ideal classification curve y_{ANN} what a neural network(NN) would find is illustrated by the dashed green line. (b) shows that if the blue and red distributions were placed to the value predicted by y_{ANN} then they appear as two peaks near the desired target values.

inverting the original condition. An example can be seen on fig. 3.1. However, it can be seen that eq. (3.1) requires the knowledge of the probability that a certain class occurred, P(A) or P(B) which is usually an uncertain quantity, the target of the measurement. The conditional probabilities, P(x|A) and P(x|B) that describe the distribution of the measured quantity x are usually not known analytically, but only a Monte Carlo approximation can be given. Exactly these probabilities need to be modelled by a density estimator, or their ratio, P(x|A) · P(x|B)$^{-1}$, by a binary regression. The following chapter is an extension of the article based on the author's work [1]

3.2 Theory of density estimations

There are numerous techniques to estimate probability densities from a given sample $\{x^i\}$ of measurements. One popular method is histogramming, where the probability density is estimated from the number of events falling into a bin with boundaries $b_j < b_{j+1}$:

$$P(b_j < x < b_{j+1}|A) \approx \frac{|\{x^i|b_j < x^i < b_{j+1}\}|}{|\{x^i\}|} \ .$$

Despite being simple, using histograms has its consequences. First of all, high dimensional histograms are memory-consuming, as the number of bins goes with the power of the number of dimensions. Secondly, histograms with fixed bin sizes come with fixed resolutions, introducing a bias into the density estimation.

The same can be said for fixed kernel methods [17, p182], where the density is estimated with a sum of normalised, non-negative $K(y - x)$ kernels placed into the sample points x^i:

$$g_{\text{est}}(y) = \frac{1}{N_{\text{s}}} \sum_{i}^{N_{\text{s}}} K(y - x^i) \ .$$

The formula is actually an estimation for the following expectation value:

$$g_{\text{est}}(y) = \int K(y - x)g(x)\,\mathrm{d}x\,,$$

where $g(x)$ is the distribution underlying the $\{x^i\}$ sample. It is technically a convolution, and unless the kernel is the δ- function, the g_{est} will be a smeared version of g. Kernels allowing negative values can also be used as a density estimator. However, the treatment of negative densities is problematic, and the overall normalisation has to be solved, typically with a high dimensional volume integral.

Feature extraction is also considered as a kernel method, where the density is estimated with a set of weighted base functions, which are pairwise orthogonal. Because of the orthogonality, the determination of the weights does not require more than an evaluation of scalar products, which are basically estimations of expectation values. As there are infinitely many of these base functions, such a method needs a stopping criteria.

Some adaptivity, or iterative parameter search, is then justified when one wants to estimate densities both without bias and with finite computation time.

3.3 Green's function based density estimation

To construct a density estimator, one can start from linear operators, as they are known to be simple to evaluate. The integral that needs to be calculated can be thought as an expectation value, or a Monte Carlo integral on the acquired sample:

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} Lp = \int L(x, x')p(x')\,\mathrm{d}\eta(x')\,,$$

with a linear operator L, parameter vectors $p(x)$ and probability measure η. A probability density estimator of this form should be then an identity operator. With a post processing operation P, which makes a scalar from the linear operators output this would look like the following:

$$\mu(x) = I(\mu) = P\left(\int L(x, x')p(x')\,\mathrm{d}\eta(x')\right).$$

Identity operators can be constructed from Green's functions of linear differential operators. A simple, and well studied one is the Green's function of the Laplace operator in dimensions $n \geq 3$. With surface of a unit sphere in n-dimensions, S_n, the Green's function G is the following:

$$\Delta G(x, x') = \delta(x - x')\,, \tag{3.2}$$

$$G(x, x') = -\frac{1}{S_n} \frac{1}{|x - x'|^{n-2}}\,.$$

For twice differentiable $c(x)$ functions that disappear at the boundaries:

$$\lim_{x \to \infty} c(x) = 0\,,$$

the following integral identity is then true:

$$c(x) = -\frac{1}{S_n} \int_{\mathbb{R}^n} \frac{\Delta c(x')}{|x - x'|^{n-2}} \, \mathrm{d}^n x' \tag{3.3}$$

$$= \frac{1}{S_n} \int_{\mathbb{R}^n} \partial_{\mu'} \frac{1}{|x - x'|^{n-2}} \cdot \partial_{\mu'} c(x') \, \mathrm{d}^n x' . \tag{3.4}$$

This still contains different derivatives of $c(x)$ on the two sides of the equation, which is not very handy when used for probability densities. An external differentiation makes it an identity for $\partial_\mu c(x)$:

$$\partial_\mu c(x) = \frac{1}{S_n} \int_{\mathbb{R}^n} \partial_\mu \partial_{\mu'} \frac{1}{|x - x'|^{n-2}} \cdot \partial_{\mu'} c(x') \, \mathrm{d}^n x' .$$

The amplitude of $\partial_\mu c(x)$ is still a scalar function, which can be identified with a probability measure $g(x)$. If, for a given $g(x)$ there exists a $c(x)$ function for which

$$|\partial_\mu c(x)| = g(x) \tag{3.5}$$

is true, then with a field of unit vectors $\phi_\mu = c_\mu/|c|$, the following identity must hold:

$$g(x)\phi_\mu(x) = \frac{1}{S_n} \int_{\mathbb{R}^n} \partial_\mu \partial_{\mu'} \frac{1}{|x - x'|^{n-2}} \cdot \phi_{\mu'}(x')g(x') \, \mathrm{d}^n x' . \tag{3.6}$$

A $c(x)$ function that satisfies eq. (3.5) exists for any differentiable $g(x)$, the proof can be found in Section 3.3.1. The particular integral operator in eq. (3.6) was derived from the Laplace operator, and in three dimensions it is similar to the *dipole* interaction operator in electromagnetism. The auxiliary ϕ_μ field can then be called a dipole field, and this acts a as parametrisation of the estimation. It is not known *a priori*, but must be found iteratively by evaluating the left-hand side of eq. (3.6) for a given ϕ_μ field, until a suitable set is found. A possible stopping criteria for the iteration is, when the expectation value $E_\mu(x)$ has the same direction as the $\phi_\mu(x)$ field at every x :

$$\phi_\mu(x) \parallel E_\mu(x) = \frac{1}{S_n} \int_{\mathbb{R}^n} \partial_\mu \partial_{\mu'} \frac{1}{|x - x'|^{n-2}} \cdot \phi_{\mu'}(x')g(x') \, \mathrm{d}^n x' . \tag{3.7}$$

Even without tuning the ϕ_μ parameters, $E_\mu \approx g\phi_\mu$ is true, since they can only differ by a divergent-free vector field. To show this, one must do an integration by parts, so a divergence on $g\phi_\mu$ appears:

$$E_\mu(x) = \frac{1}{S_n} \int_{\mathbb{R}^n} \partial_\mu \partial_{\mu'} \frac{1}{|x - x'|^{n-2}} \cdot \phi_{\mu'}(x')g(x') \, \mathrm{d}^n x' = -\frac{1}{S_n} \int_{\mathbb{R}^n} \partial_\mu \frac{1}{|x - x'|^{n-2}} \cdot \partial_{\mu'} \left(\phi_{\mu'}(x')g(x') \right) \mathrm{d}^n x' .$$

The boundary term is suppressed, since the derivatives of the probability density must also be suppressed to zero at infinity. The divergence of E_μ is an external differentiation that only acts on the interaction term, which becomes a delta function due to the properties of the chosen Green's function:

$$\partial_\mu E_\mu(x) = \int_{\mathbb{R}^n} \underbrace{-\frac{1}{S_n} \partial_\mu \partial_\mu \frac{1}{|x - x'|^{n-2}}}_{\Delta G(x,x')=\delta(x-x')} \cdot \partial_{\mu'} \left(\phi_{\mu'}(x')g(x') \right) \mathrm{d}^n x' = \partial_\mu \left(\phi_\mu(x)g(x) \right) .$$

Consequently, up to a divergence-free field they are equal. The actual difference between E_μ and $g\phi_\mu$ can be modelled by introducing a measure. With the square of the difference, it becomes the following inequality:

$$\int_{\mathbb{R}^n} \left(g(x)\phi_\mu(x) - E_\mu(x)\right)^2 \, \mathrm{d}^n x \geq 0. \tag{3.8}$$

Expanding the square, three terms can be identified:

$$\int_{\mathbb{R}^n} \mathrm{d}^n x \underbrace{g^2}_{\text{invariant}} -2 \underbrace{g\phi_\mu E_\mu}_{\text{dipole energy}} + \underbrace{E^2}_{\text{field energy}} .$$

The dipole energy is an expectation value, but the field energy requires a volume integral. Fortunately, it is strongly related to the dipole energy. In order to show how to simplify it, it must be explicitly written out:

$$\int_{\mathbb{R}^n} \mathrm{d}^n x E^2 = \int_{\mathbb{R}^{2n}} \mathrm{d}^n x \, \mathrm{d}^n y \left[g^x \phi_\mu^x g^y \phi_\nu^y\right] \underbrace{\int_{\mathbb{R}^n} \mathrm{d}^n z \partial_{x_\mu} \partial_{z_\eta} \frac{1}{|x-z|^{n-2}} \partial_{y_\nu} \partial_{z_\eta} \frac{1}{|y-z|^{n-2}}}_{I_{\text{int}}} . \tag{3.9}$$

Although the I_{int} term contains two of the dipole interaction operators, with an integration by parts, a term $\Delta \frac{1}{|x-z|^{n-2}}$, familiar from the definition of the Green's function appears. Because of eq. (3.2), this term can be substituted with a delta function, but still under a partial differentiation. Evaluating this integral gives a simple form for I_{int}, being equal to the dipole interaction operator:

$$I_{\text{int.}} = -\int_{\mathbb{R}^n} \mathrm{d}^n z \, \partial_{x_\mu} \delta(x-z) \partial_{y_\nu} \frac{1}{|y-z|^{n-2}} = \partial_{y_\nu} \partial_{x_\mu} \frac{1}{|x-y|^{n-2}} .$$

Substituting this result back into eq. (3.9) one sees that the field energy equals the dipole energy, and eq. (3.8) can be modified accordingly.

$$\int_{\mathbb{R}^n} \mathrm{d}^n x \, E^2 = \int_{\mathbb{R}^n} \mathrm{d}^n x \, g\phi_\mu E_\mu$$

$$\int_{\mathbb{R}^n} \left(g(x)\phi_\mu(x) - E_\mu(x)\right)^2 \, \mathrm{d}^n x = \int_{\mathbb{R}^n} \mathrm{d}^n x \left(g^2 - g\phi_\mu E_\mu\right) \geq 0. \tag{3.10}$$

This form shows, that in order to minimise the difference between E_μ and $g\phi_\mu$, it is enough to minimise the dipole energy, $\int g\phi_\mu E_\mu \, \mathrm{d}^n x$. Since the ϕ_μ dipole field has a constant length of unity, the criterium for the minima is that ϕ_μ is pointing in the direction of the E_μ field at every x point. This justifies the stopping criteria expressed in eq. (3.7). Although there might be several local minima where the distance between E_μ and $g\phi_\mu$ is not zero in the inequality (3.10), the global minimum (the equivalence) can be reached for differentiable $g(x)$. The proof is shown in the next subsection.

3.3.1 Existence of the optimum

The formula in eq. (3.6) can only be used as a density estimator if a $c(x)$ scalar field exists with the property that the amplitude of the gradient is exactly the density function in question, $|\partial_\mu c(x)| = g(x)$. Although the Green's function identity in eq. (3.3) is true for twice differentiable functions, it is not

trivial what function space the absolute value of the gradient of $c(x)$ covers. It can be formulated in a way that for a $g(x)$ density function one has to find a $\phi_\mu(x)$ field configuration for which the following is true:

$$\partial_\mu c(x) = g(x)\phi_\mu(x) . \tag{3.11}$$

The $c(x)$ function is a potential of $g(x)\phi_\mu(x)$, the line integral of $g(x)\phi_\mu(x)$ has to be the same on every possible path, or equivalently every possible loop integral is zero.

$$\oint_L g(x)\phi_\mu(x)\, dl_\mu = 0 .$$

As it is possible to represent a loop integral with a sum of conjoined loop integrals, it is enough to show that every infinitesimal loop integral can be annulled in the same time, or according to Stoke's theorem the n-dimensional curl is zero. A cube in n dimensions has $n(n-1)$ faces, hence $\frac{n(n-1)}{2}$ independent loop integrals can be imagined for a singe point, and this is the number of independent parameters the curl is going to have. At the same time, the unit-length $\phi_\mu(x)$ field has only $n-1$ independent parameters for every x, which makes it seem non-trivial to solve. The curl in question is the following:

$$\left[\operatorname{curl} g(x)\phi_\mu(x)\right]_{\nu\eta} = \partial_{(\nu)}(g\phi_\mu)e_\mu^{(\nu)} - \partial_{(\eta)}(g\phi_\mu)e_\mu^{(\eta)} \overset{?}{=} 0 , \tag{3.12}$$

denoting the unit basis vector in the ν direction with $e_\mu^{(\nu)}$, and the bracketed indices are *not* summed up. It can be simplified by choosing the basis to be $e_\mu^{(\nu)} = \delta_{\mu(\nu)}$:

$$\left[\operatorname{curl} f(x)\phi_\mu(x)\right]_{\nu\eta} = \partial_{(\nu)}(g\phi_{(\nu)}) - \partial_{(\eta)}(g\phi_{(\eta)}) \overset{?}{=} 0 .$$

In this expression, not all derivatives of ϕ_μ appear; the ν component needs to be derived only in the ν direction. Because of this, there are enough degrees of freedom to set the remaining components of $\partial_\nu \phi_\mu$ to a value which annulls all components of the curl. One possibility to do this is to zero out the following subcomponents while solving the differential equation for ϕ_μ:

$$\partial_{(\nu)}(g\phi_{(\nu)}) = 0 ,$$
$$\frac{\partial_{(\nu)}\phi_{(\nu)}}{\phi_{(\nu)}} = -\frac{\partial_{(\nu)}g}{g} . \tag{3.13}$$

This must be solved with the constraint $\phi_\mu\phi^\mu = 1$. The particular amplitude of ϕ_μ does not matter, since it appears in a fraction. The question is, if ϕ_μ can be *rotated* along with the change of $\partial_\nu g/g$ while moving in the ν direction. This can be done until ϕ_μ points in the ν direction, where a solution breaks down if $-\partial_\nu g/g$ is still non-zero. Nevertheless, it can be solved for a differentiable $g(x)$, but the solution should start from the points where $|\partial_\nu g/g|$ is the largest.

For a $g(x)$ which is non-differentiable, $\partial_\nu g$ would contain δ-functions, and the solution for ϕ should reflect these. The problem there is that ϕ_μ should take a different value when approaching to such a discontinuity from different directions:

$$\lim_{\epsilon\to 0}\phi_\mu(x-\epsilon) \neq \lim_{\epsilon\to 0}\phi_\mu(x+\epsilon) ,$$

making it impossible to solve eq. (3.13) with a single ϕ_μ field for all the possible discontinuities. To show this, let's consider the exception when $g(x)$ is homogenous spherical surface:

$$g(x) = \frac{1}{S_n} \delta(|x| - 1) . \tag{3.14}$$

In this case two possible solutions exist for $c(x)$ to produce $|\partial_\mu c(x)| = g(x)$. The first case is

$$c(x) = \begin{cases} 0 & \text{if } |x| > 1 \\ S_n^{-1} & \text{if } |x| \le 1 \end{cases} ,$$

while in the second case $c(x)$ is non-zero outside the sphere. In these solutions the ϕ_μ dipole field is only relevant where $g(x)$ is non-zero, on the surface of the unit sphere, and should point inwards to or outward from the centre. In accordance with the law of Gauss, such a $g\phi_\mu$ dipole field would produce no E_μ field inside and outside the sphere, but only on the surface, which models the $g(x)$ density function well. Any inhomogeneity in the density on the spherical surface would break such a cancellation of the fields, invalidating the model. Something similar happens when the density in question is simply a δ-function, which can be modelled as a limit of a spherical surface being shrunk to zero radius. The different direction of ϕ_μ at different points of the surface can be maintained as long as the radius is finite; when it reaches zero, the solution for ϕ_μ becomes ambiguous.

Although there might be solutions for eq. (3.11), or more importantly to the identity (3.6) outside those described by eq. (3.13), it is safe to say that the density estimate is valid for differentiable density functions.

3.3.2 Applicability of the method on a finite sample

As long as a ϕ_μ field was found that satisfies the stopping criteria in eq. (3.7), identity (3.6) can be used for density estimations for any $\{x_i\}$ sample:

$$g(x) = \lim_{N \to \infty} \left| \frac{1}{NS_n} \sum_{i=1}^{N} \partial_\mu \partial_\nu \frac{1}{|x - x_i|^{n-2}} \cdot \phi_\nu(x_i) \right| . \tag{3.15}$$

On a finite sample however, the $1/r$ type singularity in the Green's function may cause unnecessarily large fluctuations. This originates from the low number of sample points around the point x where one is interested in the density, which results in uncertainly estimating the integral in this layer. However, synchronous with the assumption that the underlying $g(x)$ probability density is differentiable, one can assume that in this small sphere where the fluctuations come from, the $g(x)\phi_\mu(x) = g\phi_\mu$ field is constant and homotropic. For such a spherical surface the relevant integral in the estimator gives no contribution to the centre at x:

$$\int_{S_n} \partial_\mu \partial_{\mu'} \frac{1}{|x - x'|^{n-2}} g\phi_{\mu'} \, \mathrm{d}^n x' = g\phi_{\mu'} \partial_\mu \int_{S_n} \partial_{\mu'} \frac{1}{|x - x'|^{n-2}} \, \mathrm{d}^n x'$$

$$= g\phi_\nu \partial_\mu \int_{S_n} \underbrace{- \frac{1}{|x - x'|^{n-1}}}_{const.} \hat{r}_{\mu'} \, \mathrm{d}S_n \, \mathrm{d}r' = 0 .$$

$$\underbrace{\qquad\qquad\qquad\qquad}_{\int_{S_n} \hat{r}'_\mu \, \mathrm{d}S_n = 0}$$

This means that a small ΔR radius sphere can be excluded from the calculation in eq. (3.15), since its contribution will go with $\partial_\mu(g(x)\phi_\nu(x))O(\Delta R)$ for any continuous $g(x)$ and $\phi(x)$. The only exception where the error is not $O(\Delta R)$ is for δ functions, so these points have to be treated independently.

To have the fluctuations of the density estimation under control, the ΔR exclusion radius has to be chosen in such a way that the sample points around this sphere still provides a meaningful estimate for the integral of the area they represent. This requires that the first δr thick layer around this sphere, for which the values between $\{\frac{1}{\Delta R^n}, \frac{1}{(\Delta R + \delta r)^n}\}$ are still in the same order of magnitude, must contains enough number of sample points for the integral estimation. The larger number of points means less uncertainty, but it requires larger ΔR exclusion as well, which also assumes homogeneity on $g(x)$ and hence worsening the resolution.

The δr thickness of the layer around the excluded sphere can be determined from the behaviour of the series of $(1 + {}^1/_n)^{-n}$, which quickly saturates around the value of ${}^1/_2$ as a function of n. In the same manner, the values coming from the sample points in the $\delta r = \Delta R/n$ thick layer are all in the same order of magnitude. The volume of this layer can be related to the volume of the sphere:

$$V_{\text{layer}} \approx S_n \Delta R^{n-1} \delta r = \underbrace{V_n \Delta R^n}_{V_{\text{sphere}}} \frac{\delta r}{n \Delta R} = V_{\text{sphere}} \frac{1}{n^2} \ .$$

In this approximation, one needs to choose a ΔR that covers $N_{\text{discr}} = n^2 N_{\text{large}}$ sample points to have N_{large} points in the first stable layer. Figure 3.2a shows an example dipole configuration for a Gaussian distribution. When the smoothing applied the direction of the dipoles that satisfy eq. (3.6) follows the centrally symmetric nature of the distribution and the evaluation gives a rather smooth and density estimation on fig. 3.2b.
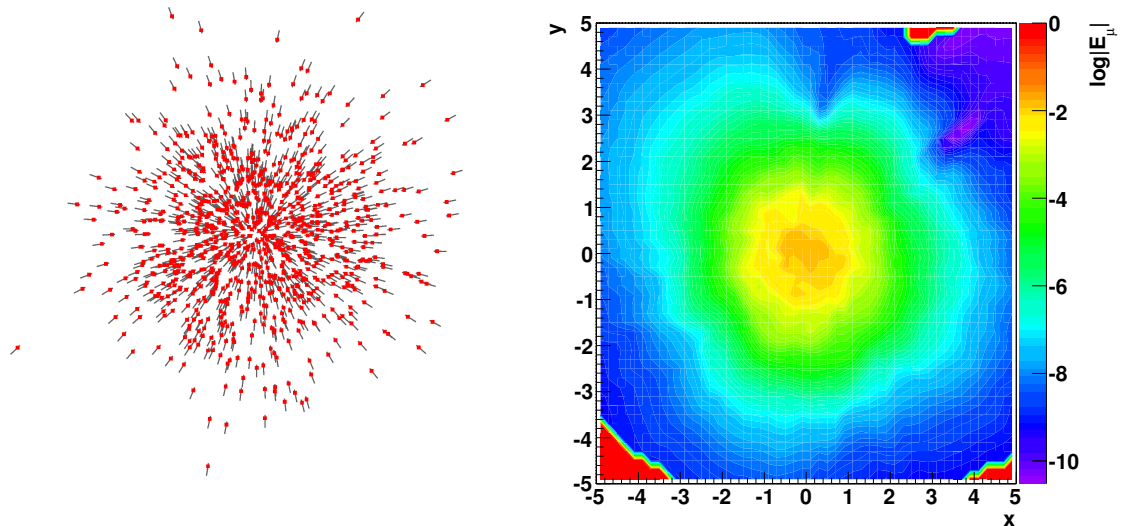
(a) 2000 points (red dots) from a two dimensional Gaussian distribution. The black sticks represent the direction of the dipoles at the sample points, and are shown after the ground state was found.

(b) Contour plot of the logarithm of the estimated densities of a Gaussian distribution.

Figure 3.2: A Gaussian distribution as a sample with the kernel parameterisation, and the estimated density.

3.3.3 Finding the kernel parameterisation

Energy minimisation

It was shown above that the $\phi_\mu(x)$ parametrisation that satisfies eq. (3.6) for a given $g(x)$ density can be thought as a dipole configuration that minimises the energy of the system. However, to use the method for density estimation on a finite sample, some modifications have to be made to eq. (3.15). The $\phi_\mu(x)$ parametrisation can only be defined for a finite set of points, $\{\phi_{\mu i} = \phi_\mu(x_i)\}$, which have to be sampled from the $g(x)$ probability density function that is wished to be estimated. For these $\{x_i\}$ set of points, the exclusion radii R^i_{discr} for a suitable N_{discr} can be defined implicitly as the following:

$$\left| \{k : |x_i - x_k| < R^i_{\text{discr}} \} \right| = N_{\text{discr}} \,.$$

With the unit vectors $\hat{r}^{ij}_\mu$ that point from x^i_μ to x^j_μ, the dipole interaction operator becomes the $D_{ij\mu\nu}$ interaction matrix:

$$\hat{r}^{ij}_\mu = \frac{x_{\mu i} - x_{\mu j}}{|x_i - x_j|} \,,$$

$$D_{ij\mu\nu} = \begin{cases} \partial_\mu \partial'_\nu \frac{1}{|x_i - x'_j|^{n-2}} = (n-2) \frac{n \hat{r}^{ij}_\mu \hat{r}^{ij}_\nu - \delta_{ij}}{|x_i - x_j|^n} & \text{if } |x_i - x_j| > R^i_{\text{discr}} \\ 0 & \text{if } |x_i - x_j| < R^i_{\text{discr}} \end{cases} \,, \tag{3.16}$$

with which the associated energy of the dipole system can be written as a simple matrix equation

$$U = -\frac{1}{2NS_n} \sum_{ij} \phi_{i\mu} D_{ij\mu\nu} \phi_{j\nu} \,. \tag{3.17}$$

The stopping criteria in eq. (3.7) can still be applied, requiring for a good set of $\{\phi_{\mu i}\}$ to be parallel and unidirectional with the E^i_ν field it generates:

$$\phi_{\mu i} \parallel E_\mu(x_i) = \frac{1}{NS_n} \sum_j D_{ij\mu\nu} \phi_{\nu j} \,. \tag{3.18}$$

Similarly to the continuous version, this can be found by minimising the dipole energy, for which one possible strategy is the steepest descent method. The gradient of the energy function, taking into account the $\phi_{\mu i} \phi^\mu_i = 1$ constraint is

$$\mathrm{d}\phi_{i\eta} = -\frac{\partial U}{\partial \{\phi_j\}} = (1 - \phi_{i\eta}\phi_{i\mu})E_{i\mu} = (1 - \phi_{i\eta}\phi_{i\mu}) \frac{1}{NS_n} \sum_{ij} D_{ij\mu\nu} \phi_{j\nu} \,, \tag{3.19}$$

which is eventually a rotation on the $\phi_{\mu i}$ dipoles until their directions match the directions of $E_{\mu i}$. The steepest descent method typically requires a line search to assure not to overshoot the minima, but it is not necessary here. The minima will necessarily have the parallel property shown in eq. (3.18), so it is enough to check the angle between $\phi_{\mu i}$ and $E_{\mu i}$ during the iteration, and choose the maximal rotation angle not to overshoot the parallel phase.

One problem still needs to be addressed though, since $E_{\mu i}$ changes every time $\phi_{\mu i}$ changed. The answer is that $E_{\mu i}$ rotates slower than $\phi_{\mu i}$, since $E_{\mu i}$ is a collective effect of many dipoles. The width of distribution of the E_μ strengths is very probably non-zero, therefore the dipoles must be rotated with different speeds, hence $E_{\mu i}$ can not rotate everywhere with the same speed. Parallelism would be set first for the dipoles s where $E_{\mu s}$ is the strongest, and since every other place has smaller $E_{\mu j}|_{j \neq s}$ field,

all $\phi_{\mu j}\,|_{j \neq s}$ rotate more slowly than $\phi_{\mu s}$, producing an $E_{\mu s}$ that rotates slower than $\phi_{\mu s}$. This behaviour ensures that the correct ϕ_μ direction is set gradually from the higher density regions toward the less dense ones. Still, a normalisation for the $\mathrm{d}\phi_\mu$ gradient-step in eq. (3.19) has to be chosen. If one wants to iterate the dipoles in the canonical way,

$$\phi_{\mu i}^{n+1} = \phi_{\mu i}^{n} + \lambda_n \, \mathrm{d}\phi_{\mu i}\,,$$

then it must be ensured that it is still a rotation, namely that the rotation angle α_i is small enough that the following approximation holds:

$$\phi_{\mu i}^{n} \cdot \lambda \, \mathrm{d}\phi_{\mu i} = \cos \alpha_i \approx 1 - \alpha_i^2\,.$$

This can be achieved if the rotation angles are $0 \leq \alpha_i \leq 0.1$. The regularisation parameter, λ_n, must be chosen in a way not to exceed this value.

Although following the gradient is successive, it may take too many iterations to reach the ground state. This is because the density underlying the sample may have a long tail, where the density is exponentially small, consequently the E_μ field is weak there causing very small rotation speeds. To avoid this, it is useful to introduce a different type of iteration a few times after a longer period of steepest descent steps. To make sure to shoot approximately into the good direction, one can use the direction defined by $\mathrm{d}\phi_{\mu i} = (1 - \phi_{\mu i}\phi_{\nu i})E_{\nu i}$, but with a random α_i amplitude independently set at each point between $[0, 1]$, still taking care not to overshoot the parallel phase. An example for this implementation can be seen on fig. 3.4a.
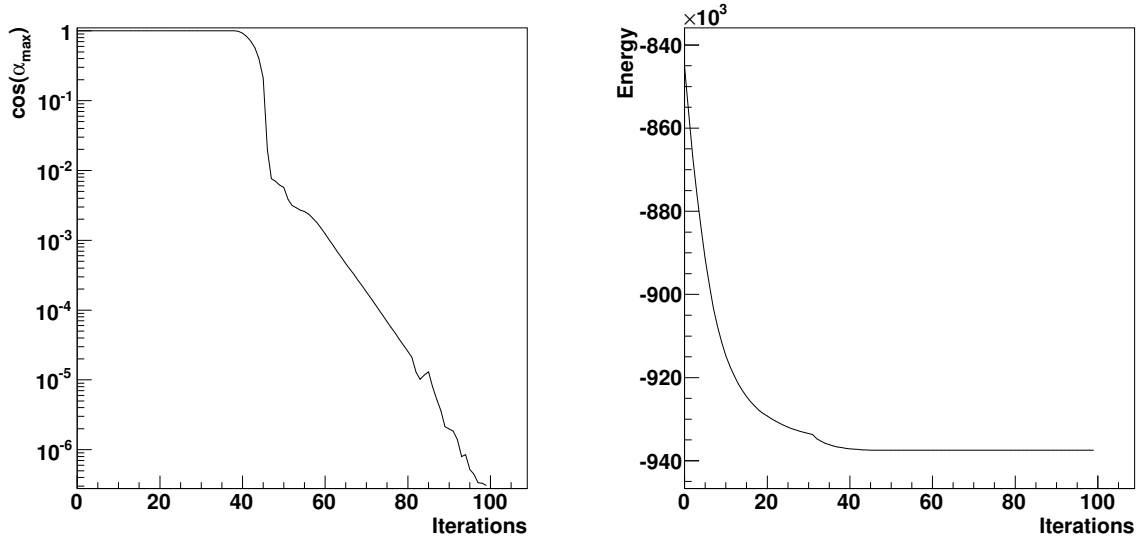
(a) Rapid decrease of the maximal angle between a ϕ_μ dipole and the E_μ field at the same point during the iterations for a sample with 2000 points from a Gaussian distribution.

(b) An energy minimum is found by the algorithm

Figure 3.3: The behaviour of fitness parameters during the energy minimasation.

Fix point search

An alternative method is based on the assumption that the solution is a fixed point. Although this is only approximately true, as there is more than one solution because of the parity symmetry of the dipoles, fast convergence can still be achieved. When eq. (3.6) is formulated with variable dipole amplitude, the additional degree of freedom allows the system to escape saddle points during the minimisation and the lack of constraints avoids the introduction of local minima. One formulation is

$$v_{\mu i}^{\text{out}} = \frac{1}{NS_n} \sum_j D_{ij\mu\nu} v_{\nu j}^{\text{in}} \, .$$

The v^{out} terms can be used as the direction for the next iteration of v^{in}, while the new amplitude of v^{in} must be relaxed towards $1/v^{\text{out}}$. When the iteration stabilises, $v_\mu^{\text{out}} = f v_\mu^{\text{in}}$ and consequently $|v^{\text{out}}| = \sqrt{f}$.

3.3.4 Evaluation

Centrally symmetric configurations

Once the stopping criteria is fulfilled in eq. (3.18), the $\{\phi_{\mu i}\}$ can be used for density estimation at any x coordinate, with a customised exclusion distance:

$$|\{k : |x - x_k| < R_{\text{discr}}(x)\}| = N_{\text{discr}} \, . \tag{3.20}$$

The calculation of the E_μ field then goes as

$$E_\mu(x) = \frac{1}{NS_n} \sum_{j \in \{j : |x - x_j| > R_{\text{discr}}(x)\}} \partial_\mu \partial_{\mu'} \frac{1}{|x - x_j'|} \phi_{\mu' j} \, , \tag{3.21}$$

and its absolute value approximates the density

$$g(x) \approx |E_\mu(x)| \, .$$

It was shown at the end of Section 3.3.1 that due to the law of Gauss a homogenous spherical surface, where the dipoles point either inwards to or outwards from the centre, gives zero field inside and outside the sphere. However it has a contribution on the surface, as it can be imagined that this imaginary spherical surface cuts the dipoles in half, enclosing all the same-sign charges and leaving out the opposite-sign ones.

This feature introduces a problem into the calculation when the estimation is done on a sample, since it requires certain regularisations to avoid fluctuations. When a certain area is excluded from the calculation where $g(x)\phi_\mu(x)$ is not a constant and resembles more and more a centrally symmetric one, it can happen that exactly the excluded area contains those the dipoles which would have given contribution to the density estimation eq. (3.21). For such an area, the exclusion radius defined in eq. (3.20) has to be shrunk, in order to include the necessary number of dipoles for the integral estimation.

To detect spherically symmetric regions, one can evaluate the following function within an exclusion radius:

$$S(x) = \sum_{i \in \{i : |x - x_i| < R_{\text{discr}}(x)\}} \frac{x_\mu - x_{i\mu}}{|x - x_i|} \phi_{i\mu} \frac{1}{N_{\text{discr}}} \, . \tag{3.22}$$

In this formula, the dipoles within the exclusion radius are compared to the unit vector pointing to

the centre of the excluded area. The extrema, $S(x) = \pm 1$, means that the configuration is spherically symmetric, and it is expected to move towards 0 when moving away from the centre; while it should be approximately zero for a homogenous configuration. Near the extrema then, the exclusion radius needs to be shrunken approximately to the radius of the spherical surface in question. Since the contribution in this case comes only from the surface of the sphere, which has one lower number of dimensions than the sample, a certain decrease of N_{discr} can be tolerated. The result of this algorithm can be seen on fig. 3.4a for the same two-dimensional Gaussian sample that was presented in fig. 3.2. The centrally symmetric dipole configuration in the origin would result in an bad density estimation from the applied exclusion, but is prevented by the shrinkage of the exclusion radius.
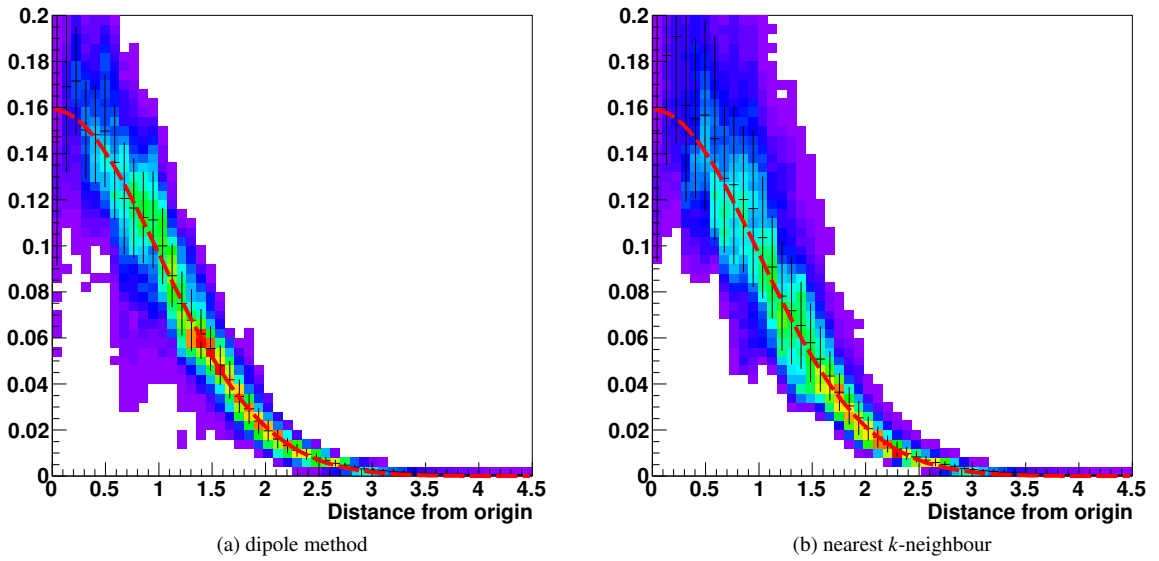

(a) dipole method

(b) nearest k-neighbour

Figure 3.4: The distribution of the estimated density versus the distance from the mean for a Gaussian distribution, made with (a) the dipole method and (b) the nearest k-neighbour method.

Low density areas

As it may happen that the stopping criteria in eq. (3.18) only leads to a local minimum, it is possible that the density model is either not normalised to unity or has a significant error according to eq. (3.8). The requirement for a perfect solution for the $\phi_\mu(x)$ field is that $g(x)\phi_\mu(x)$ is curl-free, so one can imagine the non-ideal solution as a curled field added to the ideal one. A prototype of a curled field may contain only one dipole, while everywhere else the field is zero. A loop integral on a path which passes along the dipoles while returns on a space with zero field would have no contributions to cancel the effect of the dipole, hence this configuration has a non-zero curl. This means that the curled fields can be thought of as dipole fields that were not cancelled out perfectly, as something that has a remnant polarisation. Therefore whenever the configuration is not in the global energy minimum, it can be detected by looking for sparse E_μ fields in very low, or preferably zero density areas. Furthermore, this also means that the relative error is expected to be bigger in the low density regions, when the dipole configuration is not in an ideal state. Figure 3.5 demonstrates how the relative error may increase in low density areas for a Gaussian distribution. An example for a non-ground state configuration can be seen on fig. 3.6.
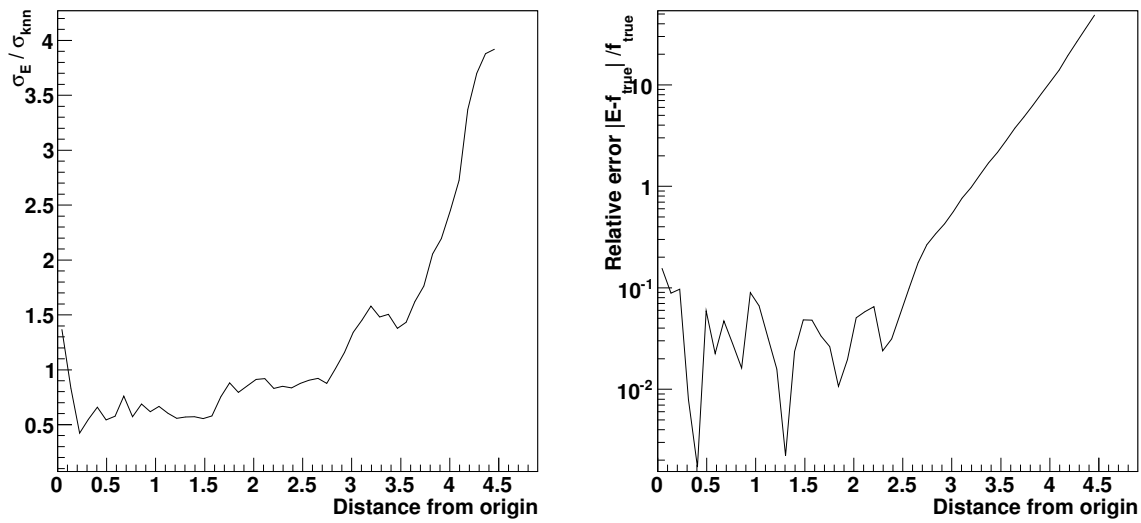
(a) Spread of the dipole-based estimator over the spread of the *k*-nearest neighbour estimator versus the radius

(b) Relative error of the average of the dipole-based density estimate compared to the true value.

Figure 3.5: Comparison of the spread of the density estimation with the *k*-nearest neighbour method and the deviation from the true value.
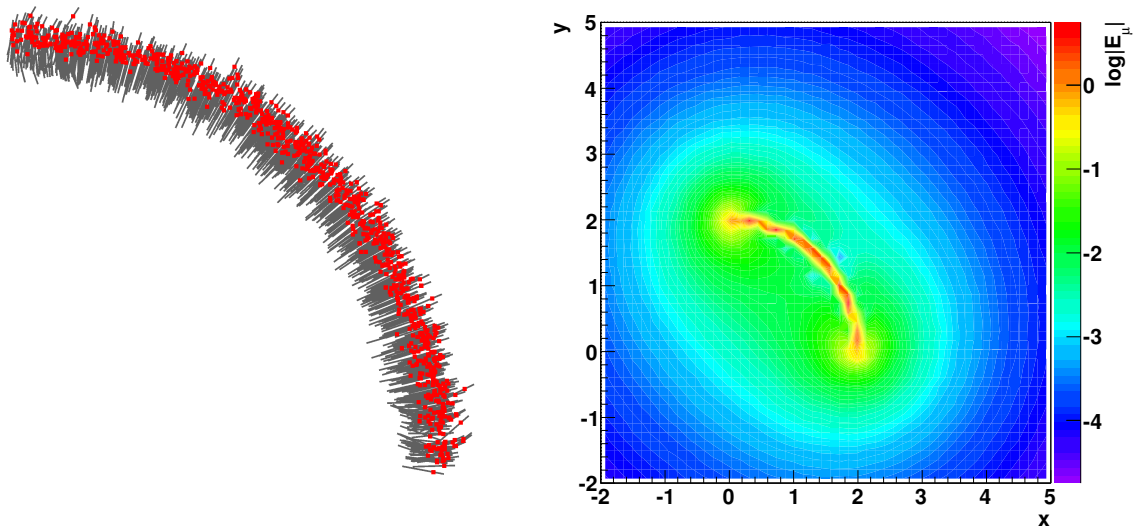
(a) Sample of a heavily bent distribution (red dots) and its kernel parametrisation (black rods).

(b) Logarithm of the density estimate of the arc distribution. A halo is clearly present towards the centre of the arc.

Figure 3.6: Example of a heavily curved and thin distribution. The distribution is too thin for the regularisation and the parametrisation probably does not correspond to the real ground state.

To make the decision that there is remnant polarisation in a low density area, one needs a different density estimator first in order to detect low density regions. Since the algorithm already requires finding the R_{discr} exclusion radius for a sphere with N_{discr} points inside, it is economical to use a *k*-nearest

neighbour algorithm. When one is interested in the density in a rarely populated region, it can be checked if the closest N_{large} data points are in the shell with radii between $[r_{\text{closest}}, r_{\text{closest}}(1+{}^1\!/_n)]$. As the volume of the sphere with radius r_{closest} and the shell are approximately the same, this would mean that at the point of interest the density is significantly smaller than in the shell. Figure 3.7 shows a uniform distribution in a rectangle, where the boundaries and edges represent a challenge to the dipole-based density estimator. The density estimation decreases only with a power law around the density-steps, but not so far from the edge the detection of low density areas turns on.

(a) A uniform distribution within a rectangle. The red dots are sample points, while the black rods represent the kernel parametrisation.

(b) The logarithm of the estimated density of a uniform distribution within $x, y \in [-0.5, 0.5]$. The estimation clearly drops down around the edges, as expected, but shows a halo nearby and cuts off the corners.
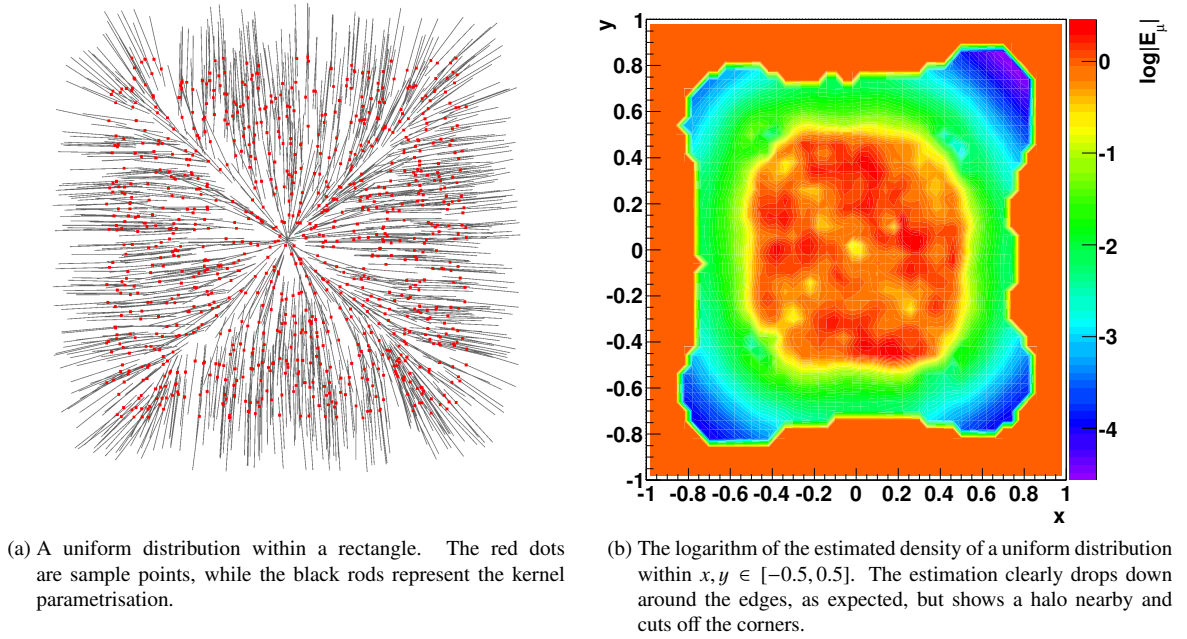
Figure 3.7: Parametrisation and density estimation of a distribution with edges and corners.

3.3.5 Initialisation and preprocessing

The solution for the $\phi_\mu(x)$ field for a spherically symmetric $g(x)$ distribution is known analytically, since it is a superposition of the solutions for homogenous spherical shells. As it was shown for eq. (3.14), the solution is a spherically symmetric ϕ_μ configuration, where all the dipoles point either inwards to or outward from the mean of the distribution. It is possible then start the iterative search of the $\phi_{\mu i}$ field from the assumption that the distribution is spherically symmetric, by setting every $\phi_{\mu i}$ dipole pointing into the centre of mass of the sample.

As a preprocessing, the sample can be made more spherically symmetric with simple and computationally cheap transformations. One of the simplest is linearly rescaling the individual x_μ coordinates of the sample in a way that the empirical width become the same for all:

$$\sigma_{(\mu)}^2 = \langle x_{(\mu)}^2 \rangle - \langle x_{(\mu)} \rangle^2 ,$$

$$x_{(\mu)}'^i = x_{(\mu)}^i \frac{\sigma_{(0)}}{\sigma_{(\mu)}} .$$

Such a step is also required because the regularisation required for samples, as described in Section 3.3.2, does not differentiate between the coordinates for the calculation of the distances. Because of this, it can happen that for the density estimation certain coordinates are not taken into account at all, and the estimation becomes insensitive for these directions. An extreme example for this is a uniform distribution in a two-dimensional rectangular region, where one of the sides is much narrower than the other. Depending on the number of points, the exclusion radius can be bigger than the smaller side of the rectangle. Artificially rescaling the smaller side would enhance the resolution in that direction while worsening it in the others.

3.4 Classification

A density estimator can be used to estimate the likelihood of a given sample point belonging to a class, hence it can be used for classification as it was described for (3.1). Following Bayes' rule, given the distribution of a signal $s(x) = P(x|s)$ and a background $b(x) = P(x|b)$ with the marginal probabilities that they occur $\sigma_s = P(s)$ and $\sigma_b = P(b)$, the posterior probability that a given point x can be a signal event is the simple fraction:

$$P(x \text{ is signal}) = P(x|s) = \frac{s(x)\sigma_s}{s(x)\sigma_s + b(x)\sigma_b} \; . \tag{3.23}$$

Several signals and backgrounds might be combined. With distributions $s^i(x)$, $b^j(x)$ and occurrences σ_s^i and σ_b^j, a combined signal and background definition is sufficient:

$$1 = \sum_i \sigma_s^i + \sum_j \sigma_b^j \, ,$$

$$\sigma_s = \sum_i \sigma_s^i \, , \; \sigma_b = \sum_i \sigma_b^i \, ,$$

$$s(x) = \sum_i s^i(x) \frac{\sigma_s^i}{\sigma_s} \, , \; b(x) = \sum_i b^i(x) \frac{\sigma_b^i}{\sigma_b} \; .$$

The posterior probability, eq (3.23) is not an ideal classifier though, since it contains the σ_s and σ_b occurrences which are related to cross sections in particle physics. These are typically uncertain or unknown quantities, and therefore it is better to construct classifiers without using them. One possibility is to set the occurrences to be equal, $\sigma_s = \sigma_b = {}^1/_2$, only for the response, and measure the real occurrences from the data.

$$r_{\text{eq}}(x) = \frac{s(x)}{s(x) + b(x)} \; .$$

Just like the posterior probability $P(s|x)$ in eq. (3.23), this response function is also between 0 and 1, but takes these extrema when either the signal or the background is pure:

$$0 \le r_{\text{eq}}(x) \le 1$$

$$r_{\text{eq}}(x) = \begin{cases} 0 & \text{if } s(x) = 0 \\ 1 & \text{if } b(x) = 0 \end{cases}$$

The response r_{eq} is a probability variable, a $\mathbb{R}^n \to \mathbb{R}$ type function of x. The probability that its value

is R can be expressed with an integral over the probability density of x, generally $P(x)$:

$$P(r_{\text{eq}} = R) = \int P(r_{\text{ex}} = R|x)P(x)\,dx\,.\tag{3.24}$$

The response distribution of the real data $d(x)$ might be different from the $r_{\text{eq}}(x)$ distribution from the simulation, where $\sigma_s = \sigma_b = {}^1/_2$ was true, since data is a superposition of both signal and background with the real occurrences:

$$P(x) = \begin{cases} \frac{1}{2}[s(x) + b(x)] & \text{for simulation} \\ d(x) = \sigma_s s(x) + \sigma_b b(x) & \text{for data} \end{cases}$$

The only dependence of the occurrences is in the $P(x)$ probability distribution, hence it is enough to calculate the response for $s(x)$ and $b(x)$ independently, then scale it with the correct σ_s and σ_b when they are determined. The $P(r_{\text{eq}}(x) = R|x)$ probability, since $r_{\text{eq}}(x)$ is a direct function of x, is just a δ-function where the equality holds:

$$P(r_{\text{eq}}(x) = R|x) = \delta(r_{\text{eq}}(x) - R)\,.$$

In other words, the probability in eq. (3.24) that $r_{\text{eq}} = R$ for a certain sample is nothing more than the integral over a contour with the measure $P(x)$. This contour can be simply expressed as

$$R = r_{\text{eq}}(x) = \frac{s(x)}{s(x) + b(x)} = \frac{1}{\frac{s(x)}{b(x)} + 1}\,,$$

$$\frac{s(x)}{b(x)} = \frac{1}{R} - 1\,,\tag{3.25}$$

which means that a specific response R selects a certain area where $s(x)/b(x)$ is constant. Equation (3.23) expresses the real probability that a certain x data point belongs to the signal class. When this $P(x|s)$ is used as a response or discriminator, a certain $P = P(x|s)$ will select contours where the following is true:

$$P = P(x \text{ is signal}) = P(s|x) = \frac{s(x)\sigma_s}{s(x)\sigma_s + b(x)\sigma_b}\,,$$

and therefore

$$\frac{\sigma_s}{\sigma_b}\frac{s(x)}{b(x)} = \frac{1}{P} - 1\,.\tag{3.26}$$

Since σ_s and σ_b do not depend on x, the contour selected by R in eq. (3.25) can be identified with a contour *labeled* with P in eq. (3.26). The two functions are both monotonic in P and R, which means that selecting events with the $r_{\text{eq}}(x)$ response or with the $P(x|s)$ is equivalent, and a selected contour would enclose the same number of signal and background events in both case. As a consequence, the efficiency of signal selection as a function of the background rejection is the same, when the selection is done with either R or P. The transformation between R and P is simple:

$$P(s|x) = P = \frac{1}{\frac{\sigma_s}{\sigma_b}(\frac{1}{R} - 1) + 1}$$

In this sense every response which only depends only on the fraction of the signal and background density $s(x)/b(x)$ is equivalent, since the signal likelihood can be restored. Figure 3.8 shows the classification capability of the dipole-based density estimator on comparison with a regression method. The
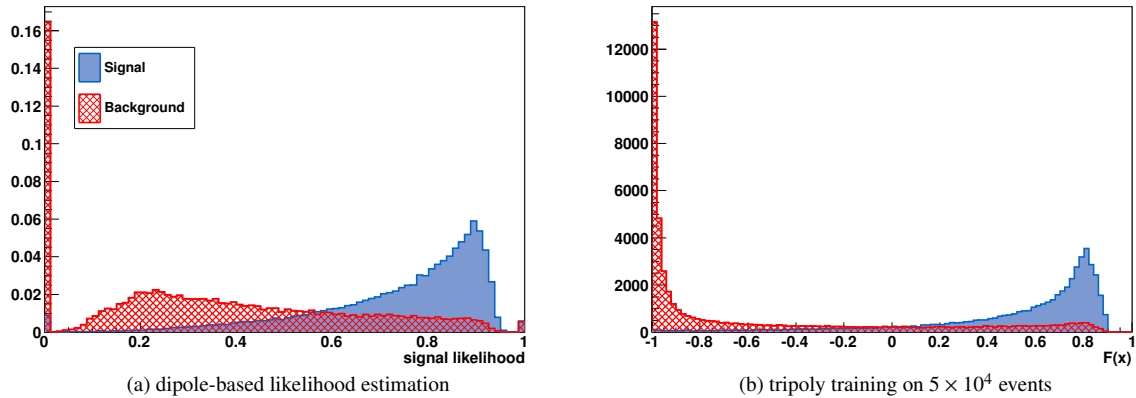
(a) dipole-based likelihood estimation (b) tripoly training on 5×10^4 events

Figure 3.8: Comparision of the dipole based density estimator, being used to estimate signal likelihood with the polynomial expansion of the theoretically best separation contours. The training input is twelve Gaussian peaks as a signal with a flat background. Although the separation with the kernel method is fairly good, it is clear that it does not behave well around the exponentially small Gaussian tails.

distribution of two responses are very similar, the main difference comes from the low density Gaussian tails, where the dipole-based density estimation is expected to be most uncertain.

3.5 Implementation

The implementation of the density estimator algorithm was done in the C++ language using the ROOT framework [18]. It provides the following features:

ROOT file handling

– reads, writes and merges the dipole containers in files;

– converts existing ROOT trees into the dipole format.

Create toy Monte Carlo samples
Several distributions can be requested from command line, with custom dimensions and sample size:

– Gaussian;

– uniform;

– sum of three Gaussians;

– smeared arc distribution, to represent non-linear correlations.

Ground state search

– optimised and parallelised iterative parameter search;

– rotates the dipoles according to the energy gradient;

– stops when the maximal angle between a dipole and the E_μ field is smaller than 10^{-6}.

Initialisation

Initialisation is done with the assumption that the distribution is spherically symmetric, no pre-processing is done.

Evaluation

The evaluation is a more sophisticated step than the ground state search:

- takes care of low density areas and centrally symmetric ones;

- N_{large} can be significantly smaller, hence the resolution is better;

- evaluation can be done on a list of sample points read from an external file;

- the result can be stored to and reread from ROOT files;

- weights are taken into account with a regression.

Classification

- cross-evaluates several samples;

- the likelihoods can be stored in or read from a file;

- uses a text file to determine which samples are signal and which are background or data;

- can handle cross sections;

- produces response histograms, efficiency vs. rejection curves.

Parallel processing

- using multi-core architectures via the *pthread* library and semaphores;

- multiple network nodes can be added through ROOT's network stack.

3.6 Overtraining during regressions

Supervised machine learning can mean two tasks. One is parameter learning, regression of a latent variable and predicting it from observables; the other is binomial regression where the target of the learning is a discreet variable. This latter can be regarded as classification. Generally the regression for two random variable-vectors, the observable X and the target Y with a probability distribution $P(x = X \wedge y = Y)$ can be regarded as a function $f(x)$ that predicts the mean of Y from the observed X:

$$f(x) = \mathrm{E}[y|x] = \int y \mathrm{P}(y = Y|x = X) \, \mathrm{d}y \, . \tag{3.27}$$

For a discrete target, where Y can only take the values $\{y_i\}$, eq. (3.27) splits up into a sum of probabilities:

$$f(x) = \sum_i y_i \mathrm{P}(y_i = Y|x = X) \, .$$

Using Bayes' rule to express conditional probabilities from marginal probabilities, this can be expressed with the condition for Y instead of X. For only two possible y_i values, $f(x)$ becomes a simple weighted fraction of y_0 and y_1:

$$\mathrm{P}(y_i = Y|x = X) = \frac{\mathrm{P}(y_i = Y \wedge x = X)}{\mathrm{P}(x = X)} = \frac{\mathrm{P}(x = X|y_i = Y)\mathrm{P}(y_i = Y)}{\sum_j \mathrm{P}(x = X|y_j = Y)\mathrm{P}(y_j = Y)} \, ,$$

$$f(x) = y_0 P(y_0 = Y|x = X) + y_1 P(y_1 = Y|x = X)$$
$$= \frac{y_0 P(x = X|y_0 = Y)P(y_0 = Y) + y_1 P(x = X|y_1 = Y)P(y_1 = Y)}{P(x = X|y_0 = Y)P(y_0 = Y) + P(x = X|y_1 = Y)P(y_1 = Y)}.$$

In this view, $f(x)$ can be expressed with $g_i(x) = P(x = X|y_i = Y)$ probability density functions belonging to y_i and with $\sigma_i = P(y_i = Y)$, the marginal probabilities of the occurrence of y_i. This weighted sum can be further simplified, as it is only a function of $g_0(x)/g_1(x)$ and σ_1/σ_0 :

$$f(x) = \frac{y_0 \frac{g_0(x)}{g_1(x)} \frac{\sigma_0}{\sigma_1} + y_1}{\frac{g_0(x)}{g_1(x)} \frac{\sigma_0}{\sigma_1} + 1}. \tag{3.28}$$

A regression to the mean is typically based on the χ^2 loss function, E_{χ^2}. The sample with input values $\{x_i\}$ and target values $\{y_i\}$ is modelled with a parameterised function $f(x, p)$:

$$E_{\chi^2} = \sum_i (y_i - f(x_i, p))^2. \tag{3.29}$$

The parameterisation p is selected via the minimisation of E_{χ^2}. When the $f(x, p)$ function can parameterise any function, the minimum of E_{χ^2} will successfully approximate eq. (3.27) with enough sample points. This is because if $f(x, p)$ can approximate any function, then there is no constraint between the values for two different points, x_1 and x_2, so $f(x_1, p)$ and $f(x_2, p)$ can be optimised independently. For points $x_i = x_1$ the sum of squares that has to be minimised is the following:

$$E_{\chi^2}^{\text{part}} = \sum_{i \in \{j|x_j = x_1\}} (y_i - f(x_1, p))^2.$$

This can be rearranged after expressing the squares:

$$E_{\chi^2}^{\text{part}} = \sum_{i \in \{j|x_j = x_1\}} y_i^2 - 2y_i f(x_1, p) + f(x_1, p)^2$$

$$= \sum_{i \in \{j|x_j = x_1\}} y_i^2 + 2\left(\sum_{i \in \{j|x_j = x_1\}} y_i\right) f(x_1, p) + N f(x_1, p)^2$$

$$= \sum_{i \in \{j|x_j = x_1\}} y_i^2 + N\left(f(x_1, p) - \frac{1}{N}\sum_{i \in \{j|x_j = x_1\}} y_i\right)^2 - \left(\frac{1}{N}\sum_{i \in \{j|x_j = x_1\}} y_i\right)^2.$$

The only variable, $f(x_1, p)$, appears within a single square, and the minimum is reached when $f(x_1, p)$ equals to the y_i average:

$$f(x_1, p) = \frac{1}{N}\sum_{i \in \{j|x_j = x_1\}} y_i, \text{ when } \min_p E_{\chi^2}^{\text{part}}$$

and

$$N = |\{j|x_j = x_1\}|.$$

The derivation shows that the χ^2 loss function in eq. (3.29) regresses to the mean, as the fitted function

$f(x, p)$ shows the estimated average of the y values distributed by the conditional probability $P(y|x)$. But this is only true if the population of y values for a single x is sufficiently large, so an average can be calculated with sufficient precision. For a finite sample, where the distribution of x is not discrete, this requires either a regularisation in x or a careful analysis of the p parametrisation, as none of the p parameters should depend only on a few number of $\{x_i, y_i\}$ sample point pairs. This is typically circumvented by choosing a class of functions with only a few parameters: low degree polynomial or a multi-layer perceptron neural network with only a few internal nodes. However, this is just a rule of thumb, and introduces otherwise unwanted constraints into the fitting function, $f(x, p)$. The dependence of a parameter on a very few sample points is *overfitting* or *overtraining*. A way to check if this happens usually requires separating the available data into independent training and testing samples, and the procedure of *early stopping* requires minimising the loss function based on the training data, but stopping the iteration when the loss function no longer decreases on the test sample.

The problem of overtraining remains when doing binary classification with a regression. The density estimation technique, based on the dipole ground state search, explicitly requires regularisation and calculates the components for eq. (3.28) independently. Although the parametrisation of the dipole system primarily estimates probability densities of the two class, when combined for eq. (3.28) it also estimates the conditional probability eq. (3.27) just as the regression. One may ask why this parametrisation is less prone to overtraining? How come that for N sample points it can provide $2N(d-1)$ parameters, much more than what is advised to avoid overtraining? It is because the $\phi_\mu(x)$ parametrisation is correlated, the regularisation causes a smoothing in the $\phi_\mu(x)$ field, so the direction of a $\phi_\mu(x_i)$ dipole at the point x_i can be more-or-less predicted from the nearby points. The nearby ϕ_μ directions are based on the same information, with a little reweighting depending on the position. This correlation ensures that every ϕ_μ dipole is supported by several data points, and depends on the distribution itself rather than the sample. Removal or addition any of the sample points would not change any of the parameters significantly. Figure 3.8a was produced with an independent training and testing sample. The central region of the response is well-behaved, it is smooth as expected and signal purity increases in function of the response. Nevertheless, the appearance of the small signal contribution at the very left side and the background events at the very right side can be understood as a slight overtraining. These responses represent very low signal and background regions accordingly, and those were detected and by the auxiliary algorithm for low densities and have the largest uncertainties.

CHAPTER **4**

Principal curves and coordinates

The previous chapter dealt with a supervised machine learning algorithm that is capable of classifying data when the class definitions are already available in a sampled way. Nevertheless, it is technically possible to classify data without predefined classes via unsupervised learning. This typically means clusterisation, but as it was discussed before, physics requires the detection of internal constraints and correlations in measured data, which sometimes requires unsupervised parameterisation above the classification. This chapter aims to reformulate the heuristics in existing algorithms upon a fundamental basis.

4.1 Principal curves

One of the most basic methods to take non-linear correlations into account is the method of principal curves. The goal of the method is to define a curve that goes through the middle of an empirical distribution. A simple definition by Hastie [19] is the folowing:

$$c(p), \text{ which satisfies } \min_p \sum_i \Delta r_i^{\text{closest}}(p) \,, \tag{4.1}$$

where $c(p)$ is a set of curves, parametrized by p, and $\Delta r_i^{\text{closest}}$ is the closest distance of the ith data point to the curve $c(p)$. This definition still has some ambiguity. The parametrisation of all possible curves is not possible without an infinite number of parameters, hence a static parametrisation would not provide the fundamentally best result. With a non-static parametrisation one would need additional optimisation criteria, which are not straightforward to select. One example is that the description of open and closed curves are different due to their different topology, and an automatised choice between the best closed curve and the best open curve can not be easily done. In such a case all possible curves would have to be compared not just the ones that are accessible by the restricted parametrisation. An actual implementation would be then heuristic, rather than fundamental. Another problem is the particular choice of measure of the distance used in the definition. The components of a data event x_μ may come from fundamentally different sources, e.g. measurements made by different detectors. Rescaling one component is then allowed with arbitrary units, which changes the position of the closest approach of a given curve, opening the possibility to provide a different curve after optimisation.

A more fundamental problem is that this definition does not deal with higher dimensional hypersurfaces, but nevertheless tries to regress a one dimensional curve on any input. It may give inappropriate

results for input distributions where the shape is unknown, generally it can not be decided only from the output if the distribution can be simplified into a curve or not.

4.2 Redefinition of principal curves

4.2.1 The dipole approach

Keeping all this in mind, one has a freedom to redefine the principal curve in eq. (4.1). First of all, instead of $\Delta r_i^{\text{closest}}$ a strictly monotonous function of it can be used. For a decreasing function, one has to look for the maximum in p instead of the minimum. This way any non-zero power of $\Delta r_i^{\text{closest}}$ is allowed. This is beneficial, because the negative powers can be approximated with integrals, making the equations easier to handle and avoiding the double minimisation. For a given curve $c(p)$ one can select a parametrisation along the curve, with parameter t, and approximate $\Delta r_i^{\text{closest}}$ with the following integral:

$$(\Delta r_i^{\text{closest}})^{-n} \approx \int_{t_0}^{t_1} \frac{1}{|r_i - c(p,t)|^n} \, dt \, . \tag{4.2}$$

In case there is only one point of closest approach, eq (4.2) is an approximation of $(\Delta r_i^{\text{closest}})^{-n}$ for $n > 1$. This is because the point of closest approach dominates the integral and every point further is suppressed compared to that. The line integral in this definition shares similarities with the work done along a path in the field of a point charge at r_i with potential $1/r^{n+1}$. Such a parametrisation with a $1/r^{n-2}$ potential in n-dimensions would have the benefit that the line integral only depended on the endpoints. This would make the calculation simpler since only the endpoints would have to be varied. Nevertheless, it is possible to modify the definition further and use the same parameters that vary the curve to vary the potential in the same time. The simplest modification to a point charge that is also direction dependent is the infinitesimal dipole. This integral can be reformulated as a path integral that calculates the potential difference of two point charges at the end of the curve. Consider a unit dipole with direction ϕ_i at every data point r_i. The field of these dipoles will define a direction in every r point in the input space:

$$E_\mu(r) = \sum_i \underbrace{\frac{n\hat{r}_\mu \hat{r}_\nu - \delta_{\mu\nu}}{|r_i - r|^n}}_{\text{dipole interaction term}} \phi_{i\nu} \, . \tag{4.3}$$

From a given starting point s a line integral can be performed along these directions, realising a general curve

$$c_\mu(s,t)[r_i, \phi_i] = s_\mu + \int_0^t \hat{E}_\mu(c(s,t')) \, dt' \, .$$

As it was noted about eq. (4.2), the dipole interaction term in eq. (4.3) can be used in the estimation for the closest approach of $c(p)$ to a point r_i:

$$(\Delta r_i^{\text{closest}})^{-n} \approx \int_{t_0}^{t_1} \frac{\partial c_\mu(p,t)}{\partial t} \frac{n\hat{r}_\mu \hat{r}_\nu - \delta_{\mu\nu}}{|r_i - c(p,t)|^n} \phi_{i\nu} \, dt \tag{4.4}$$

Here $\frac{\partial c(s,t)}{\partial t} \, \mathrm{d}t$ had to be used in order to make the integral independent from the t parametrisation. This approximation is only valid when the tangent vector of $c(p)$ nearby the point r_i is parallel to the dipole ϕ_i at this point, otherwise it does not take its extremum near this point, but elsewhere. In the following it will be shown that this is true for the principal curve. As for negative powers of r_i^{closest} the principal curve has to be maximised, its definition becomes the following:

$$c_\mu(s,t)[r_i, \phi_i], \text{ which satisfies } \max_{s,[\phi_i]} \sum_i \int_{t_0}^{t_1} E_\mu(c(s,t)) \frac{n \hat{r}_\mu \hat{r}_\nu - \delta_{\mu\nu}}{|r_i - c(s,t)|^n} \phi_{i\nu} \, \mathrm{d}t \,.$$

It can be simplified as

$$c_\mu(s,t)[r_i, \phi_i], \text{ which satisfies } \max_{s,[\phi_i]} \int_{t_0}^{t_1} E_\mu(c(s,t)) E_\mu(c(s,t)) \, \mathrm{d}t \,.$$

This contains the square of the field strength, but since $E_\mu(c(s,t)) = \dfrac{\partial c_\mu(s,t)}{\partial t}$, and $\left| \dfrac{\partial c(s,t)}{\partial t} \right|$ is the distance on the curve by changing the t parameter by a unit, the overall integral is the potential difference between the endpoints of the curve. An approximation can be introduced after reordering the equation:

$$\int_{t_0}^{t_1} E_\mu(c(s,t)) E_\mu(c(s,t)) \, \mathrm{d}t = \sum_i \phi_{i\mu} \int_L \underbrace{\frac{n \hat{r}_\mu \hat{r}_\nu - \delta_{\mu\nu}}{|r_i - c(s,t)|^n}}_{\partial_\mu \frac{\hat{r}_\nu}{|r_i - c(s,t)|^{n-1}}} \hat{E}_\nu(c(s,t)) \, \mathrm{d}r \,,$$

where a directional derivative appears, which can be integrated out with a line integral along the curve. In case the endpoints are not the same, the solution is that the integral is substituted with the field from a point charge at the endpoints, $\Phi_{t_0\mu}$ and $\Phi_{t_1\mu}$. This can be denoted as

$$\int_{t_0}^{t_1} E_\mu(c(s,t)) E_\mu(c(s,t)) \, \mathrm{d}t = \sum_i \phi_{i\mu} [\Phi_{t_0\mu}(r_i) - \Phi_{t_1\mu}(r_i)] \,.$$

One should maximise this quantity to obtain the parameters for the principal curve. This result can be interpreted as the energy of the $\{\phi_i\}$ dipole system in the field of two point charges at the endpoints of the curve t_0 and t_1. These points are still related to the dipole configuration, but only indirectly. They should be a valid endpoint of curves that the field of the dipoles determine. Around these points the dipoles will be aligned and point directly to the point charges. The explicit way to assure that the dipole configuration in the maximisation process selects these endpoints would be a constrained extremum search. A technically easier way is to approximate the field of the point charges with the dipoles themselves, and then do the maximisation. This can be expressed as the self energy of the dipole system:

$$c(p,t), \text{ where } \max_{[\phi_i]} \sum_{ij} \phi_{i\mu} \frac{n \hat{r}_{ij\mu} \hat{r}_{ij\nu} - \delta_{\mu\nu}}{|r_i - r_j|^n} \phi_{j\nu} \,. \tag{4.5}$$

Performing this calculation is technically less difficult than doing line integrals at every step of the iteration, not to mention that sometimes the lines are infinitely long. The requirement for the ground state of the system ensures that it is stable, and as a consequence the field generated by the dipoles at

the data points are parallel to the dipoles themselves. This is consistent with the definition of $\Delta r^{\text{closest}}$ in eq. (4.2).

This formulation of a principal curve has many benefits:

- selects the best curve from an non-constrained set;

- increasing the number of data points refines the curve;

- the definition does not requires time-consuming calculations of running through curves;

- provides a parametrisation of the data instead of a curve;

- the field lines are defined on the full input space;

- such a parametrisation can be useful on datasets which inherently consist of either several curves or none;

- since the field lines generated by a dipole system are not crossing each other, they can be used as a parameterisation of the data.

However, in the definition of the dipole-based principal curves a technical issue arises with the choice of negative powers of $\Delta r^{\text{closest}}$ when applied together with the maximisation criteria of the sum. The closest possible distance of a curve and a data point is zero, so while $\Delta r^{\text{closest}}$ approaches 0, its negative power $(\Delta r^{\text{closest}})^{-n}$ diverges, making it impossible to perform the calculation. It is then necessary to introduce regularisation into the calculations, to provide smooth curves. To avoid artificial scales, a fixed number of dipoles can be excluded from the calculation around the point for which the E_μ is calculated, a similar technique to the one that was described in section 3.3.2. Such a regularisation would disappear in the continuum limit, which makes it possible to define the principal curve with using a probability density function $f(r)$ instead of a data sample $\{r_{\mu i}\}$. It has the benefit that it is normalised and its properties can be investigated in the analytic form without the statistical noise. Applying the regularisation to eq. (4.5) while denoting the dipoles as a vector field $\phi_\mu(r)$ defined at every space point, the formula becomes the following:

$$c_\mu(p, t), \text{ where } \max_{[\phi_\mu(r)]} \iint_{\mathbb{R}\,\mathbb{R}} f(r)\phi_\mu(r)\frac{n\hat{r}_\mu\hat{r}_\nu - \delta_{\mu\nu}}{|r - r'|^n}f(r')\phi_\nu(r')\,\mathrm{d}^n r\,\mathrm{d}^n r'\,. \qquad (4.6)$$

In this limit the only problem remaining is the infinite dipole self interaction. Nevertheless, the result of the integral is definitely finite, since the $2n$-dimensional integral of a $1/_{r^n}$ singularity is finite. Similar reasoning is true for the line integral on the field, as the field calculation in eq. (4.3) results in a logarithmic singularity, which becomes finite after integration.

4.2.2 Open and closed curves

The curve parametrisation described in the previous section is very similar to the density estimation technique described in Section 3.3. The difference is that the principal curve requires maximisation of the dipole energy, while the density estimation needs minimisation. The two are definitely connected, and it is possible to take advantage of this relation. When the energy is minimised, the curl of the dipole field ϕ_μ is also minimised, as it should be zero in the ground state. This means that the dipoles would not form a closed loop, also not the generated E_μ field, but they should end either at infinity or at places where E_μ is zero. These are typically areas which either have a pole-like spherical symmetry in the ϕ_μ configuration or contains a domain wall.

How to create closed curves from the open curves? It is possible to create quasi-closed loops with curl-free dipole fields, but it requires more than unit dipoles. Figure 4.1a shows an ideal parametrisation for closed loops that do not intersect and act like a phase: an angle that starts from a certain value and ends at another one for every curve around the same axis, $\beta \in [0, 2\pi]$. As it was described for the density estimator, the identity in eq. (3.6) holds for the energy ground state, when the $g(x)\phi_\mu(x)$ field is a gradient of a potential field $c(x)$, so $g(x)\phi_\mu(x) = \partial_\mu c(x)$. Such a potential is similar to a broken helix that only contains one curl and has drop between $2\pi \to 0$, as can be seen on fig. 4.1b. The axis will remain an ambiguous point, but it has a nearly zero measure. The constant drop in $c(x)$ will correspond to a δ-function with homogenous density on a surface that has one end at infinity and one at the axis. Individually neither the continuous part $C_\mu(x)$, nor the δ-function-like dipole-surface $s(x)$ satisfy the Green's function based vector identity, but their sum is a gradient of a scalar $c(x)$ function:

$$\partial_\mu c(x) = C_\mu(x) + \phi_\mu(x)\delta(s(x))2\pi \, .$$

(a) A possible parameterisation of closed curves.

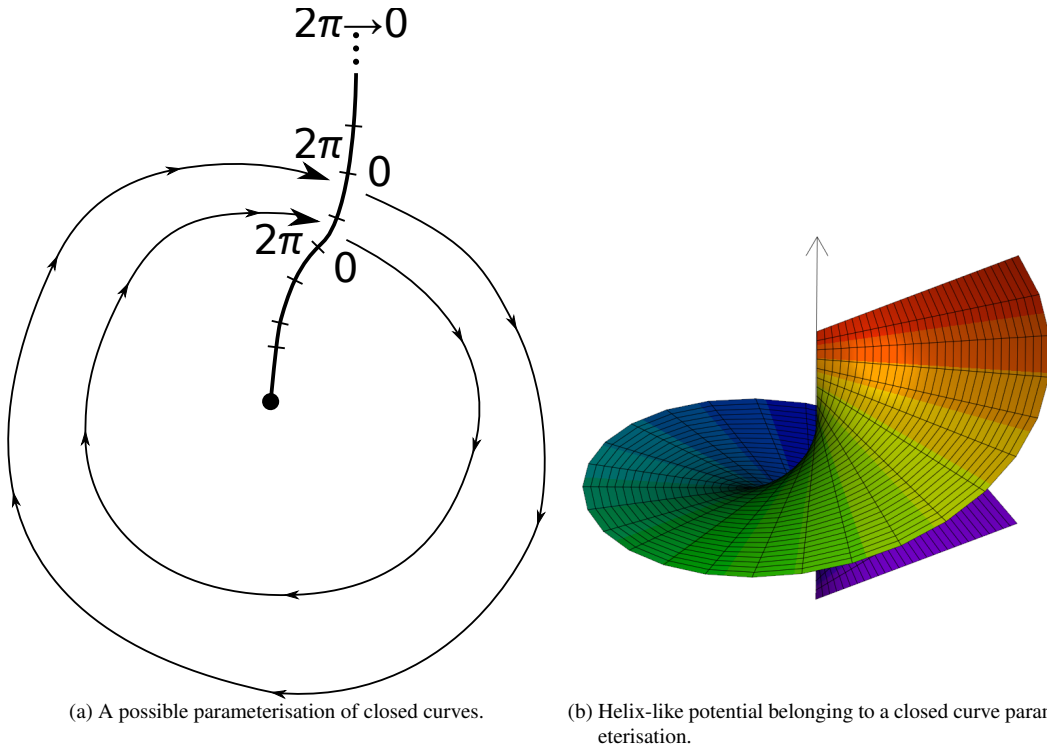(b) Helix-like potential belonging to a closed curve parameterisation.

Figure 4.1: When closed curves are parameterised with a circular phase a surface appear where the parameters promptly change. In the case this parameter is represented by a scalar potential, this sudden change appears as a step function. The potential at the central axis is undefined.

The δ-surface may be rotated anywhere around the axis, but it has to be perpendicular to the incident curves, in order to ensure that the real curves are still closed loops. There can be many δ-surfaces simultaneously, with different weights which add up to 2π, the maximal phase. The distribution of these $s(x)$ surfaces can be done in a way, that its density corresponds to the density of the $C_\mu(x)$ continuous part, since the loop integral of the continuous parts were required to be a constant 2π. Evidently this leads to a paradox, that a rotational potential field can only be constructed with a help of a secondary field that cancels its effect, hence giving zero field-strength. The configuration is very similar as requiring

that the produced field E_μ should be antiparallel to the field ϕ_μ, practically that the system should be in an energy maxima. Overall, this is not a potential field, since the ϕ_μ field is not curl-free at the axis. The reason why maximising the energy of a dipole system still leads to field lines that are looped, and have non-zero strength is subtle. For a finite number of sample points, the dipole interaction operator has to be regularised, implicitly requiring smoothness of the dipole field. This requirement is true for the ground state, but as the example above shows, it is not true for the energy maxima. The used dipole operator,

$$D_{\mu\nu ij} = \begin{cases} \dfrac{n\hat{r}_{ij\mu}\hat{r}_{ij\nu} - \delta_{\mu\nu}}{|r_i - r_j|^n} & \text{if } |r_i - r_j| > \epsilon \\ 0 & \text{if } |r_i - r_j| \le \epsilon \end{cases}, \tag{4.7}$$

needed to be regularised in order to get rid of the singularity. When this matrix is actually derived from the Green's function of the Laplace operator, because of the differentiation applied on the absolute function in the denominator, it has an additional δ-function term:

$$\partial_\mu\partial_{\mu'}\frac{1}{|x - x'|^{n-2}} = (n-2)\frac{n\hat{x}_\mu\hat{x}_{\mu'} - \delta_{\mu\mu'}}{|x - x'|^n} + (n-2)S_n\delta_{\mu\mu'}\delta^n(x - x').$$

This δ-term is also ignored when the function is regularised, and requiring smoothness on a finite sample would implicitly choose something. One of the several possible choices is the mangetic dipole. As fig. 4.2 illustrates, in three dimensions the microscopic magnetic dipole differs from the electric dipole only by the scale of the δ-function:

$$D^{\mathrm{e}}_{\mu\mu'} = \frac{1}{4\pi}\frac{3\hat{x}_\mu\hat{x}_{\mu'} - \delta_{\mu\mu'}}{|x - x'|^3} - \frac{1}{3}\delta_{\mu\mu'}\delta^3(x - x'),$$

$$D^{\mathrm{m}}_{\mu\mu'} = \frac{1}{4\pi}\frac{3\hat{x}_\mu\hat{x}_{\mu'} - \delta_{\mu\mu'}}{|x - x'|^3} + \frac{2}{3}\delta_{\mu\mu'}\delta^3(x - x').$$

The latter equation determines the magnetic field for infinitesimally small, closed loop currents. Although in three dimensions the magnetic dipole parameterisation is the one that is capable of describing closed curves, this can not be extended trivially to higher dimensions. Both the curl and the vector cross product require a n-dimensional vector and a multi-index tensor, a 1-form and an $(n-2)$ form, in order to produce a pseudo-vector $(n-1)$-form. This practically means that several additional fields are required to produce closed curves.

However, the regularised interaction matrix in eq. (4.7), when looking for an energy maxima, can be regarded as an extension of the magnetic dipole operator in higher dimensions, and at least in the 3-dimensional case it provides closed curves. The derivation of the 2-dimensional form is different, due to the logarithmic potential, but leads to a formally similar result. For an example, see fig. 4.3. In higher dimensions though, for distributions that are not inherently curve-like, the method fails to provide the expected result. Figure 4.4 shows that for a three dimensional Gaussian distribution the algorithm fits a curve that is although closed, but incapable of representing the symmetry of the peak. This is because the method inherently looks for thin, curved one-dimensional distribution and interprets anything as a sum of these, leading to the knotted feature.
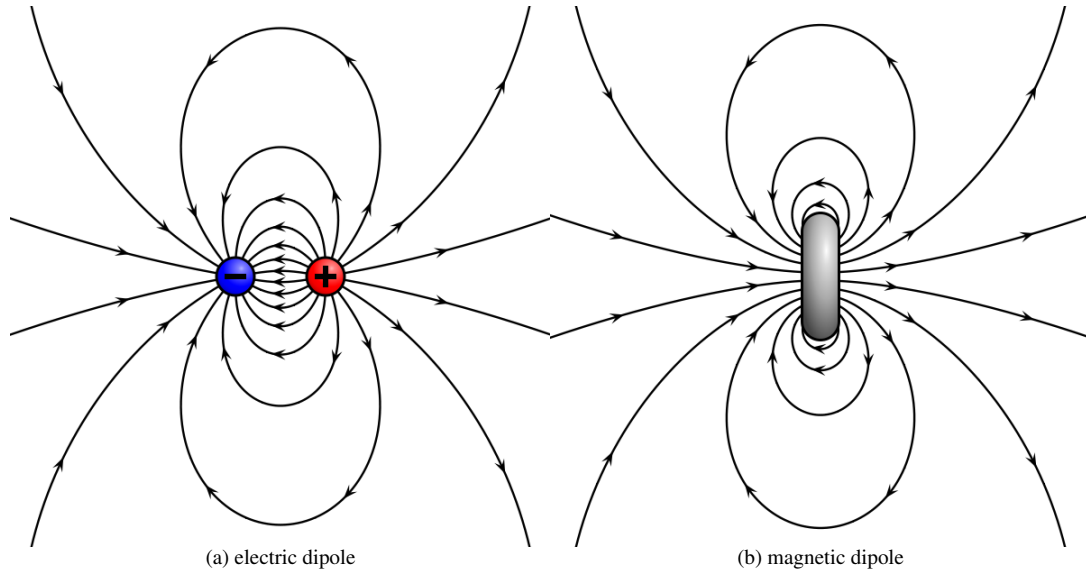
(a) electric dipole

(b) magnetic dipole

Figure 4.2: The far field of an electric and a magnetic dipole is identical, but the microscopic behaviour is different. The electric field becomes divergent nearby the sources while the magnetic field is rotational. The electric field also has a different direction between the two poles, a step in the potential, which is missing from the magnetic field.

(a) A curved distribution example.

(b) Example of a peaked distribution.

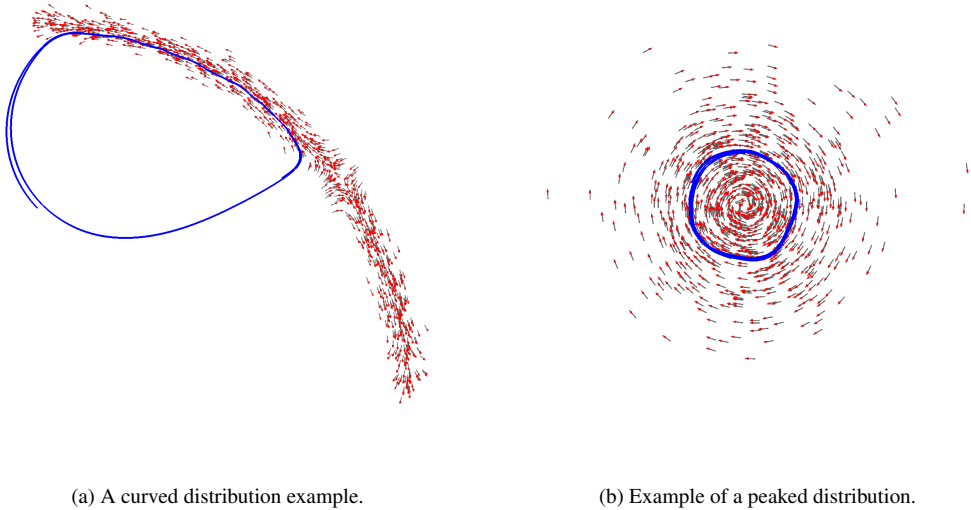Figure 4.3: The dipole method is capable of finding the symmetries of the distribution, providing a parameterisation. (a) shows thin, curve like distribution, where although the dipole system is in a local maximum the field lines can be followed with fourth order Runge-Kutta method, giving approximately closed curves. On a non-curved distribution (b), the method provides a parameterisation for the phase.

(a) front view   (b) side view
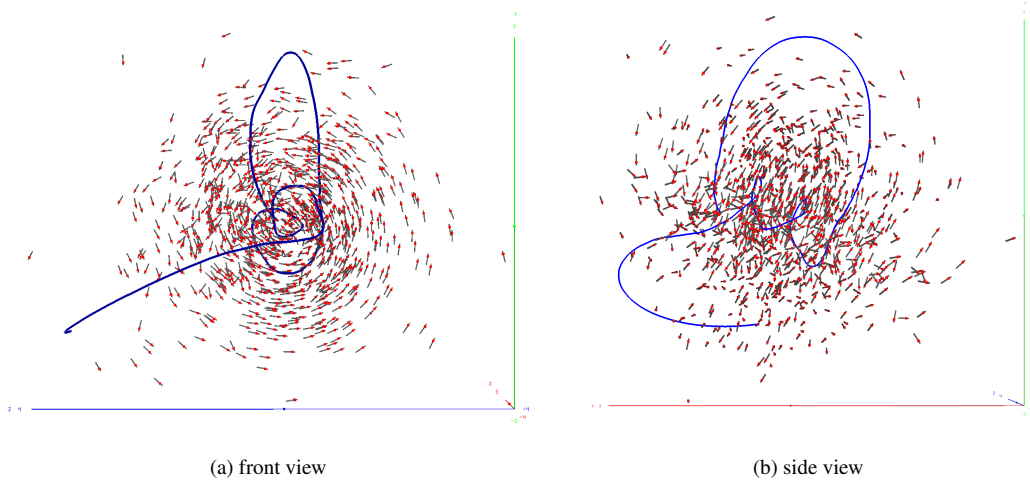
Figure 4.4: The principal curve finding on a non-curved distribution, where the results are hardly interpretable. Although the dipole system is in a energy minimum, it is just partially smooth, as the method tries to force a one-dimensional parameterisation for the distribution. This leads to a tube-like coordinate system tightly packed into the shape of the 3-dimensional distribution.

4.3 Supervised and unsupervised machine learning

As it was described in the previous section, a principal curve finding algorithm may be inadequate for data sets which contain structures other than parametric curves. The algorithm can be prepared though to look for hyper-surfaces independently of their dimensionality. Such an algorithm would create a new mapping of the distribution, providing new coordinates. A good guiding principle can be finding the symmetries of the system, coordinates that show the self-similarity of the distribution. A one dimensional symmetry would show up as a translational invariance of the density, for continuous distributions:

$$f(x(t, u)) = f(x(t + \mathrm{d}t, u)), \tag{4.8}$$

$$\frac{\partial f(x(t, u))}{\partial t} = 0.$$

The same requirement for every other coordinate implies the independence of the coordinates:

$$\frac{\partial^2 f(x(t_0, t_1, ..., t_n))}{\partial t_i \partial t_j} = 0,$$

$$f(x)\,\mathrm{d}^n x = \mathrm{P}(X \in [x_\mu, x_\mu + \mathrm{d}x_\mu)) = \prod_i \mathrm{P}(T_\nu \in [t_i, t_i + \mathrm{d}t_i)) = \prod_i g_i(t)\,\mathrm{d}^n t,$$

where $g(t_i)$ represents the probability distributions along each t_i coordinate. Every non-discrete one-dimensional distribution can be transformed into a uniform distribution, by solving

$$g(t_i)\,\mathrm{d}t_i = \kappa_i\,\mathrm{d}c_i$$

with k_i arbitrary constants. With the help of the independent coordinates $f(x)$ can be transformed into

an n-dimensional uniform distribution:

$$f(x)\, \mathrm{d}^n x = \kappa\, \mathrm{d}^n c\,.$$

Although this equation does not contain the boundary conditions, namely that all c_ν coordinates are finite and must be between $c_\nu \in [C_\nu^0, C_\nu^1]$ with the condition that $\prod_\nu (C_\nu^1 - C_\nu^0) = \frac{1}{\kappa}$ and the boundaries are rectangular $\partial C_\nu^i / \partial c_\mu = 0$, the equation can actually be solved as a partial differential equation:

$$f(x) = \kappa \det\left(\frac{\partial c_\nu}{\partial x_\mu}\right)\,. \tag{4.9}$$

The determinant in eq. (4.9) has $n(n-1)$ degrees of freedom, hence there are numerous ways to satisfy the equation. The constraint from the boundary conditions described above has small importance, since once the differential equation was solved for a certain irregular boundary, it can be later transformed into a rectangle. A meaningful solution would also require that the new coordinate system is smooth enough, and the coordinate lines do not cross each other too often. Such a crossing typically looks like a pole on a map and behaves like a point-like boundary, hence it is beneficial to decrease their number to as low as possible. This requirement is equivalent of requiring all the $\partial c_{(\mu)} / \partial x_\nu$ (for fixed μ) fields to be curl-free at most of the places.

The $\partial c_{(\mu)} / \partial x_\mu$ fields can be *modelled* with the help of dipole fields, where the various ground states can provide smooth and mostly curl-free field lines. The vector identity in eq. (3.6), derived from the Green's function identity being true for the lowest ground state, can be used as a base to create the approximations for the mapping:

$$\partial_\mu c = \frac{1}{S_n} \int_{\mathbb{R}^n} \partial_\mu \partial_\nu \frac{n \hat{r}_\mu \hat{r}_\nu - \delta_{\mu\nu}}{|r - r'|^n} \partial_\nu c\, \mathrm{d}^n r\,.$$

There are several possibilities to introduce the density and the vector fields into the equation. When the $f(x)$ density is not known exactly, and similarly to the cases in the previous sections only a sample is given, a self consistent equation can be created that is true for the ground state. Let's assume that n vector fields satisfy the criteria of being curl-free, with $v_\mu^{(\nu)} = \partial_\mu c^{(\nu)}$ for the coordinates in the ν direction. The $v_\mu^{(\nu)}$ fields can be regarded as the legs of a local, flat coordinate system, the *vielbein*:

$$v_\mu^{(\nu)} = \frac{1}{S_n} \int_{\mathbb{R}^n} \partial_\mu \partial_{\mu'} \frac{n \hat{r}_\mu \hat{r}_{\mu'} - \delta_{\mu\mu'}}{|r - r'|^n} v_{\mu'}^{(\nu)}\, \mathrm{d}^n r'\,.$$

When the coordinates are chosen wisely, and $f(x) = \det(v_\mu^\nu)$ is true, it can be inserted into the equation above:

$$v_\mu^{(\nu)} = \frac{1}{S_n} \int_{\mathbb{R}^n} \partial_\mu \partial_{\mu'} \frac{n \hat{r}_\mu \hat{r}_{\mu'} - \delta_{\mu\mu'}}{|r - r'|^n} \frac{f}{\det v} v_{\mu'}^{(\nu)}\, \mathrm{d}^n r'\,. \tag{4.10}$$

This equation is *self consistent*. It can be true for a certain choice of $v_\mu^{(\nu)}$ fields, but it is not an energy ground state of a dipole system anymore. Nevertheless there is a chance that it can be found as a fixed point of the equation, and can be used on samples after discretisation:

$$v_{i\mu}^{(\nu)} = \frac{1}{NS_n} \sum_{j \in \{k \| r_i - r_k | < R_i\}} \partial_\mu \partial_{\mu'} \frac{n \hat{r}_{ij\mu} \hat{r}_{ij\mu'} - \delta_{\mu\mu'}}{|r_i - r_j|^n} \frac{1}{\det v_j} v_{j\mu'}^{(\nu)} \tag{4.11}$$

To be a real fixed point, it is required that only one solution exists. This is definitely not true, since reversing the direction of any $v_\mu^{(v)}$ field in a solution is also a solution. Nevertheless, a fix point search can be convergent when started nearby a solution and the gradient around the solution is below certain limits.

To investigate if this is the case for the eq. (4.10) and eq. (4.11) one has to analyse them further. The determinant in the denominator can always be expressed as a scalar product of one of the column vectors in the matrix with a $(n-1)$-form pseudovector created from the other columns:

$$
\begin{aligned}
\det v_\eta^\kappa &= v_{\mu_1}^1 \wedge v_{\mu_2}^2 \wedge ... \wedge v_{\mu_n}^n \\
&= v_{\mu_1}^1 \underbrace{\epsilon_{\mu_1 \mu_2 ... \mu_n} v_{\mu_2}^2 ... v_{\mu_n}^n}_{p_{\mu_1}} \\
&= v_{\mu_1}^1 p_{\mu_1}\,,
\end{aligned}
$$

where $\wedge$ symbolises the wedge product, the antisymmetrised tensor product. It is equivalent of multiplying the vectors with the totally antisymmetric n-dimensional Levi-Civita tensor, $\epsilon_{\mu_1 \mu_2 ... \mu_n}$. The full determinant also depends on the amplitude of each column vector:

$$
\det v_\eta^\kappa = \epsilon_{\mu_1 \mu_2 ... \mu_n} \hat{v}_{\mu_1}^1 ... \hat{v}_{\mu_n}^n |v^1|...|v^n|\,. \tag{4.12}
$$

This means that the discrete and the integral version of the mapping equations, eq. (4.10) and eq. (4.11) do not depend explicitly on the amplitude of the dipole vectors for each leg. The equation for the vth leg depends only on the unit dipole $\hat{v}_\mu^{(v)}$ and on the amplitude of the other legs. The remaining part of the determinant in the denominator acts as a weight on the density, emphasising a certain point depending on the amplitude and direction of the other, $v' \neq v$ legs. This determinant represents the only interaction between the legs, when looking for the self-consistent solution, and otherwise it is very similar to the equations described in the previous sections, and bears similarity with the dipole ground state search. Since the ground state search for the system of unit dipoles is a fixed point search, one can argue that the solution for eq. (4.11) can be found as a fixed point search nearby the solution, if the determinant converges faster than the direction of the $\hat{v}_\mu^{(v)}$ dipoles.

The solution for eq. (4.11) is still not a ground state of a dipole system, but for a fixed determinant the individual legs can be regarded as ground state configurations. Nevertheless, unlike the individual legs the full system is not an optimum for a single preferred variable; for example it can not be said that it is the energy minimum of the sum of energies for the individual legs.

4.3.1 The boundary conditions

The integral in the Green's identity is performed on the full space, as it is valid for functions with vanishing tails:

$$
c(x) = \int\limits_{\mathbb{R}^n} G(x, y) \Delta c(y)\, \mathrm{d}^n y\,.
$$

It is a solution of the Poisson differential equation with boundary conditions that both the value of $c(x)$ and its derivatives are zero at infinities. Nevertheless, if $c(x)$ denotes a coordinate it may only be defined on a finite volume and be undefined outside. As it was described in Section 4.2.2 these boundaries in $c(x)$ would be seen as a δ-function in the derivatives, hence these cases have to be treated with extra

caution. A different problem is that although the continuous equation (4.10) is defined everywhere, since it contains $f(x)/\det v_\eta^\kappa$ explicitly, the discrete version eq. (4.11) can only be defined on points where the density $f(x)$ is not zero. Where $f(x) = 0$, the $f(x)/\det v_\eta^\kappa = 1$ should be compensated with infinitely large v_η^κ, which can only be avoided by setting up some boundaries.

The identity contains a surface term when the integral is performed only on a truncated volume, leaving the parts where the second derivatives of $c(x)$ are zero:

$$
\begin{aligned}
c(x) &= -\frac{1}{S_n} \int_V \frac{1}{|x-x'|^{n-2}} \partial_{\mu'} \partial_{\mu'} c(x') \, \mathrm{d}^n x' \\
&= \frac{1}{S_n} \int_V \partial_{\mu'} \frac{1}{|x-x'|^{n-2}} \partial_{\mu'} c(x') \, \mathrm{d}^n x' - \frac{1}{S_n} \int_S \frac{1}{|x-x'|^{n-2}} \partial_{\mu'} c(x') \, \mathrm{d}S_{\mu'} .
\end{aligned}
$$

The identity for the first derivative of $c(x)$ becomes

$$
\partial_\mu c(x) = \frac{1}{S_n} \int_V \partial_\mu \partial_{\mu'} \frac{1}{|x-x'|^{n-2}} \partial_{\mu'} c(x') \, \mathrm{d}^n x' - \frac{1}{S_n} \int_S \partial_\mu \frac{1}{|x-x'|^{n-2}} \partial_{\mu'} c(x') \, \mathrm{d}S_{\mu'} .
$$

The surface term gives a finite contribution to $\partial_\mu c(x)$ for the points within the volume, and behaves like a field of electric charges, and this should also be present in the discrete equation (4.11). This is something undesirable, since it does not allow to do the calculation on the sample points exclusively, but would require an approximation for the integral on a – typically unknown – surface. When combined with the boundary conditions for the $c(x)$, that it should start from C^0 and end at C^1 independently of the other coordinates, it means that on most of the surface $\partial_\mu c(x)$ must be orthogonal to the normal of the surface, while at other places it has to be parallel. At these latter points, where $c(x) = C^0$ and $c(x) = C^1$, point charges are exclusively either positive or negative, and the full system looks like field lines between two capacitor plates. The extreme example of this would be the closed loops, where the two oppositely charged capacitor plates overlap, and the full system can be modelled with the approximation of magnetic dipoles, as it was described in the previous section. For a non-looped coordinate, when the surface term is neglected, the system can still be approximated as a non-ground state solution with a remnant polarisation.

4.3.2 Considerations towards finding the solution

The naive way to solve the Jacobian identity eq. (4.11) is to iteratively feed back the results of the dipole system into the equation again, and wait for stability, an equilibrium. Unfortunately numerical errors in the calculation introduce instabilities, and the iteration may not converge. Such a numerical error may arise in centrally symmetric regions, as it was described in Section 3.3.4, which should be detected and corrected during the iterative search, but nevertheless may contain inaccuracies.

Another method to solve such an equation is the minimisation of differences between the left and the right side of the eq. (4.10):

$$
E_\mu^{(v)} = \frac{1}{S_n} \int_V \partial_\mu \partial_{\mu'} \frac{1}{|x-x'|^{n-2}} v_{\mu'}^{(v)} \frac{f}{\det v_\eta^\kappa} \, \mathrm{d}^n x'
$$

and solving with

$$\min_{v_\mu^{(\nu)}} \sum_\nu \int_{\mathbb{R}^n} \left(v_\mu^{(\nu)} - E_\mu^{(\nu)}\right)^2 \mathrm{d}^n x\,. \tag{4.13}$$

The problem with this formulation is that its gradient contains not only an numerically unresolvable volume integral over $\mathbb{R}^n$, but also contains a difficult derivative of the determinant and $f(x)$ on a non-unit power. This latter feature means that the gradient of eq. (4.13) can not be estimated with an expectation value on a sample, but the approximate knowledge of $f(x)$ density is also required. An other apparent problem with this formulation is the normalisation of the $v_\mu^{(\nu)}$ legs. Because of the use of $\det v_\eta^\kappa$ in the equation, it is not a linear function of $v_\mu^{(\nu)}$ anymore, hence scaling the legs up with a constant factor would not be a solution. For this reason, the solution for eq. (4.11) can only be found with allowing a variation of a scale between the left side of the equation and the results of the right side. This can be fulfilled with comparing only the normalised $v_\mu^{(\nu)}$ with a normalised $E_\mu^{(\nu)}$. However, any kind of normalisation would suffice. For example, fixing the spatial integral of either $|v_\mu^{(\nu)}|$ or $|v_\mu^{(\nu)}|^2$, and comparing it with $|E_\mu^{(\nu)}|$ or $|E_\mu^{(\nu)}|^2$ accordingly only differs slightly in the measure of the difference between $v_\mu^{(\nu)}$ and $E_\mu^{(\nu)}$.

It is also allowed to do a weighted normalisation, which is much easier to perform on samples:

$$\int_{\mathbb{R}^n} f(x)|v_\mu^{(\nu)}|^2 \, \mathrm{d}^n x = \text{const. and } \int_{\mathbb{R}^n} f(x)|E_\mu^{(\nu)}|^2 \, \mathrm{d}^n x = \text{const.}\,,$$

instead of

$$\int_{\mathbb{R}^n} |v_\mu^{(\nu)}|^2 \, \mathrm{d}^n x = \text{const. and } \int_{\mathbb{R}^n} |E_\mu^{(\nu)}|^2 \, \mathrm{d}^n x = \text{const.}\,,$$

because the latter can only be calculated with the knowledge of the $f(x)$ density when it is calculated for a sample.

Rather difficult minimisation functions can be constructed, where the gradient can be calculated without the knowledge of the $f(x)$ density. Consider the following formulation of the equation, where different degrees of freedom can be varied:

$$E_\mu^{(\nu)} = \int_V \partial_\mu \partial_{\mu'} \frac{1}{|x - x'|^{n-2}} v_{\mu'}^{(\nu)} f \, \mathrm{d}^n x'\,, \tag{4.14}$$

and where the density-determinant relation should be introduced as a constraint. An approximation for the determinant, where all the legs of the coordinate system are perpendicular can be considered as an upper limit:

$$\det E_\eta^\kappa \approx \prod_\nu |E_\mu^{(\nu)}|\,,$$

which is now calculated from the output $E_\mu^{(\nu)}$ field, rather than from $v_\mu^{(\nu)}$. This is actually desirable, when one is not looking for the exact solution of eq. (4.11), but only for a model of the coordinate lines. The normalisation for the fields can be chosen as the following:

$$\int_{\mathbb{R}^n} f \frac{|E_\mu^{(\nu)}|}{\det E_\eta^\kappa} \, \mathrm{d}^n x = \int_{\mathbb{R}^n} \frac{1}{\prod_{\nu' \neq \nu} |E^{(\nu')}|} = N^{(\nu)}\,,$$

$$\int_{\mathbb{R}^n} f|v_\mu^{(\nu)}|\, \mathrm{d}^n x = 1\,.$$

This latter normalisation can be done after every iteration, and it can be assumed that it is always true. What has to be achieved, is the equivalence between the $v_\mu^{(\nu)}$ and the $E_\mu^{(\nu)}$ fields. In the case of the constrained eq. (4.14), a minimisation function may look like the following:

$$m(v_\mu^\nu) = \int_{\mathbb{R}^n} f \sum_\nu \left(\frac{E_\mu^{(\nu)}}{N^{(\nu)} \det E_\eta^\kappa} - v_\mu^{(\nu)} \right)^2 \mathrm{d}^n x\,,$$

where f could be any positive weight function, but let's choose the density for convenience. The gradient of $m(v_\mu^\nu)$ will then have three rather complex terms:

$$\frac{\partial m}{\partial v_\nu^\kappa(r)} = - f(r)\left(\frac{E_\nu^{(\kappa)}(r)}{N^{(\kappa)} \det E(r)} - v_\nu^{(\kappa)}(r) \right) \tag{4.15}$$

$$+ f(r) \int f(r') \frac{1}{\det E(r') N^{(\kappa)}} M_{\mu'\nu}(r - r') \left(\delta_{\mu\mu'} - \hat{E}_\mu^{(\kappa)} \hat{E}_{\mu'}^{(\kappa)} \right) \left(\frac{E_\mu^{(\kappa)}(r')}{N^{(\kappa)} \det E(r')} - v_\mu^{(\kappa)}(r') \right) \mathrm{d}^n r'$$

$$+ f(r) \sum_{\lambda \neq \kappa} \int f(r') \frac{E^{(\lambda)\mu}(r')}{\det(r') N^{(\lambda)2}} \left(\frac{E_\mu^{(\lambda)}(r')}{N^{(\lambda)} \det E(r')} - v_\mu^{(\lambda)}(r') \right) \mathrm{d}^n r' \cdot \int f(r'') \frac{1}{\prod_{\iota \neq \lambda} |E^{(\iota')}(r'')|} M_{\mu'\nu}(r - r'') \hat{E}_{\mu'}^{(\kappa)} \mathrm{d}^n r''$$

All three terms, with the exception of the leading $f(r)$, can be calculated with using expectation values. But due to the non-linearity of the problem, the three terms depend differently on the normalisation of the $v_\nu^{(\kappa)}$ fields, hence it can be arbitrarily chosen whether the first or the second two terms are dominant. The minima does not scale with the normalisation, but during the minimisation the last two terms can be neglected. Using only this part of the gradient will always get closer to the minima in some extent. It should be noted that an additional term is missing, namely the constraint that the different legs should be perpendicular, as it was assumed for the approximation of the determinant.

4.3.3 Two ways of determinant estimation

Although eq. (4.15) shows that the solution for eq. (4.11) can be found by minimisation, it is still prone to numerical errors. This can be demonstrated with analysing the leading part in eq. (4.15),

$$\frac{\partial m}{\partial v_\nu^\kappa(r)} = -f(r)\left(\frac{E_\nu^{(\kappa)}(r)}{N^{(\kappa)} \det E(r)} - v_\nu^{(\kappa)}(r) \right)\,.$$

As it was described in Section 3.3.4, the centrally symmetric regions in the discretised equations typically underestimate the densities. Similarly, eq. (4.11) underestimates the output field strength compared to the continuous eq. (4.10). This is true not only for the legs where the grounds state is sought, but also for the energy maxima. When one leg forms such a configuration, assuming the other legs are perpendicular to it, it is likely that these other legs are also non-homogenous; hence in small regions all the $E_\mu^{(\nu)}$ field strengths are underestimated. As a consequence, the determinants are underestimated with a much greater power, giving a unnecessarily large contribution to the gradient in this region. In the next step of the iteration, these regions will have the dominant $v_\mu^{(\nu)}$ input, like a peak, and will give the dominant contribution for the $E_\mu^{(\nu)}$ fields at most of the places. This behaviour makes the iteration unstable, and some correction is inevitable.

The detection of centrally symmetric and inhomogenous regions, described in the Section 3.3.4, and the dynamic change of the exclusion area can shrink the area where the determinant is underestimated, but it can not make it disappear. This technique only moderates the problem, but it is not a final solution. As it was mentioned in the previous section, eq. (4.11) looks like a ground state configuration of a dipole system, for fixed determinants. This suggested that an iterative search might be stable when the used determinants converge faster than the dipoles itself, and this situation is favourable even if the determinants converged to crude approximations.

Orthogonal upper limit

The full determinant in eq. (4.12) depends not only on the magnitude of the $E_\mu^{(\nu)}$ fields, but also on the angle between them. This is a large source of discrepancy, since despite that the angles between the $v_\mu^{(\nu)}$ fields can be constrained during the iterations, the $E_\mu^{(\nu)}$ angles are always a product of a complex function which is rather hard to set constraints on. Uncertainty in the direction of any of the $E_\mu^{(\nu)}$ fields can introduce a factor between $[0, 1]$ on the determinant, eventually causing division by nearly zero in eq. (4.11). This supports the idea to estimate the determinants only with the absolute values of $E_\mu^{(\nu)}$:

$$\det E_\eta^\kappa \approx \prod_\nu |E_\mu^{(\nu)}| \,.$$

This implicitly assumes that the $E_\mu^{(\nu)}$ fields are perpendicular to each other, or in other words, that the new coordinate system has perpendicular legs everywhere. This can be maintained approximately, by keeping the $v_\mu^{(\nu)}$ legs perpendicular for every data point. The two ways to do this is either the direct orthogonalisation or the Lagrange constraint. The direct orthogonalsation would look like the following:

$$\tilde{v}_\mu^{(\nu)} = v_\mu^{(\nu)} - \sum_{\nu' < \nu} \frac{1}{|\tilde{v}^{(\nu')}|^2} v_\eta^{(\nu)} \tilde{v}_\eta^{(\nu')} \cdot \tilde{v}_\mu^{(\nu')} \,.$$

This creates a new $\tilde{v}_\mu^{(\nu)}$ collection of vectors from the old ones, $v_\mu^{(\nu)}$. The problem is that it requires an ordering between the different legs, and may eventually result in null vectors. Nevertheless, when the update of the vectors is done slowly, namely the step size of the gradient method is small enough, the orthogonality of the vectors is only slightly violated, hence a regular orthogonalisation after an iteration would not interfere much with the gradient.

As it was mentioned above the other method is the approximate orthogonalisation via Lagrange multipliers. This is a more symmetric approach, based on adding an extra term to the minimisation function:

$$m_L(v_\eta^\nu) = \lambda \int \sum_{\nu\nu'} \left(\frac{v_\mu^{(\nu)} v_\mu^{(\nu')}}{|v^{(\nu)}||v^{(\nu')}|} - \delta_{\nu\nu'} \right)^k \mathrm{d}^n x \,. \tag{4.16}$$

For a large λ and an even k, this behaves as a penalty potential when the $v_\mu^{(\nu)}$ vectors are not perpendicular. Although this is a simple, elegant and symmetric orthogonalisation technique it has a clear disadvantage over the direct orthogonalisation. This constraint can introduce many local minima, and due to its symmetries, it stays trapped. However, the direct orthogonalisation, with the leg ordering can jump over these barriers.

Artificial brakes without angle constraints

Since the previous method assumes that the legs of the coordinate system are perpendicular, a more robust method without this constraint is also necessary. As the instability of the iteration mainly comes from the underestimation of the used determinant, a possible remedy is to introduce artificial breaks that prevents using underestimated determinants. Practically this means that during the iterative search, the $v_\mu^{(\nu)}$ fields should be updated in a way that is consistent with some expectations. The re-parameterised version of eq. (4.10), namely eq. (4.14)

$$E_\mu^{(\nu)} = \frac{1}{S_n} \int_V \partial_\mu \partial_{\mu'} \frac{1}{|x - x'|^{n-2}} v_{\mu'}^{(\nu)} f \, \mathrm{d}^n x'$$

can be used here as well, but with the following formulation of the requirements for the solution:

$$E_\mu^{(\nu)} = f v_\mu^{(\nu)} \,,$$

$$\det v_\mu^{(\nu)} = \frac{1}{f^{n-1}} \,.$$

The combination of the two equations results in $\det E_\eta^\kappa = f$, a similar requirement as in the previous section, but this formulation allows a different handle during the iterations. As practically the $v_\mu^{(\nu)}$ fields are sought parameters, these degrees of freedom can be constrained to anything during the calculations, and these are the ones that should be prevented of diverging due to a numerical error. This can be modelled with a minimisation function:

$$m(v_\eta^\kappa) + m_L^{\mathrm{det}}(v_\eta^\kappa) = \sum_\nu \int \left(f v_\mu^{(\nu)} - E_\mu^{(\nu)} \right)^2 \, \mathrm{d}^n x + \lambda^\mathrm{d} \int \left(\det v_\eta^\kappa - \frac{1}{f^{n-1}} \right) \mathrm{d}^n x \,.$$

There is still a problem with this minimisation function, because the interaction between the different $v_\mu^{(\nu)}$ legs only occurs through the constraint. In practice, the update of $v_\mu^{(\nu)}$ in one iteration should be done first, according to the gradient, then a modification can happen to the $v_\mu^{(\nu)}$ amplitudes depending on their deviations from the constraint. This means that the determinant of the initial values of $v_\mu^{(\nu)}$ should be close to the ideal determinant, which can be done relatively easily; but the angles between the several $v_\mu^{(\nu)}$ legs should also not change rapidly, otherwise the rescaling of the legs after each iteration would introduce the same numerical instability that was supposed to be prevented. The fitness of the model can be measured by the unconstrained part of the minimisation function, which should be zero for an ideal solution. As it was previously mentioned, a symmetric angle constraint can introduce additional minima, and the iteration can easily got stuck there. A non-symmetric version of eq. (4.16) can be created by introducing an ordering into the set of the legs:

$$m_L^{\mathrm{asym}}(v_\eta^\nu) = \lambda^\mathrm{a} m_L \int \sum_{\nu > \nu'} \left(\frac{v_\mu^{(\nu)} v_\mu^{(\nu')}}{|v^{(\nu)}||v^{(\nu')}|} \right)^k \mathrm{d}^n x \,.$$

The combination of the minimisation function and the two constraints can be solved with the steepest descent method, which can be modelled with the following algorithm. To calculate the initial conditions:

1. Approximate the density of the sample, $a(x)$, which will be used later to regulate the used determinants.

2. Set the initial values of $v_\mu^{(\nu)}$ to be perpendicular, and scale it with $\sqrt[n]{\frac{1}{\det v_\eta^\kappa a^{n-1}}}$.

3. Normalise $v_\mu^{(\nu)}$ with $\sum_i v^{2(\nu)}(x_i)$.

The procedure for every iteration is:

1. For every ν leg calculate the expectation value

$$E_\mu^{(\nu)}(x_j) = \frac{1}{S_n N} \sum_i \partial_\mu \partial_{\mu'} \frac{1}{|x_j - x_i|^{n-2}} v_{\mu'}^{(\nu)}(x_i).$$

2. Normalise $E_\mu^{(\nu)}$ with $\sum_i (E_\mu^\nu(x_i)/a(x_i))^2$.

3. Update $v_\mu^{(\nu)}$ with the main part of the gradient,

$$v_\mu^\nu := v_\mu^\nu - \lambda_{\text{step}} \left(\frac{E_\mu(x_i)}{a(x_i)} - v_\mu^\nu(x_i) \right).$$

4. Normalise $v_\mu^{(\nu)}$ again.

5. Approximate fitness with

$$m_L = \sum_{i\nu} \frac{1}{a(x_i)} \left(a(x_i) v_\mu^{(\nu)} - E_\mu^{(\nu)} \right)^2.$$

6. Apply a fitness based correction for the angles, preferring the ones with larger amplitude

$$v_\mu^{(\nu)} := v_\mu^{(\nu)} + m_L \sum_{\nu' < \nu} \frac{a(x_i)}{|E^{(\nu)}|} \hat{v}_\mu^{(\nu')} \hat{v}_\alpha^{(\nu)} \hat{v}_\alpha^{(\nu')}.$$

7. Normalise $v_\mu^{(\nu)}$ again.

8. Calculate the approximate determinant

$$d(x_i) = \det v_\eta^\kappa(x_i).$$

9. With the ratio of the approximate determinant and the target determinant $r = (a^{n-1}(x_i)d(x_i))^{-1}$, scale $v_\mu^{(\nu)}$ with ϵ, where ϵ allows a 10 % fluctuation up but no fluctuation down:

$$\epsilon = \begin{cases} \sqrt[n]{r} & \text{if } r < 1.1 \\ \sqrt[n]{1.1} & \text{if } r > 1.1 \end{cases}.$$

10. Normalise $v_\mu^{(\nu)}$ again.

4.3.4 Classification

Unsupervised classification is sometimes called clusterisation, but unlike many of these methods the requirement here is not peak-finding but the separation of hyper-surfaces, as these can be identified as physical objects. The design principle of the principal coordinate method was to find a coordinate system in which the sample is uniformly distributed. Although the structural similarity with the principal curves shows that the coordinate lines curve along the hyper-surfaces within the original distribution,

additional postulates are needed to identify the hyper-surfaces. This is because the coordinates are continuous variables and none of them suggest a direct boundary.

A meaningful criterium can be derived from the notion of information entropy. In the ideal new coordinate system that satisfies eq. (4.10) the sample can be modelled with a uniform distribution, which means that its information entropy is maximal, since all possible values are used equally to describe the original distribution. This is a goal of many data compression algorithms, like the Huffman-coding [20].

To establish the idea of compression, one must select variables which can be truncated to a certain precision without losing the possibility to reconstruct the original data-set. The truncation of certain variables can go as far as discretising them, which being done in the principal coordinate system eventually means the recognition of the hyper-surfaces within the original data set. The actual error between the original data point and its representation through the truncated principal coordinates can vary with the machine representation of the number. Generally it can be said that a Δc_ν uncertainty in the new coordinate c_ν will lead to the Δx_μ uncertainty in the back-transformed coordinate x_μ through the transformation matrix $v_\mu^{(\nu)}$, which is simply

$$v_\mu^{(\nu)}\Delta x_\mu = \Delta c_\nu .\tag{4.17}$$

For the principal coordinate system $\det v_\mu^\nu = \kappa f$ holds. Hence along the coordinate line on which the f density drops, $v_\mu^{(\nu)}$ is large and as a consequence of eq. (4.17) the uncertainty of the reconstruction of the original data decreases. This is why it is reasonable to truncate most of those c_ν coordinates along which $v_\mu^{(\nu)}$ is large. Those coordinates are typically perpendicular to the hyper-surfaces. It must be noted that traditional peak-finding clusterisation can be achieved by fitting a radial coordinate system to the sample.

4.4 Conclusions

The two determinant estimation methods show similarity in the goals of well known algorithms, like multidimensional scaling [21] and self-organised maps or Kohonen maps [22]. The crucial detail that differentiates them from the known methods is that the dimensionality of the sought hyper-surface is not pre-defined, but is determined from the measurements. Furthermore, in the principal coordinate system the sample is transformed into a uniform distribution, which can be described with independent random variables, providing a unique point of view as it makes it possible to identify hyper-surfaces and reduce the dimensionality of the data with minimal information loss. In other words, the identification of hyper-surfaces helps in the simplification of the description of the data by formulating its inner constraints, the laws that describe it. In the method explained in this work, the dipole formulation could be regarded as one of the possible ways to solve the differential equation in eq. (4.10), but its advantage is that it selects a smoothed coordinate system from the possible solutions and is capable to find open and closed curves at the same time.

CHAPTER **5**

Polynomial regression

Chapter 3 dealt with a density estimator, which was then used to calculate the posterior probability of a certain phase space point being signal or background. As it was mentioned this task can also be achieved with a regression technique. This chapter describes such a method based on the polynomial expansion of the ideal regression function which appears to be fast and accurate. Due to its inner structure, it is capable of extracting only the statistically relevant information while neglecting the statistical noise associated with sampled data. This chapter contains a short summary of the polynomial expansion method described in the article [3], while extending it with further methods increasing its accuracy and speed for high dimensional inputs.

5.1 Polynomial expansion of the ideal classification function

According to the Neyman-Pearson lemma [23], the optimal separation contours between a signal sample with density $s(x)$ and a background sample with density $b(x)$ are following the $s(x)/b(x)$ density ratio contours. Given large statistics, these contours can be found by fitting a $y_i = F(x_i)$ function using $y_i = \pm 1$ as target for the signal and the background for data points x_i. Nevertheless, this fitting function only approximates the ideal classifier function, which is the following for the ± 1 targets:

$$F(x) = \frac{s(x) - b(x)}{s(x) + b(x)} \,. \tag{5.1}$$

A polynomial expansion of this function is also possible, without the knowledge of the theoretical $s(x)$ and $b(x)$ distributions, only using samples. The coefficients of the $F(x) = \sum_k F_k x^k$ polynomial can be found when the equation is multiplied by $s(x) + b(x)$, and a $\mathcal{F}$ Fourier transformation is performed. The multiplication of the two x dependent functions under Fourier transformation is a convolution, which becomes a discrete matrix equation for the polynomial coefficients. The Fourier transform of a distribution is the same as the expectation value of the phase $e^{ix\omega}$. It is known as the characteristic function, and has the unique feature that its polynomial expansion can be expressed with the moments of the untransformed distribution:

$$\mathcal{F}[s(x)] = \mathbb{E}_s[e^{ix\omega}] = \sum_j \langle x^k \rangle_s \frac{(-i\omega)^k}{k!} \,.$$

When eq. 5.1 is multiplied by $s(x)+b(x)$, the polynomial of $F(x)$ introduces a series of x^k to the left-hand

side. The Fourier transform of each $x^k[s(x) + b(x)]$ can be substituted with the kth derivatives of $s(x)$ and $b(x)$. These derivatives can be easily approximated with the polynomials made from the moments of the sample, and finally arrive to the equation for the polynomial coefficients:

$$\sum_k F_k \left(\langle x^{j+k} \rangle_s + \langle x^{j+k} \rangle_b \right) = \langle x^j \rangle_s - \langle x^j \rangle_b .$$

Here $\langle x^{j+k} \rangle_s - \langle x^{j+k} \rangle_b$ can be treated as a matrix, and F_k can be found with linear equation solvers. Such a matrix is called a Hankel matrix, for which so-called superfast solvers exist [24]. The method itself shows similarity with the Stieltjes and the Hamburger moment problems. Nevertheless, as it is shown later in this chapter, the method can be extended with robust checks against overfitting.

It is possible to extend the method for general regression, and to be able to handle higher dimensional x input spaces [3], by using tensor coefficients in the polynomial expansion and for the moment calculations. As all these tensors are symmetric, the number of free parameters is small. A n-dimensional tensor with d-degree has n^d parameters, of which only $\binom{n+d-1}{d}$ are free. Although this is a large reduction, it still grows rapidly with the tensor degree. The two limitations for the method both appear only at large tensor degrees, and they are related to the numerical precision of the matrix solvers:

- The high tensor degrees require calculation of high moments, which becomes a limiting factor around degree twenty for double precision and degree thirty on quadruple precision.

- The other limitation comes from the size of the appearing matrix, since this matrix is dense and depending on the required tensor degree, its elements may cover many orders of magnitude. This typically happens when the matrix size reaches a few thousand rows and columns and result in large numerical errors during solving the equations.

5.2 Propagation of uncertainty

The great advantage of using the moments as the input for the calculations is that due to the central limit theorem these have easily calculable Gaussian uncertainty and correlations. The moments are formed by weighted sums. For N number of entries and $\sum w_i = N$ weights it is calculated as the following

$$\langle x^k \rangle = \frac{1}{N} \sum_i w_i x_i^k .$$

This formulation allows $w_i x_i^k$ to be treated as a single random variable, making it easier to calculate the covariance matrix:

$$\text{Cov}\left(wx^j, wx^k \right) = \mathbb{E}\left[\left(wx^j - \langle wx^j \rangle \right) \left(wx^k - \langle wx^k \rangle \right) \right] = \frac{1}{N} \sum_i w_i^2 x_i^{j+k} - \langle x^j \rangle \langle x^k \rangle .$$

The covariance of the averages is then the scaled down version of the weighted values[1]:

$$\text{Cov}\left(\langle x^j \rangle, \langle x^k \rangle \right) = \frac{1}{N-1} \text{Cov}\left(wx^j, wx^k \right) .$$

It is possible to propagate this uncertainty into the χ^2 of $F(x)$, and check if a new tensor order significantly decreases the χ^2 or not. Due to the simplicity of how $F(x)$ is being calculated, this takes a

---

[1] It should be noted that this covariance matrix is not related to the standard deviation of the x^k averages, as it contains the square of the w_i weights.

rather simple form. Using y_i as the target value to be regressed on the random field x_i, the following simplifications can be done to calculate the coefficients of the $F(x)$ polynomial. With the introduction of the G_{kj} matrix, which contains the correlations of the x^j and x^k power series, the F_j coefficients are calculated with

$$G_{kj} = \langle x^j x^k \rangle = \langle x^{j+k} \rangle \,,$$
$$F_j = \sum_k G_{kj}^{-1} \langle y x^k \rangle \,. \tag{5.2}$$

Treating the F_j coefficients as a vector, the $F(x)$ value can be calculated as a scalar product between F_j and a vector of diadic tensors of powers of x, namely x^j:

$$F(x) = \sum_j F_j x^j = \sum_{kj} x^j G_{kj}^{-1} \langle y x^k \rangle \,.$$

This vector formulation allows now a calculation of the squared loss function, or χ^2 of $F(x)$ and the training sample:

$$\chi^2 = \frac{1}{N} \sum_i w_i (y_i - F(x_i))^2 = \langle y^2 \rangle - 2\langle y F(x) \rangle + \langle F(x)^2 \rangle \,,$$

where the last term can be simplified further to

$$\langle F(x)^2 \rangle = \sum_{jkl} \langle y x^k \rangle G_{kj}^{-1} \underbrace{\sum_i w_i x_i^j x_i^k}_{G_{jk}} G_{kl}^{-1} \langle y x^l \rangle = \sum_k \sum_i w_i y_i x_i^k F_k = \langle y F(x) \rangle \,.$$

Altogether, the χ^2 of the $F(x)$ function can be minimised by maximising the correlation of $F(x)$ to the target y, since

$$\chi^2 = \langle y^2 \rangle - \langle y F(x) \rangle \,.$$

This loss function both depends on the $\langle y x^k \rangle$ and the $\langle x^k \rangle$ averages; the latter one appears only via the G^{-1} inverse. Assuming that $\langle y^2 \rangle$ can not be varied, the uncertainty of the χ^2 can be approximated with Gaussian error propagation, for a given d maximal degree of the $F(x)$ polynomial. The standard deviation of the χ^2 is then

$$p_j = \left(\langle y x^0 \rangle, ..., \langle y x^d \rangle, \langle x^0 \rangle, ..., \langle x^{2d} \rangle \right) \,,$$

$$\sigma_{\chi^2}^2 = \sum_{jk} \frac{\partial \chi^2}{\partial p_j} \text{Cov} \left(p_j, p_k \right) \frac{\partial \chi^2}{\partial p_k} \,. \tag{5.3}$$

The covariance term can be calculated as it was discussed before, but additionally it has to contain the correlation between the $\langle y x^k \rangle$ and $\langle x^k \rangle$ terms. Using the rule for the derivative of an inverse matrix, $\frac{\partial G^{-1}}{\partial p_j} = -G^{-1} \frac{\partial G}{\partial p_j} G^{-1}$, the partial derivatives of the χ^2 can be written as follows:

$$\frac{\partial \chi^2}{\partial \langle yx^j \rangle} = -\frac{\partial \langle yFx \rangle}{\partial \langle yx^j \rangle} = -2F_j$$

$$\frac{\partial \chi^2}{\partial \langle x^j \rangle} = \sum_{aklm} \langle yx^a \rangle G_{ak}^{-1} \frac{\partial G_{kl}}{\partial \langle x^j \rangle} G_{lm}^{-1} \langle yx^m \rangle = \sum_{ak} F_a \frac{\partial G_{ak}}{\partial \langle x^j \rangle} F_k = F_j^c. \qquad (5.4)$$

Since $\frac{\partial G_{ak}}{\partial \langle x^j \rangle}$ is a three index tensor filled with ones and zeros, F_j^c can be regarded as a convolved version of the F_j coefficients. Combining the partial derivatives of χ^2 in eq. (5.4) with the covariance matrix in eq. (5.3) gives the uncertainty of the χ_d^2 fitness of $F(x)$ for a given d degree. This also allows the calculation of the uncertainty of the $\chi_{d+1}^2 - \chi_d^2$ difference, showing whether a new F_{d+1} term significantly decreases the χ^2 or not. Using the t-probe to test $\chi_{d+1}^2 - \chi_d^2$ being different from zero, one can regulate the $F(x)$ regression function with a given confidence level while determining the maximal significant degree of the $F(x)$ polynomial. The equations are very similar for the higher dimensional case, though it is unclear from this formulation how to order the different terms of the n-dimensional, d-degree tensor F_d. This is why a different method was also developed, which can determine the importance of the different terms.

5.3 Maximising the significance

A typical question about polynomial regression is the maximal degree used. In the solution for the F_i coefficients in eq. (5.2) it is unclear where to truncate the matrix G_{ij}. Neglecting those moments which are smaller than the main terms is ambiguous, since a transformation on the input space can largely affect these ratios. An example is a distribution that is between $x \in \{0, 1\}$, which has moments that are converging to 0 with increasing powers. However, when this distribution is shifted to $x \in \{1, 2\}$, the x^k powers are definitely larger than the x^{k-1} powers, though the complexity of the regression function is intact.

In the case of classification one can see that the $\langle yx^k \rangle = \langle x^k \rangle_s - \langle x^k \rangle_b$ averages hold crucial information. The $\langle x^k \rangle_s$ moment series contain the Taylor expansion of the Fourier transform of the signal distribution, while $\langle x^k \rangle_b$ is the same for the background distribution. It can be checked whether the difference is significantly different from zero or not, and that can be used to terminate the moment series. Nevertheless, as it was said in the previous section, the uncertainties of the different moments are correlated and in the higher dimensional formulation the different moments can not be ordered.

To circumvent the ordering problem and the correlations, a certain subset of variables can be selected which are significantly different from zero when being combined. This can be performed, since the $\langle x^k \rangle_s - \langle x^k \rangle_b$ difference is a vector with an associated Gaussian uncertainty, and it can be checked whether this vector is so different from zero that it is unlikely to be a statistical fluctuation. Furthermore, this can be performed on various subspaces, from which the most significant can be selected.

5.3.1 Significance of variables

For variables $v_k = \langle x^k \rangle_s - \langle x^k \rangle_b$, the uncertainty of the d amplitude can be calculated by the propagation of the correlated uncertainties between the variables. Assuming that the signal and background samples are independent, the covariance matrix of v_k is simply the sum of the two covariance matrices:

$$E_{jk} = \text{Cov}\left(\langle x^j \rangle, \langle x^k \rangle \right)_s + \text{Cov}\left(\langle x^j \rangle, \langle x^k \rangle \right)_b ,$$

$$d = \sqrt{\sum_j v_j^2} \, .$$

With these definitions the standard deviation of the d amplitude is

$$\sigma_d^2 = \sum_{jk} \frac{\partial d}{\partial v_j} E_{jk} \frac{\partial d}{\partial v_k} = \frac{\sum_{jk} v_j E_{jk} v_k}{d^2} \, , \qquad (5.5)$$

and with which the confidence levels for the hypothesis that the true amplitude is zero can be calculated. These terms make it easy to calculate the significance, namely that

the null hypothesis can be excluded with $\sqrt{\dfrac{d^2}{\sigma_d^2}}$ sigma coverage.

It must be noted that eq. (5.5) is only an approximation of the uncertainty of d, since d can only be approximated with a Gaussian if d/σ_d is large enough. For very low d/σ_d values the higher order covariances would dominate the uncertainty.

5.3.2 Selection of the most significant group of variables

To reduce the number of rows in the calculation of eq. (5.2) it is possible to select only those variables which give the largest contribution to the significance of the $v_k = \langle x^k \rangle_{\mathrm{s}} - \langle x^k \rangle_{\mathrm{b}}$ vector. Either the significance has to be maximised or its inverse has to be minimised by keeping only the S set of variables:

$$\text{select the set } S \text{ for which} \quad \min_{S \in \{\text{possible sets}\}} \frac{\sum_{j,k \in S} v_j E_{jk} v_k}{d^4} \, . \qquad (5.6)$$

For a large number of rows an iteration through every subset is unfeasible, hence some heuristics must be used. One can start from a conveniently large set, and from those select the most significant individual row and put it into S. The next step is to extend this set with the variable that increases the significance the most. The increase of the significance may stop at a certain size of S for two reasons: either the leftover variables can be predicted from the selected set or the arising numerical errors overcome the small increase of the significance.

This method can be extended with various checks. The growth of S might be stopped at a predefined limit, before including all the variables. A possible stopping criteria is a minimum increase in the significance. The selection process can also be repeated on the remaining variables, as it is possible that the same significance can be reached with a smaller number of variables.

An obvious source of error must be avoided though, and it is selection bias. When the number of rows is large enough, the chance that a selected subspace has large significance also grows, hence the significance threshold must be selected in a way that is bigger than this bias.

5.3.3 Optimisation of weights

The minimisation process in eq. (5.6) is equivalent of setting 1 and 0 weights to some of the v_i variables during the iterations. Since the formula to be minimised is relatively simple, it is possible to calculate its minimum when allowing real-valued a_j weights for each v_j value. This is equivalent to a non-linear transformation of the x input space where the original v_j averages are transformed into $a_j v_j$. This also

means, that the different rows and columns of the uncertainty matrix E_{ij} have to be reweighted too, therefore only the square of the weights, $b_j = a_j^2$, appear in the significance calculation:

$$M(b_0, ...) = \min_{a_j} \frac{\sum_{j,k} v_j a_j^2 E_{jk} a_k^2 v_k}{\left(\sum_l v_l a_l^2 v_l\right)^2} = \min_{b_j} \frac{\sum_{j,k} v_j b_j E_{jk} b_k v_k}{\left(\sum_l v_l b_l v_l\right)^2} \tag{5.7}$$

At the minimum, the derivative of $M(b_j)$ has to be zero, which results in an equation that can be solved analytically for b_j.

$$d = \sqrt{\sum_l v_l b_l v_l}$$

$$\frac{\partial M}{\partial b_n} = 4 \frac{\sum_j v_j b_j E_{jn} v_n b_n}{d^4} - 2 \frac{\sum_{jk} v_j b_j E_{jk} v_k b_k v_n^2}{d^6} = 0$$

$$\sum_{jk} v_j b_j E_{jk} v_k \left(\delta_{kn} - \frac{b_k v_n^2}{d^2}\right) = 0. \tag{5.8}$$

In eq. (5.8) a rescaled version of the uncertainty matrix appears as $W_{jk} = v_j E_{jk} v_k$, whereas $b_k v_n^2$ is a dyadic matrix with a single non-zero eigenvalue. The equation resembles an eigenvector equation for this dyadic matrix. Using matrix notations, with the vector $v_n' = v_n^2$ this becomes

$$b^T W \left(I - \frac{bv'^T}{d^2}\right) = 0. \tag{5.9}$$

This formulation shows that the $b^T W$ vector might be an eigenvector of the bv'^T matrix with the eigenvalue d^2 at the extrema of M. Other solutions are also possible if the W matrix have zero eigenvalues, which would also annull the expression. For invertible W the solution is then a weight-square vector that is perpendicular to the following:

$$b_j = \sum_k W_{jk}^{-1} v_k' = \sum_k (v_j E_{jk} v_k)^{-1} v_k^2, \tag{5.10}$$

without appropriate normalisation. The full solution for a singular W would be the same for the invertible subspace plus any vector that is annulled by W, combined with appropriate normalisation.

In the case when the uncertainties of the v_i variables are uncorrelated, the E matrix is diagonal and the square of the weights are simply the inverse of the diagonal elements, or the standard deviations.

5.3.4 Remarks on the significance maximisation methods

The two methods might be combined, first creating a subset of variables with large significance than calculating their optimal weights. Using only the weight optimisation may be tempting, but a nearly singular uncertainty matrix may introduce numerical errors in the solution. Applying the method on the $v_k = \langle x^k \rangle_s - \langle x^k \rangle_b$ vector is not a minimising the χ^2 though, it is just making sure that the components used are statistically well established and their individual effect on the χ^2 will be significant. The full solution would be setting a minimum precision required for the χ^2 and selecting those components that minimise it with at least this precision.

The modification of the methods to non-binary regression target is also possible, and it is expected to help there against overfitting there as well.

5.4 Extensions towards a neural decision tree

The previous section can help to avoid overfitting the input sample, but it can not solve the problem of increasing numerical errors in the calculation of higher moments. To overcome these numerical errors one must emulate the calculation of higher moments without calculating the high powers of the input variables. A possible method can be explained with a simple example. The high degree moments are only necessary if the degree of the $F(x)$ function requires it, and this can be avoided by splitting the input space into distinct regions. A region small enough can always be successfully fitted with a low degree polynomial.

Such a tilling has a disadvantage though, because non-differentiable jumps may appear at the region boundaries. One can reduce this effect by not using straight boundaries, but only smoothed transitions from one to an other, applying sigmoid weights with a given width. A combined fit would look like the following with the $w(x)$ sigmoid function, boundary at m and transition width Γ:

$$w(x) = \text{atan}\left(\frac{x - m}{\Gamma}\right) + m$$

$$F(x) = w(x)F_1(x) + (1 - w(x))\,F_2(x)\,. \tag{5.11}$$

The $F_1(x)$ function can be fitted on the $w(x)$ weighted dataset while $F_2(x)$ on the $1 - w(x)$ weighted dataset. Nevertheless, the weights of the dataset might not necessarily be the same as the weights on the function combination in eq. (5.11), as with large enough statistics $F_1(x)$ would be equal to $F_2(x)$ in the transition region. A simple way to handle the smoothness of the combined fit function is then to use wider sigmoids for the dataset weights than for the function weights.

Considering computational complexity, the number of regions should also be minimal. Using rectangular boundaries on a n-dimensional input space would produce 2^n regions, hence this must be optimised somehow. A possible way to select the interesting regions is performing a fit first on the full input space and use the contours of this $F_{\text{pre}}(x)$ fit to decide about the subregions where $F_1(x)$ and $F_2(x)$ is going to be fitted. The width of the $F_{\text{pre}}(x)$ distribution can also be used to determine a practical width for the dataset sigmoid weights and the function combination weights. As the $F_{\text{pre}}(x)$ pre-fit is already a valuable approximation, a further improvement can be done by using its value. The $F_1(x)$ and $F_2(x)$ functions might be regressed only for the difference between $F_{\text{pre}}(x)$ and the y target. This can reduce their functional complexity, resulting in faster calculations. The sub-region fits can be used to split the regions further where new fits can be performed, forming a multi-layer decision tree.

A proof for the effectiveness of this method can be seen on fig. 5.1. A second degree curve can be successfully regressed with a multi-layer tree-based network, where each node does only a linear fit.

5.5 Details of the implementation

Not all the features were implemented into the C++ code that was used to produce the figures shown in this thesis. The significance calculation was only applied at the entry level, to determine what pre-calculated input variables might be useful for the training, maximising their linear correlation to the fitting target. However this feature can be turned off, to be able to compare the algorithm with other methods.

The code is theoretically capable of calculating the fitting polynomial up to arbitrary number of dimensions and degrees, as it was described in the cited article [3]. However, the tree-based regression method uses second degree polynomials for the high dimensional fits, as it is the lowest degree that can describe closed curves. Such closed curves are useful for these binomial regressions, where the signal is
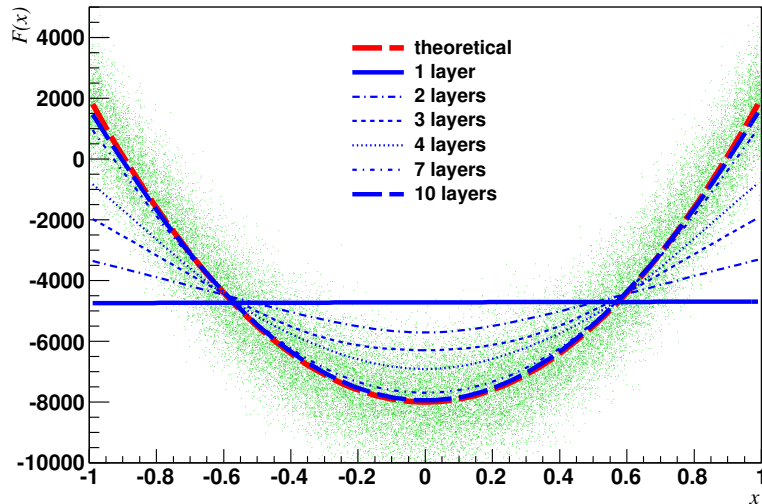
Figure 5.1: Example to show the how a decision tree can help regressing high degree functions. The target distribution (green) on this figure follows a simple second degree curve, but it can be successfully regressed with a few layers of linear functions. The small, smoothed transition region around 0 can be observed between the linear fits at $x > 0$ and $x < 0$ for the 2-layer tree.

peaked above a background, hence the ideal classification contours are closed curves around the peak. As the low degree polynomials may overshoot the target in the non-central regions, the output of this high-dimensional low-degree curve is fed into a one-dimensional, but high-degree shaping layer. The advantage of the shaping layer can be observed on fig. 5.2.

The number of decision layers used was determined using independent training and testing samples, stopping at a level where the $F(x)$ distributions of the two samples were still similar. The decision boundaries were always at the mean of the $F(x)$ distributions. The used Γ transition width in the function combination sigmoid was equal to the width of the $F(x)$ distribution, while the data weight sigmoid width was 3Γ.

The C++ code utilises the Eigen Matrix Library [25] for the linear algebraic calculations, but some features can be compiled using the LAPACK [26] library instead. It compiles with both the `clang` and the `g++-4.4` compiler under MacOS X and various flavours of Linux.

5.6 Conclusions

The core of the presented regression method is a polynomial regression that is based on using the moments of the underlying distributions, making it possible to calculate the uncertainties and their propagations into the fit function. The polynomial fitting itself is very robust and fast, and only requires a single linear equation solver for the fitting. Unlike many other polynomial fitters, the size of the utilised matrix does not grow with the number of sample points, but only with the target degree of the fitting polynomial. The applied neural weight suppressions combined with a decision tree helps to avoid numerical errors that might appear with high degree polynomials. What is important, is that none of the calculations require optimisations via parameter scanning or explicit minimizations, making the full regression procedure significantly faster compared to other methods like neural networks, boosted decision trees or support vector machines.

(a) 3-dimensional 20-degree polynomial

(b) 3-dimensional 20-degree polynomial with a one-dimensional 20 degree shaping layer
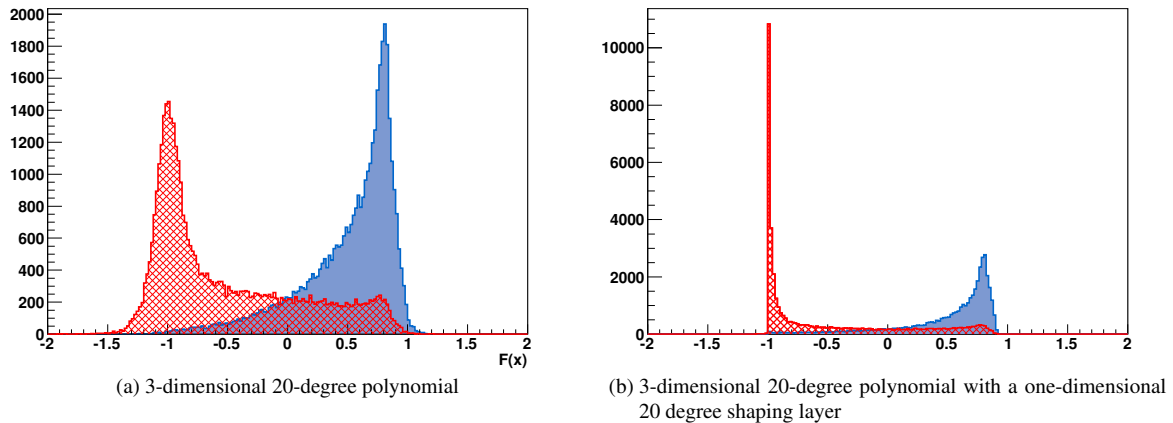
Figure 5.2: Application of a shaping layer to a high dimensional fit can be advantageous. The original high dimensional classifier output on fig. (a) is shaped with a high-degree one-dimensional polynomial in fig. (b). This latter output is now between the boundaries of the ±1 targets for the signal (blue) and the background (red) distributions, and the $F(x)$ output correlates better with the purity of the signal. The sample is the same sum of twelve 3-dimensional Gaussian peaks and flat background that was used for demonstration in fig. 3.8

.

The possibility to select the statistically significant set of variables naturally arises from the method, and it is unique compared to other methods which typically require the separation of the available data set into independent training and testing samples. As Section 5.2 has shown, the method allows finding a regression function with statistical confidence in a very simple and fast way.

The high dimensional classifications are approximated in a three step way, combining:

- preprocessing by variable selection and variable normalisations,

- high-dimensional low degree regressions,

- post-processing with one-dimensional high degree output shaping.

Hence a name was coined for the method, calling it *tripoly*. This short name is used in some of the figures in this thesis, indicating if the plot was produced with the method described in this chapter.

CHAPTER **6**

QCD, jets and Monte Carlo

The measurement of particles that do not exist at everyday energies is only possible by observing high energy collisions where these might be produced. To predict their effect on the detectors a complex simulation infrastructure is needed, which is separated into two distinct steps. The simulation of individual collisions is the responsibility of *event generators* or Monte Carlo generators, while their interaction with the detector is done through detector simulations. This chapter discusses the former, showing how much the pattern recognition and the theoretical calculations are intertwined.

6.1 About event generation

Event generators try to integrate the interesting phase spaces of the collisions, the ones with the highest momentum exchange are called the *hard process*. At the electron-positron colliders it was typically enough to do calculations on leading-order of the perturbation, since the electroweak coupling constant is indeed small and the contribution from subsequent orders can be neglected. In hadron colliders though, most processes happen via the strong interaction, where even next-to-leading-order (NLO) and next-to-leading-logarithm (NLL) calculations can be interpreted sometimes as just qualitative models. Since NLO generators require taking into account destructive interference between different Feynman diagrams with the same outgoing legs, they usually need matching criteria for the similarities. The next section aims to highlight the difficulty of this matching, because of the formation of *jets*. Above the hard process generators, there are dedicated generators for resonant decays, parton showers (see below), underlying events, hadronization, cosmic rays and ordinary decays.

6.2 Jet production

QCD being strong at low energies carries the consequence that its charges can not be directly observed. The observable states that participate in the strong interaction and can be used for experimentation are all bound states (technically, they belong to a different vacuum). Calculating cross sections of the scattering of these composite objects from first principles can only be done to a limited accuracy with lattice QCD. However, some modelling is possible which can be parametrized by experimental data. With the help of the factorisation theorem, the cross section calculation of such a process can be factored into high energetic and perturbative parts, and non-perturbative structure functions. The latter can be expressed with the parton distribution functions (PDF), belonging to the colliding hadrons [13][27][28].

For a collision where q momentum was transferred, the cross section can be expressed with scalar functions F that only depend on $Q^2 = -q^2$ and x, where x is the relative longitudinal momentum of the parton within the hadron. These $F_i^j(x, Q^2)$ are called the ith structure function [13] for the jth current, either the neutral or the charged. For large virtual momentum transfers, $Q^2 \gg M^2$, the structure function factors into a sum of the parton distribution functions of the individual parton types a convoluted with the C coefficient functions:

$$F_i = \sum_a C_i^a \otimes f_a \,,$$

$$C \otimes f = \int_x^1 \frac{dy}{y} C(y) f(\frac{x}{y}) \,.$$

These $f(x, \mu^2)$ parton distribution functions also depend on the renormalisation scale μ^2, which makes it valid only in a certain renormalisation scheme and perturbation order. Though the PDF can not be calculated *a priory*, but its dependence on μ^2 can be described with the help of the DGLAP equation, derived by Dokshitzer [29], Gribov, Lipatov [30], Altarelli and Parisi [31], hence its name. It is called an evolution equation, because the μ^2 dependence requires the P_{ab} splitting function, the probability of a given parton splitting into two others: For the $b \rightarrow a$ process it is

$$\frac{\partial f_a}{\partial \ln \mu^2} \sim \frac{\alpha_s(\mu^2)}{2\pi} \sum_b P_{ab} \otimes f_b \,.$$

With this differential equation, an experimentally measured PDF [32] can be extrapolated to either lower or higher scales, and it helps describing the decay of strongly interacting states. Due to the showering-like behaviour of these decays, the process is called fragmentation. The shower itself is called a jet.

Many models exist to describe the evolution of jets. For low mass partons the dominating subprocess is the soft gluon radiation (infrared divergence) and collinear splitting (new partons with approximately the same momenta) [9], while massive partons can have a so-called dead-cone [33], into which the gluon emission is suppressed. Calculations can be done with string, cluster fragmentation or with the colour dipole model. The formation of hadrons is another difficulty; it can be done with angular ordering of the daughters, but p_T ordering is also possible. The result is that jet fragmentation happens in numerous, topologically different ways, making the cross section calculations difficult, especially when detector acceptance has to be taken into account. Fortunately, many different jets give approximately the same detector response. This allows calculating many cross sections with Monte Carlo integration, generally using a hard process event generator, applying a shower algorithm and finally detector response simulation.

6.3 Jet finding

Without mystification, it must be said that jet finding is nothing more than the clustering of objects. A typical definition is, that a jet is a composite object with four momentum, that can be used to reconstruct the daughter partons in the hard process. The problem is that there is no objective definition of partons, they are a result of the calculation methods. Their definition depends on the applied perturbation, the order of the perturbation and on renormalisation scales. Therefore it is very important to use jet finding algorithms that are as free as possible of these effects, so different predictions can be compared [34]. This requires, that the cluster should be the same when the algorithm is applied to

- leading order partons;

- next-to-leading order partons (hard radiations);

- parton showers;

- clustered hadrons.

Theoretically this is possible when an algorithm satisfies the following:

- collinear safety :
  random split of a parton into two similar momentum partons are always clustered together by the algorithm;

- infrared safety :
  soft gluon radiation would not alter the cluster.

A comparison of various jet finders can be seen on fig. 6.1. Historically, the first jet finders were cone algorithms developed for electron-positron colliders like Petra and PEP [35]. The clustering was based on angle differences in (ϕ, θ), trying to maximise the momenta of particles in a cone with a fixed opening angle. Although this was rather easy to implement, it was neither collinear nor infrared safe. This is not a big problem at lepton colliders though, but still other algorithms were tested. The Durham algorithm satisfies the requirements, as do almost all *recombination* based algorithms. It aims to recombine objects into higher level objects, just like reversing the fragmentation, repeating this procedure until every particle is included into the hierarchy. A clear disadvantage though is that several *ad hoc* splitting scales have to be defined, which can typically alter the observed number of jets. An advanced attempt was the early k_t algorithm [36], that recombines objects not with the predefined scales, but when their distance is minimal compared to every other possible objects pairs. At hadron colliders the algorithm has to take the beam remnant into account, which typically acts as a stopping criteria for clustering. Jets in hadron collisions require an additional feature, the boost invariance in the z-direction, to handle the uncertainty of p_z of the centre-of-mass frame. This can be satisfied by using the transverse momenta, p_T, or the $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ pseudo-angle in the distance measure, with the approximately z-boost invariant *pseudorapidity* $\eta = -\ln\left(\tan\frac{\theta}{2}\right)$. The algorithm described above still has a caveat that must be addressed. It is somewhat sensitive to low energetic noise, which can be present as calorimeter noise or pile-up. Furthermore, the detector is not sensitive to objects below a certain threshold, which perturbs the possible clusters a bit. To circumvent this, the anti-k_t algorithms start the clustering with the highest momentum objects first, and reconstructs the hierarchy with a top-to-bottom approach [37].

6.4 More on parton showers

Parton showering can happen not just for the particles that leave the interaction point, but it is important to take it into account for the colliding particles, and for the spectators as well. Showering may occur before and after the hard collision, which are called as initial- and final- state radiation (ISR and FSR). The scale of the separation of these processes from the hard collision is purely arbitrary, as it is only there to simplify the calculations. The usage of ISR and FSR in calculations is to approximate higher order terms from QCD as in the Pythia event generator [38]. Although leading order QCD calculations fit into this picture easily, but NLO and higher order calculations double count them. To incorporate it, one must match certain outgoing parton showers of the hard collision with the final state radiation
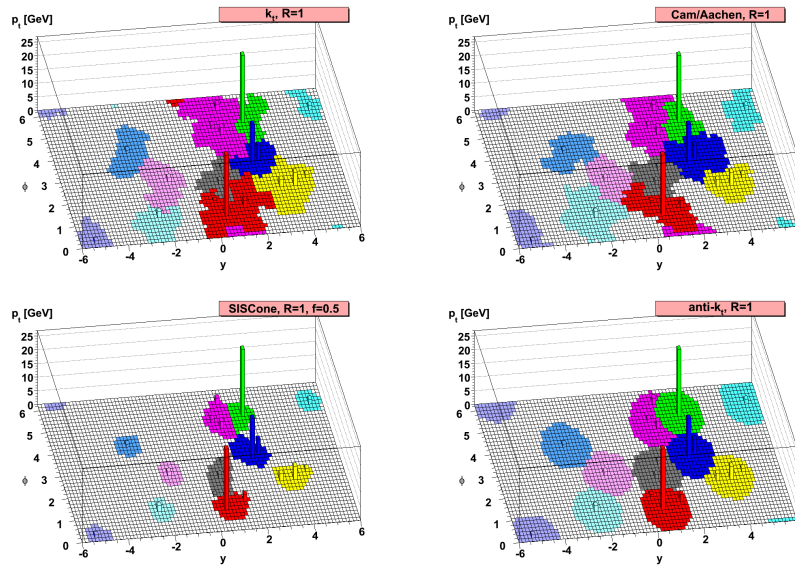
Figure 6.1: Different jet algorithms usually find the same number of clusters, but may differ significantly in the details. The figure shows four algorithms that were applied on the same simulated event, the ϕ and y axes are the coordinates of calorimeter hits. The k_T and the Cambridge/Aachen algorithm are very similar, both are based on the k-means clustering algorithm, and differ only by the definition of the weighted distance of two hits. The anti-k_T is like a k_T-algoritm, but builds clusters from top-to-bottom, and less sensitive to noise. The Seedless Infrared Safe Cone algorithm (SISCone) is an advanced cone finder method that incorporates a split-or-merge step to find a more probable boundary for a cluster.

to avoid the double counting. Matching with the ISR showers is not necessary, due to the way ISR is simulated.

A similar matrix element matching can be done for NLO event generation. The problem which it solves is that although a Monte Carlo event generator is supposed to integrate the cross section formulas, these integrals are inherently divergent. When properly regularised and renormalised, these integrals are finite, and the divergencies are known to cancel each other. The cancellation of the divergencies can be interpreted as a destructive interference between the infrared gluon radiation divergence and the virtual (loop) gluon emission. This cancellation can be avoided either by matching these phase spaces during event generation as in ALPGEN [39], or by introducing negative weights for the loop correction graphs, like in MC@NLO [40]. However, matching within the generator needs a specific choice of jet algorithm, which can be inconsistent with the later analyses. Event re-weighting is also possible, as is done in POWHEG [41][42][43].

Top-quark production and background processes

The top quark is the heaviest known elementary particle. Due to its high mass it has its own unique characteristics. This chapter summarises its discovery and highlights its properties in order to show why it was selected as a target of this analysis.

7.1 History of the top-quark

The first prediction of the top quark was made by Kobayashi and Maskawa in 1972 [44] , in order to explain the CP-violation observed in kaon decays. As they explain, the CP-violating term may arise when three generations of quarks are mixed, this is due to the difference between the mass and weak eigenstates. In short, it is a time-reversal symmetry violation caused by the interference between the six quark fields, because they propagate in space-time differently (the mass eigenstates) than how they interact with each other (the weak eigenstates). This causes a charge-symmetry violation, but since the weak interaction is already parity violating, it is altogether a CP-violation. At that time, only three quarks were known, and the hunt was still ongoing for the charm quark, predicted by the GIM mechanism [45].

The first evidence for the third generation of quarks was the discovery of the bottom quark at Fermilab in 1977 by Leon M. Lederman. Despite the heated expectations of the LEP collider at CERN, the top quark remained elusive until 1994. It was already known at that time that the top-quark mass must be larger than 77 GeV, but that year two evidences hit the light. The CDF group reported the observation (with low confidence limits) of the top-quark pair production with a mass around 175 GeV, and with great accordance with that precision measurements of electroweak vector boson masses and couplings were only consistent with a top-quark mass between 145 GeV to 185 GeV. The final, high confidence evidence came a year later, in 1995 by the two large Fermilab experiments CDF and D0.

The rediscovery of the top quark was the warm-up for the LHC, demonstrating the readiness of the collider and the experiments for future discoveries with the early 7 TeV collisions in 2010. Figure 7.1a shows how the cross sections of various processes change with the increasing energy changing from Tevatron to the LHC, and fig 7.1b reveals the good state of the LHC and the ATLAS detector by measuring several cross sections with good accuracy.
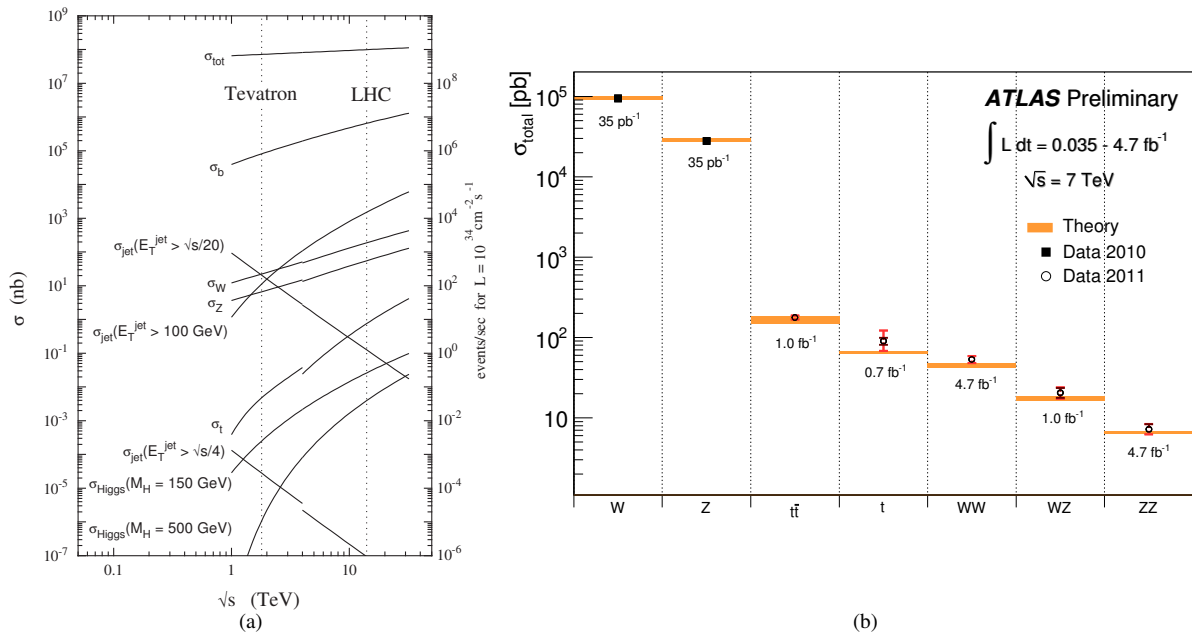
Figure 7.1: Cross section of various processes in proton-antiproton and proton-proton colliders[46], and cross section of various processes measured with early LHC data at 7 TeV collision energy.

7.2 The top-quark, the Higgs boson and new physics

The two main types of top-quark production are the top-quark pair production and the single top-quark processes. While the $t\bar{t}$ decays have a fairly large cross section and a distinctive topology, the single top-quark is just the opposite. At LHC energies its cross section is relatively small and its kinematic constraints are very similar either to the constraints of $t\bar{t}$ or to the various vector-boson processes. Figure 7.2 shows the possible single top-quark production processes: the two t-channel, the s-channel and the Wt-channel. The top-quark has a very short half-life and it is therefore not expected to form hadrons. The Cabibbo-Kobayashi-Maskawa (CKM) matrix leaves practically one decay channel for the top, via the weak interaction, $t \rightarrow Wb$, while every other mode is suppressed. Althouhg the W has five different decay modes, none of them are easy to detect. The $W \rightarrow qq$ is difficult because of the uncertainty of jet energy scales, hence such jets might be easily confused with the other jets in the event – the combinatorial background in identifying the W. The leptonic decays include a neutrino $W \rightarrow l\nu$, which only appears as missing momentum of the full system, which has even larger uncertainty than the jets. This makes finding the Wt channel the hardest of the four, and the compromise is to look at the lepton+jets channel, for which the jet combinatorics are low and there is only one neutrino to deal with [13][47].

The main background for this single top-quark processes are then the ones which produce one energetic lepton, missing energy and multiple jets. These include:

fake leptons

- multi-jet production from QCD

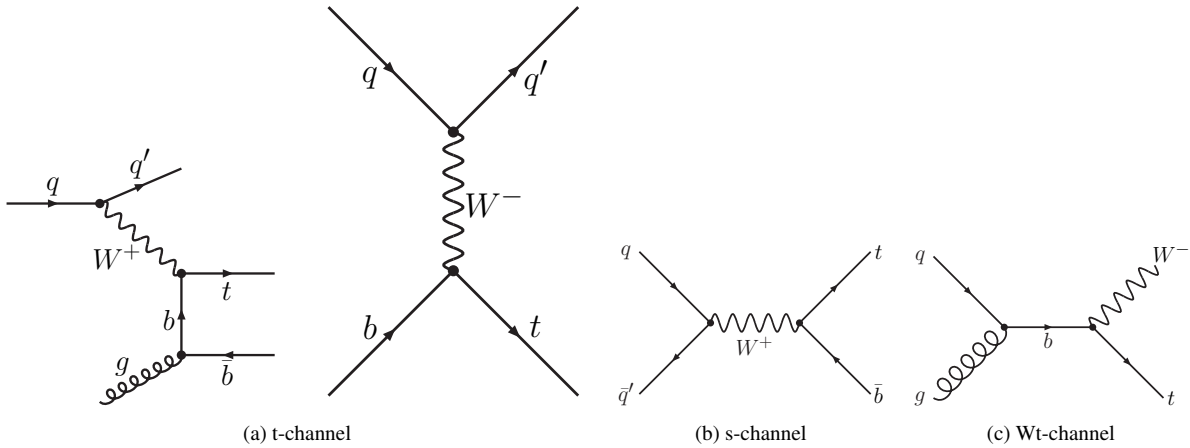- $Z \rightarrow qq$ + jets

- $W \rightarrow qq$ + jets

Figure 7.2: The Feynman-diagrams of the dominant single top generating processes. Since the t-channel production requires a *b*-quark in the initial state, which is not present in the protons, the LO and NLO contributions have comparable cross sections. These may happen without an on-shell *W*-production, hence they have less kinematical constraints than the s-channel and have larger chance to occur. The *Wt*-channel defines an on-shell *W* and a *t*-quark in the final states which requires a highly off-shell *b*-quark, making it the less probable one and the hardest to find.

prompt leptons

- $t\bar{t} \rightarrow W^+ b W^- \bar{b} \rightarrow l\nu b\bar{b} +$ jets

- $W \rightarrow l\nu +$ jets

- $WW/WZ \rightarrow l\nu +$ jets

The processes that fake lepton identification need to be treated differently. Since these processes are multi-jet events originally, their individual cross section is rather uncertain. A systematic uncertainty also emerges here because the description of faking lepton recognition heavily depends on the understanding of the detector. The background from the multi-jet processes are better estimated from the data itself then, applying techniques that estimate the fake lepton efficiencies for a given channel by slightly loosening the lepton identification criteria.

The single top quark productions are sensitive to the V_{tb} coupling of the CKM matrix, which is otherwise inaccessible, therefore it is an important tool for testing the Standard Model and looking for new physics [8]. Furthermore, as the top-quark is the heaviest known elementary particle yet, it is known to have the strongest Yukawa coupling to the yet unconfirmed Higgs boson, which introduces slight, logarithmic dependence into the cross sections of various processes through renormalization [48]. The precise measurements of the single top-quark production cross sections indirectly help modelling the backgrounds for the $t\bar{t}$ cross-section and mass measurements, leading to an indirect determination of the Higgs mass.

Although the t-channel and the s-channel single top-quark production was confirmed by Fermilab, the proof needed a combined effort of several groups and training multivariate techniques with altogether more than one hundred input variables [49][50]. The cross-section of single top-quark processes are relatively higher at Fermilab than at the LHC [51], nevertheless the much higher luminosity and statistics makes it somewhat easier to detect at the LHC.

LHC and ATLAS

The only place with the technology that is currently capable of producing and detecting top quarks resides at CERN, the European Organisation for Nuclear Research. This chapter aims to describe architecture of the Large Hadron Collider and the ATLAS detector and the engineering reasons behind it.

8.1 Overview of the Large Hadron Collider

The building of the Large Hadron Collider (LHC) was approved in 1994 by the CERN Council with a proposed proton-proton centre of mass collision energy of 14 TeV. This was a year after of the cancellation of the 40 TeV Superconducting Super Collider in the USA and a year before Tevatron's discovery of the top quark.

The LHC was built in the 27 km circumference tunnel of the former Large Electron-Positron Collider (LEP) on the border of Switzerland and France. As the proton mass is much larger than the electron mass, the energy loss via synchrotron radiation is smaller than for electrons at the same energy and radius. This makes possible to achieve more than an order of magnitude higher collision energy than at LEP, where it was around 200 GeV. The centre-of-mass energy of the LHC in 2011 was 7 TeV and was increased to 8 TeV. The nominal 14 TeV collision energy is planned to be reached in 2014.

A chain of accelerators provides the protons to LHC [52][53]. In order these are the LINAC2, PS BOOSTER, Proton Synchrotron (PS), the Super Proton Synchrotron (SPS) and finally the LHC, what gradually increase the beam energy. An overview can be seen on fig. 8.1.

Nearly 1300 dipole magnets are used to keep the protons on a circular track, while the 400 quadrupole magnets are to focus the beams. Small sextupole and octupole corrector magnets are attached to the larger dipoles and quadrupoles. In order to maintain the needed 8.3 T magnetic field for the 7 TeV and half for the 3.5 TeV beams, superconducting niobium-titanium magnets are used, cooled to 1.9 K with high pressure superfluid helium. The protons in the beams are in bunches, and recently in bunch trains. These can cross at four interaction points (IP), where the the four major experiments reside, the A Toroidal LHC ApparatuS (ATLAS), the Compact Muon Solenoid (CMS), A Large Ion Collider Experiment (ALICE) and the Large Hadron Collider beauty experiment (LHCb). ATLAS and CMS are general purpose detectors, and as can be seen on fig. 8.2, they are ready to observe every aspects of proton-proton collisions and to detect anything unexpected. LHCb is still a large experiment, but built for a specific goal. It aims to precisely measure the CP violation and rare decays of *B* hadrons,
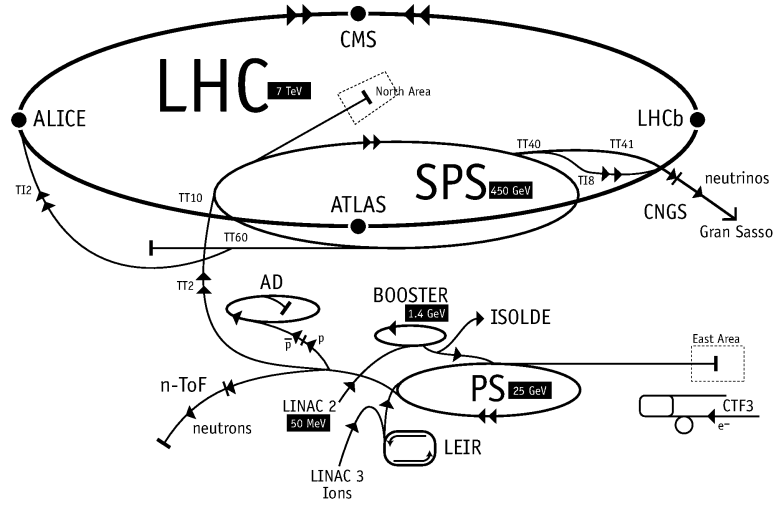
Figure 8.1: The LHC accelerator complex

which can be very sensitive to new physics. ALICE, as its name suggest was built to observe heavy ion collisions and look for collective quark phenomena, like the quark gluon plasma that may appear at high temperatures. Nevertheless it needs proton-proton collisions for calibration, and is able to provide more general results as well. Two detectors are suited to observe the physics of forward scattering and provide additional luminosity measurements, Totem and LHCf. The former is near to CMS, while the latter is close to ATLAS. Lastly, MoEDAL is an experiment that can detect magnetic monopoles.

LHC aims not only for record breaking energies, but for unprecedented luminosity as well, as that is needed for studying processes with small cross sections. The goal is $\mathcal{L} = 10^{34}\,\mathrm{cm\,s^{-2}}$, nearly a hundred times bigger than what was achievable at Tevatron. The luminosity is a function of the following [54]:

$$\mathcal{L} = \frac{fN^2}{4\pi\sigma^2}\mathcal{F}(\theta)\,,$$

with the parameters

f = bunch collision frequency,

N = number of particles per bunch,

σ = root mean square transverse beam size at IP,

$\mathcal{F}(\theta)$ = luminosity reduction factor due to crossing angle, θ.

Here $\mathcal{F}$ is a constant parameter $\sim$85%, defined during the design. The bunch frequency is $f = 40\,\mathrm{MHz} = 1/25\,\mathrm{ns}^{-1}$, which follows from the 7.5 m distance between the bunches. Generally beam size at the interaction point should be as small as possible, it is different for heavy ion physics to artificially reduce luminosity. It is because the high number of particles in a bunch, which can be around $N \approx 10^{11}$ at the LHC, compressed into small volumes has the side effect that more than one high inelastic collision might happen at the same bunch crossing. This effect is called pile-up, and should be handled with care in the analyses even in proton proton collisions. In heavy ion collisions pile-up is avoided by decreasing the luminosity, otherwise the particle multiplicity would be too high for the detectors. In the same manner, the luminosity is decreased for LHCb.
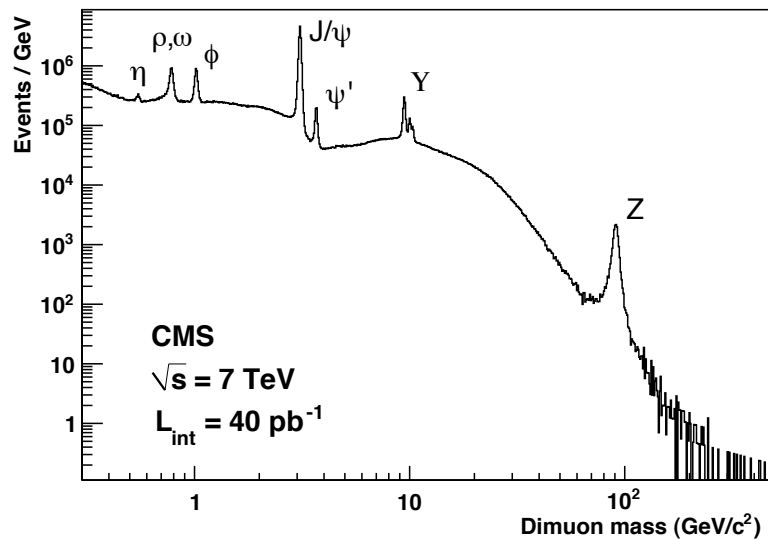
Figure 8.2: A plot from CMS, demonstrating the capability of LHC by rediscovering the Standard Model from the early data.

8.2 ATLAS overview

A general purpose detector is aimed to be able to reconstruct the high energy processes at the collision point. Eventually this requires the ability to reconstruct the four momenta of the known particles at high energies. One can only detect directly the long lived particles, that travel far away from the IP and leave a significant trace in the detectors before their decay. Typically these are the electrons, muons, photons, some of the mesons from which the most abundant one are the pions, mesons with *b*-quark content and in some sense the neutrinos. It is possible to reconstruct the track of charged particles, measure the their energies and their inclination to interact with different types of matter. In this way, electrons can be distinguished from photons because they leave tracks in the dedicated detectors. Pions differ from photons and electrons since they interact hadronically, and they may create showers. Muons, due to their higher mass than electrons do not brake down so easily in matter, and since they do not participate in the strong interaction they avoid creating showers and pass through most detectors by just leaving a simple trace. Hadrons containing a *b*-quark sometimes decay far enough from the IP that their decay vertices can be directly seen within the tracking devices, but otherwise these decay vertices can still be distinguished from the vertex of the collision point.

For centre-of-mass collisions the ideal detector would be spherically symmetric to be ready for the spherically symmetric collision outcome. However, the centre-of-mass system in *pp* collisions is rather uncertain, since it is not the protons that collide directly, but their constituents. The protons consist of many particles, all of them carrying a fraction of the full momenta, which gives an uncertainty of the sum of momenta for the colliding partons in the direction of the beam. This means that the original spherical symmetry is convoluted with a translational symmetry along the beam pipe, which overall resembles to a cylindrical symmetry. A practical reason why ATLAS was built approximately cylindrically symmetric is that it had to be built around the beam pipes. Another physics reason is that due to the mentioned compositeness of protons, collisions can be either elastic or inelastic. Elastic collisions only alter the direction of the flight of the protons which tend to fly approximately parallel to the beam direction, whereas inelastic collisions convert a more significant part of the beam energy into the creation of new particles providing a window to high energy physics. These particles tend to be created with a momenta
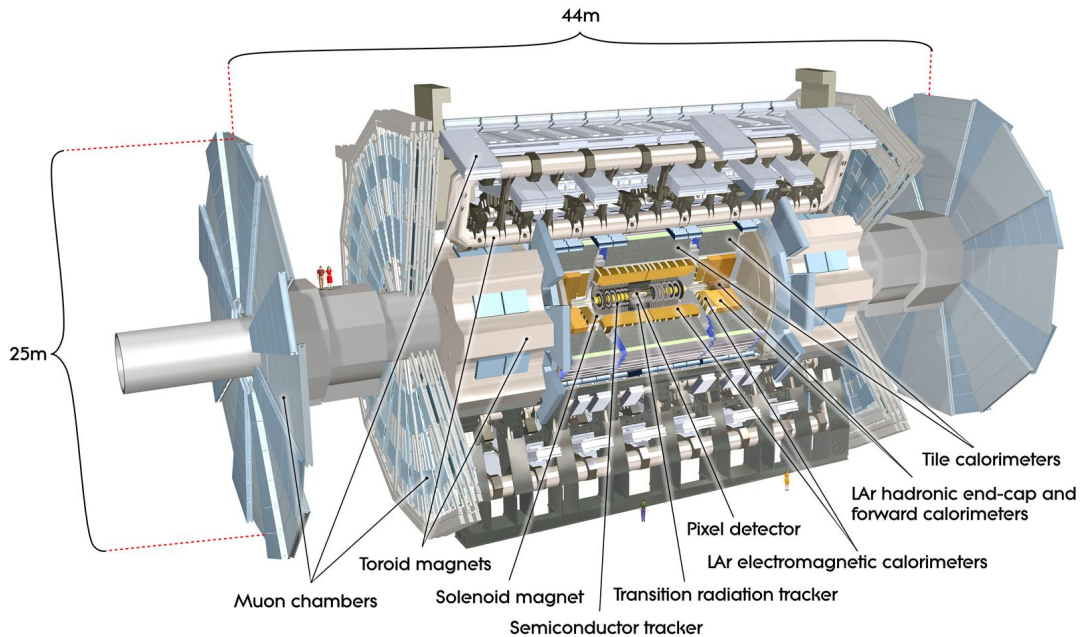
75

Figure 8.3: Overview of the ATLAS detector

perpendicular to the beam pipe, because of the compositeness the protons, giving an additional benefit of cylindrical detector. These resulted in that ATLAS was constructed with a cylindrical part with layers parallel to the beam pipe called the barrel, and with cylinder sections with layers perpendicular to the beam pipe, the so called endcap detectors. In fig. 8.3 this structure is apparent in most of the subdetectors.

The tracking is done by the inner detectors (ID), which is also responsible to detect the exact collision point. The detectors are under a considerably high 2 T magnetic field, which under the charged particles bend, revealing the sign of their charge and their momenta. The measurement of the energy is the task of the calorimeters, but as different particles interact differently with ordinary matter, one needs to measure more than just the deposited energy. ATLAS contains an Electromagnetic Calorimeter (EMCal) and Liquid Argon Calorimeter (LAr), with high enough granularity to help determining the direction of the particles, which can guide the tracking algorithms where to look for high energetic particles. Typically, high energetic muons can still escape from these detectors. Their identification and momentum measurement is helped by the outermost Muon Spectrometers (MS).

The magnetic field in ATLAS has two sources. The inner detectors are surrounded by a superconducting solenoid magnet, which provides an approximately homogenous field, ideal for momentum determination of particles. Air-core toroids with eightfold symmetry can be found at the outer parts of the barrel, while two smaller disks of magnets are at the endcaps. The reason behind this complexity is to direct the muons through the detector material instead of the dead materials.
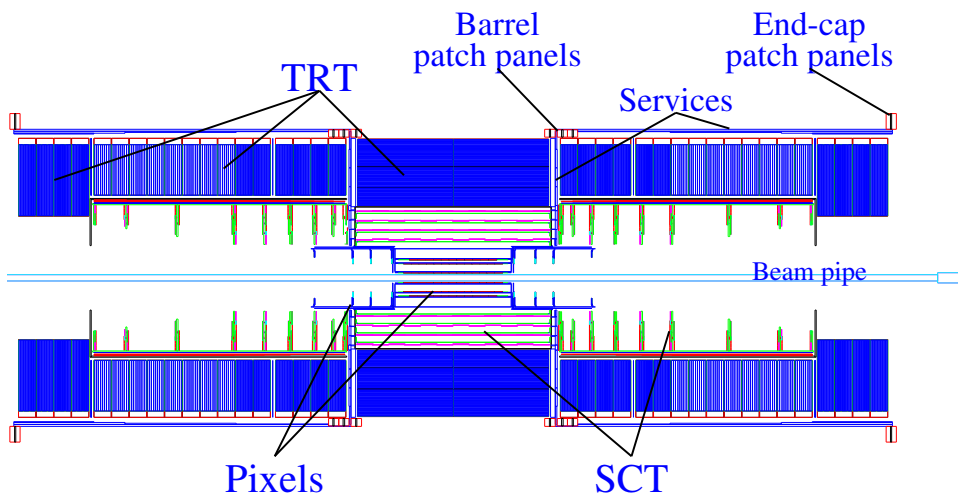
Figure 8.4: Longitudinal view of the Inner Detector

8.3 Inner detector

The inner detector in ATLAS performs many tasks at the same time [55]. Through detecting the positions of individual interactions of the charged particles within its volume, the recorded hits can be combined into tracks. These are important in not just determining the momenta of the particles, but also to reconstruct the vertex where the collision happened. The hits have a $10\,\mu\text{m}$ to $500\,\mu\text{m}$ precision, while the position of the vertices can be reconstructed with approximately $50\,\mu\text{m}$, see [47]. This is possible mainly with the help of the innermost of the three inner detectors, the Pixel Detector. The precision is needed not just to distinguish the vertices from pile-up events, but also to find displaced secondary vertices. These are produced by the decay of relatively long-living mesons, hence it gives an important hint of the flavour of quarks involved.

Farther from the primary vertex the Semiconductor Tracker (SCT) encapsulates the pixel detector, see fig. 8.4. It is almost as precise as the Pixel Detector, and due to the smaller radiation it ages more slowly, which is an important feature in the view of the expected long running time and high luminosity of LHC. The intriguing outermost sub-detector of the Inner Detector is the Transition Radiation Tracker (TRT), which gives not only additional precision hits for tracking but also particle identification hints.

Apart from precision tracking, vertex detection and particle identification, the overall tracking performance should have a good enough performance to be able to tell if there was no track at all at a given position. Many particles, such as photons and neutral pions, mesons and hadrons deposit their energies in the calorimeters, but being neutral they expected not to leave significantly many hits in the trackers. One especially important case is with the neutrinos and weakly interacting particles predicted by exotic theories, which particles can only be detected through the unbalanced transverse momenta or missing transverse energy $\not{E}_T$. In such a case the tracker plays an important role in deciding if the $\not{E}_T$ was really from a neutral particle or caused by an error or by a hole in the calorimetry.

8.3.1 The pixel detector

The innermost detector is the pixel detector with an inner radius of $5\,\text{cm}$. Its cylindrical structure built from layers containing pixel modules. These modules are $21.4 \times 62.4\ \text{mm}^2$ semiconducting silicon chips, each with 16 instance of 120×24 pixels with a resolution of $50\,\mu\text{m}$ in the plane perpendicular to

the beam, and 400 μm in the *z* direction. The detector consists of three layers and covers $|\eta| < 1.5$. The barrel layers are approximately cylindrical, while the end-cap modules are disks, closing the cylinder. The innermost layer is called *b*-layer, reminding its importance in *b*-tagging, and it is designed to be replaceable. With the two additional layers with larger radius, the pixel detector gives a high chance of three hits for every charged particle. The outer radii are 12.2 cm for the barrel and 14.9 cm for the end-caps. The full detector has 80.9 million pixels. A much larger pixel detector would come with unreasonably high *material budget*, meaning that the increasing readout and maintenance infrastructure would interfere too much with the particles passing by.

8.3.2 The semiconductor tracker

Unlike the pixel detector the particle detection in the SCT is based not on pixels but on silicon stripes. This allows a much more accurate spatial resolution of 16 μm in the $R\phi$ plane, while it is still 580 μm in *z*. Similarly to the pixel detector, the modules in the barrel region lie on the cover of the cylinder, the 12.5 cm long strips being approximately parallel to the *z*-axis. The geometry of each module is 126 mm × 63.6 mm, but a second module is glued to the back of each module rotated with a 40 mrad angle, making possible to measure the *z*-direction with the mentioned precision. The barrel consists of four cylindrical layers, with radii from 30 cm to 70 cm and a length of 149 cm, altogether having 12 modules along the *z* direction. The end-cap parts form disks, with slightly overlapping trapezoid-shaped modules, which are two-sided, similarly to the modules in the barrel. Each end-cap disk consists of four concentric rings. The inner radii of these rings are adjusted in order that they cover $|\eta| < 2.5$. The detector provides 6.2 million readout channels altogether.

8.3.3 The transition radiation tracker

The TRT also consist of stripes, but the particle detection is based on a different physical process as it is a straw-acttube tracker. The TRT is made of grounded tubes containing wires on high voltage, filled with a neutral gas that can be easily ionised. Similarly to a Geiger-Müller counter, when a charged particle flies through this gas, it triggers an avalanche of ionisation. It was named after the transition radiation effect, since it detects the photons radiated by the relativistic charged particles passing through the boundary of two media with different dielectric constant.

The maximum straw length is 150 cm with only 4 mm of diameter, and there are around 50 thousand of them in the barrel. The endcap takes a larger part in tracking, as it has 320 000 of straws. The barrel straws, just like in SCT are lying parallel to the *z*-axis, and are organised into modules with a triangular cross section. The barrel cylinder covers the radii from 56 cm to 107 cm, while most of the 36 end-cap wheels are between 64 cm to 104 cm maintaining tracking hits in the $|\eta| < 2.5$ range. The high number of straws assures an average number of 20 to 36 hits per tracks, each with a 170 μm precision. This precision is a result of not just measuring the fact of transition, but the timing of the hit and the shape of the response. TRT employs a non-flammable gas mixture of 70% Xe, 27% CO_2 and 3% CF_4. The Xe is also useful in the electron identification, as it absorbs the transition radiation photons, while the CF_4 enables faster drift-time for electrons. These give a stronger signal that discriminates electrons from pion tracks. The CO_2 stabilises the mixture of the two former gases under high voltage. Unlike the other detectors, the TRT requires uniform temperatures among the modules for stable performance, which is solved by circulating an additional cooling gas within the system.

8.4 Calorimetry

The calorimeters in ATLAS were designed to cover a large rapidity range $|\eta| < 5$, and provide good granularity and energy resolution for various physics scenarios, including the invariant mass reconstruction of various Higgs decays and exotic weak vector bosons, the Z' and the W'. These require the precise energy measurement of many different particles: photons, electrons, muons, taus, light and heavy mesons and baryons and much more complex structures, like jets from different origins [56]. The main difference in these classes is that some of them interact not only with the atomic electromagnetic field of the detector material but also with the nuclei. These particles do not deposit their full energy in a calorimeter since they tend to produce hadron showers that contain particles which may decay further and produce neutrinos, or high energy muons. While the energy carried away by them can not be measured within the size of ATLAS, still the shape of the showers carry information about the energy and origin. These also lead to the distinction between electromagnetic and hadronic calorimetry, depending on the main type of interaction involved.

To gain the most possible information about the particles leaving the inner detector, the first they meet is the electromagnetic calorimeter and more outside the hadronic calorimeters. The relative energy resolution typically goes by the following formula:

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} + b \,,$$

where E is the energy of either a particle or a jet, $\sigma(E)$ the root mean square of the reconstructed energy probability. The constant a is called the statistical term, while b is the constant term. For high energies the constant term will dominate, for electrons it is around 1%, for jets it is around 3% for the barrel and end-caps, while 10% for the forward detectors, dominated by the hadronic calorimetry. The statistical term for electrons is rather low, at the level of 10%, but for jets it is 50% in the central region and even 100% in the forward detector. An important question is the linearity of the calorimeter response with respect of the incident energy, where the detector performs rather well. An important variable in calorimetry is the *radiation length* X_0 for every detector material, used in the calculation of the average energy loss by bremsstrahlung of the charged particles flying through.

8.4.1 Electromagnetic calorimeter

The electromagnetic calorimeter covers a large rapidity region of $|\eta| < 3.2$ and it is outside the inner detector and the 44 mm thick superconducting solenoid magnet. It is a sampling calorimeter, having alternating layers of lead, liquid argon and read-out electronics in an accordion shape. A particle coming from the direction of the collision point first faces a lead pre-sampler layer, creating an initial electromagnetic shower. The lead was chosen as absorber because of its high atomic number, so charged particles lose significant amount of energy by flying through. The signal is created in the liquid argon, as it is ionised by the shower. The EC has three different readout layers, the closest to the interaction point is the pre-sampler with $4.3X_0$, this is relatively thin compared to the second layer with $16X_0$, and the last one is used as a trigger in bunch crossing identification and has thickness of $2X_0$. The barrel covers the rapidity region $|\eta| < 1.475$ and the end-cap is between $1.375 < |\eta| < 3.2$. The granularity of the detectors are different in the sampling layers and in comparing barrel with end-cap. The barrel presampler has the finest resolution in $\Delta\phi$, but the worst in $\Delta\eta$. In terms of $\Delta\eta \times \Delta\phi$, it is 0.003×0.1 in the main sampling layer, while it is 0.025×0.025 and 0.05×0.025 in the middle layer and in the trigger sampler respectively. The resolution of the end-cap is similar in $\Delta\eta \times \Delta\phi$, but it changes as a function of η, optimised for physics reconstruction. The full end-cap contains approximately 200 000 channels,

and this data helps seeding the tracking algorithms and the particle identification.

8.4.2 Tile calorimeter

The tile calorimeter is the hadronic calorimeter occupying the outer barrel regions up to $|\eta| < 1.6$. It consists of alternating layers of steel absorbers and scintillators, read out by photomultiplier tubes. The cylinder spans from radii 2 m to 4 m, segmented into a barrel in $|\eta| < 1$, and two extended barrel regions in $0.8 < |\eta| < 1.7$. Although iron is not as resistant to radiation as copper, the integrated dose is much smaller than in the high rapidity regions (the endcaps and the forward parts), so it was safely chosen as cheap absorber. The resolution of the tile calorimeter is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$.

8.4.3 Liquid argon calorimeter

Hadron calorimetry in the endcap region is done by another liquid argon calorimeter, the Hadronic End-Cap (HEC) which shares many things in design and infrastructure with the electromagnetic calorimeter [57]. It consists of two wheels, made of different thickness of copper plates. Its gaps are filled with liquid argon, and covers the rapidities $1.4 < |\eta| < 3.1$. The readout is done with kapton-insulated copper plates, kept in place with a honeycomb structure. The granularity $\Delta\eta \times \Delta\phi$ worsens as a function of $|\eta|$ from 0.1×0.1 to 0.2×0.2.

8.4.4 Forward calorimeter

The very high rapidity region is only covered by a hadron calorimeter, the Forward Calorimeter. It consists of three disk-shaped liquid argon detectors at each side of ATLAS, between $3.0 < |\eta| < 4.9$. The volume of the detector is rather small, so to maintain the required resolution in η, it is a relatively dense detector. The disk closest to the interaction point uses copper as an absorber, while the further disks use tungsten. The detection is done in a hexagonal grid structure of tubes drilled into the disks, pointing to the collision point. Altogether it provides 11 288 channels.

8.5 The muon spectrometer

The muon tracking and momentum measurement rely strongly on the magnetic field structure. The barrel solenoid, the eight barrel toroid and the end-cap toroids create a complex magnetic field, while the muon trackers are aligned in a way to be able to measure the coordinates of the tracks in the principal bending direction for the most precise momentum measurement. The emphasis on muon tracking is because they can provide the cleanest access to many physical processes as they may be produced in the hard collision with high momenta. Nevertheless they only slightly interact with the detector materials at the expected energies and lose only a small portion of their momenta. That is why the muon detectors are trackers and they are the outermost with the arrangement as fig. 8.5 shows. The other detectors filter everything else, with the exception of the neutrinos and occasional high energetic jet punch throughs. The central region, $|\eta| < 1$, is dominated by the field of the barrel toroids, the $1.4 < |\eta| < 2.7$ region by the endcap-toroids, while in between $1.0 < |\eta| < 1.4$ is the interplay of both. This was optimised in a way that the magnetic field is mostly orthogonal to the muon tracks in the most interesting energy ranges. In the barrel region lies three cylindrical layers of Monitored Drift Tubes (MDT), while the large rapidity regions are covered by Cathode Strip Chambers (CSC). These latter despite being much closer to the interaction point, have finer granularity to cope with the larger number of background and to provide better precision.
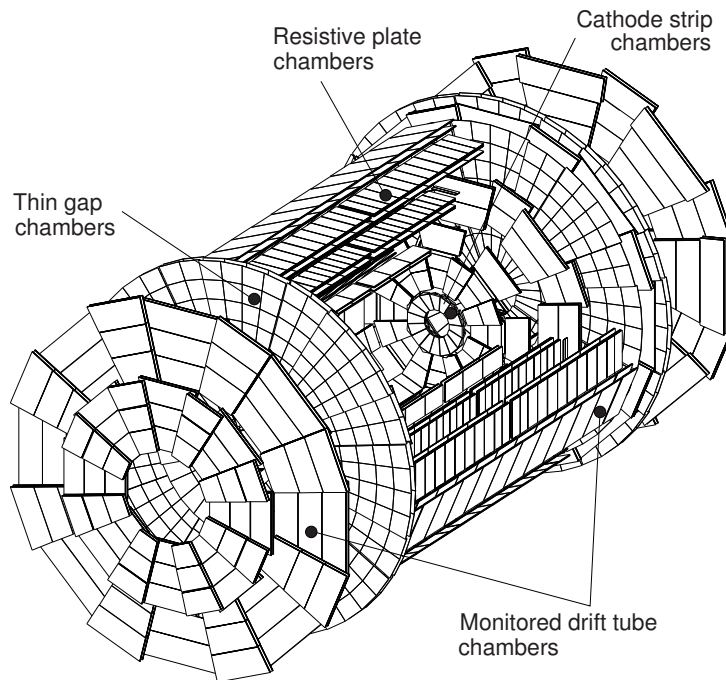
Figure 8.5: Arrangement of the muon detectors

Additionally there are two more types of muon detectors, both are dedicated for triggering. Around the barrel stand the Resistive Plate Chambers (RPC), while up to $|\eta| = 2.4$ in the endcap the Thin Gap Chambers (TGC) are located. Both can measure a two coordinate component of the tracks.

8.6 The trigger system

ATLAS has to face a 40 MHz stream of collisions, each intrinsically different, though most of them are out of interest. Reading out all the channels at this rate would be impossible with current technologies, not to mention the storage demand. To deal with this amount of information, a three level trigger system was implemented that decides which events are worth to store. The Level 1 Trigger (LVL1) was built directly into ATLAS, and uses the dedicated trigger detectors from the electromagnetic calorimeter, hadronic calorimeters and from the muon detectors [58]. It must be noted that the inner detector can not participate in the trigger decision, as it is computationally infeasible to reconstruct tracks for every event. The LVL1 algorithm reconstructs the following objects:

- local objects
  - muons
  - EM-clusters, isolated or not
  - narrow jets, useful for reconstruction of taus or single hadron decays

> – jets

- global objects

  > – missing transverse energy

  > – scalar transverse energy

The existence of these objects with certain minimal energy or transverse momentum can be used as a LVL1 trigger. These objects can then be used by the High Level Triggers [46] (HLT), which is the common name for the Level 2 Trigger (LVL2) and the Event Filter (EF). The LVL2 is based on forming Regions of Interest (RoI) around the interesting LVL1 objects, and accessing all the detectors within this (η, ϕ) region, including the trackers, which enables it to refine the LVL1 decision with better granularity and higher quality objects. This can further decrease the rate of events leaving the detector. Finally, these objects can be fed to the Event Filter, making a decision if the event belongs to a class that is worth to be recorded. To avoid event shaping, namely that physicists only look at events that look like what they prefer, the triggers are typically very loose, being based on p_T thresholds, multiplicity of objects and so on. The measurement of trigger acceptance and efficiency is non-trivial, since it can only be based on understood physics objects. A consistent method to measure acceptance and efficiency from data is the bootstrap method, which starts from minimum bias triggers and gradually tightens them with increasing the thresholds when the previous one was well understood. Another method is based on independent event signatures, where one understood trigger can be used to study an different one, though it also needs accurate physical description. A third one, based on di-object is basically a combination of the previous two, called as tag-and-probe method. Its credibility is based on the assumption that certain pair-production processes are dominating, hence triggering on one of the objects can be used to test the trigger efficiency on the other. Such processes are di-jet production in QCD, $Z \rightarrow ll$ and $J/\Psi \rightarrow ll$ decays for example.

8.7 Luminosity measurement

Data taken with ATLAS is divided into chunks, called Luminosity Blocks (LB). This segmentation is necessary for maintaining the same data quality, since many detector elements may suffer from temporary high voltage drops, dead times, data acquisition problems and the luminosity gradually decreases during a run period too [59][60].

The luminosity measurement comes from many sources, because all of them are model dependent. Their combination is used to determine and decrease the overall systematical uncertainty. The main methods available are the following:

- **Roman pots**
  The Roman Pots were considered to be a milestone in luminosity measurement in hadron colliders in the 1970s. The three roman pots – one horizontal and two vertical – consist of a vacuum chamber and a movable beam position monitor (BPM). They are attached to the beam pipes in the very forward region of ATLAS, 240 m from the IP [61]. The idea is to measure the intensity of particles in the beam that elastically scatter in a very small angle [62], based on the optical theorem [63]. That formula very simply connects the total cross section σ_{tot} of the interaction to the imaginary part of the forward scattering amplitude f, for momentum k.

$$\sigma_{\text{tot}} = \frac{4\pi}{k} \text{Im} f(0) \,.$$

This can provide a very precise absolute luminosity measurement, an accuracy of $\Delta\mathcal{L}/\mathcal{L} = 2\,\%$ to $3\,\%$. It can also operate as a Level 1 trigger, however it can only be used in the very low luminosity runs, at $10^{27}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ to calibrate the other methods.

- **LHC machine parameters**
  Both relative and absolute luminosity measurement can be done with the estimation of the LHC beam parameters. What have to be known are the number of protons in a bunch and in a beam, the transverse beam dimensions β and emmittance, and finally the transverse displacement of the beams. Here the emmittance measures the compactness of the beam in phase space, combining the physical cross section with the momentum spread. Its uncertainty dominates the measurement, and it is about $5\,\%$ to $10\,\%$. The physical width of the beams are measured via *Van der Meer* scans.

- **W/Z counting**
  During high luminosity runs, when $\mathcal{L} > 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, the weak vector boson production can be used for online luminosity monitoring. Depending on the uncertainty of the parton distribution functions and the $W^{\pm}/Z^0$ production cross section it can be used either for relative or absolute luminosity measurements. The reconstruction of leptonic decays need understanding of the detector, which is expected to get better over time, and new PDF measurements can also be taken into account. The uncertainty of the absolute measurement is around $\Delta\mathcal{L}/\mathcal{L} = 10\%$, while many systematic uncertainties cancel for the relative luminosity determination. At the highest luminosities the statistical precision can be as good as $5\,\%$ in a $10\,\mathrm{s}$ measurement, or $1\,\%$ after $3\,\mathrm{min}$.

- **Dilepton counting**
  The cross section of $\mu\mu/ee$ production from photon-photon fusion is known from theory within 1% uncertainty, though it is rather low. For central muons above $3\,\mathrm{GeV}$ it is around $1\,\mathrm{pb}$. With good understanding of the needed objects, and knowledge of the acceptance, the absolute and relative integrated luminosity can be measured with $2\,\%$ precision.

- **LUCID**
  $17\,\mathrm{m}$ from the interaction point along and around the beam pipe, but still within ATLAS stations the LUCID, the Luminosity measurement using Cherenkov Integrating Detector [64]. It consists of small, gas filled tubes pointing to the interaction point, to measure the flux of forward scattered particles through the integrated Cherenkov light with photomultipliers. This detector at $|\eta| \sim 6$ can also act as a LVL1 trigger, and monitor the relative luminosity at a bunch-by-bunch level.

- **Beam Condition Monitor**
  BCM is a small, $1 \times 1\,\mathrm{cm}^2$ thin artificial diamond sensor, laying directly on the beam pipe, only $1.8\,\mathrm{m}$ away from the interaction point and $5\,\mathrm{cm}$ from the centre of the beam. It can measure the number of charged particles in $3.9 > |\eta| > 4.1$, which typically come from inelastic collisions, providing information about the relative luminosity.

- **MBTS**
  The Minimum Bias Trigger Scintillator is a 2×8 scintillator grid in front of the LAr end-caps. Its main purpose is to provide a LVL1 trigger for minimum bias events, with an acceptance in $|\eta| = 1.9$ to 3.8. Since it is damaged by radiation, it can only be used for relative online luminosity monitoring.

- **TileCal**
  The photomultiplier tubes in the hadronic calorimeter in the barrel are sensitive to the level of

minimum bias events. The integrated anode current provides a relative luminosity monitoring, without a need for a trigger.

- **LAr**
  Similarly to the TileCal, the Liquid Argon calorimeter can act as a luminosity monitor too. The relative luminosity measurement is done by recording the current of the high voltage system.

CHAPTER **9**

The data acquisition infrastructure of ATLAS

The important hardware components of data acquisition in ATLAS were described previously in Section 8.2. This chapter summarises the additional external infrastructure and the online and offline software necessary for data taking and physics analysis, based on the technical design reports of the trigger system [46] and computing [65].

9.1 System components and functions

9.1.1 The data flow system

Every subdetector of ATLAS has a readout driver (ROD) that behaves as an interface between the detector and the external infrastructure, where the connection is done with over 1600 readout links (ROL). The LVL1 trigger decisions are made within the dedicated detectors, and the information on the selected events are transferred from the RODs via the ROLs into the readout buffers (ROB). These buffers serve the information for the LVL2 trigger of the HLT system. The accepted events are then transferred to one of the ~100 computers in the Sub-Farm Interfaces, which builds the event from the fragments, and conveys it to the second element of HLT, the Event Filter (EF). All the events accepted by the EF are sent both for permanent storage to the Sub-Farm Output (SFO) and for offline reconstruction for further analysis.

9.1.2 The HLT system

The LVL2 trigger has an average time of 10 ms for every event that passed LVL1. It uses highly optimised, limited precision algorithms to make a decision on the fraction of the full event data, as it typically uses only 2% of it. The several Regions of Interests (RoI) identified by the LVL1 trigger are combined together by the RoI Builder (RoIB) algorithm, so LVL2 can make a global decision on the event. The necessary information are dispatched by the LVL2 Supervisors (L2SV), and the LVL2 node allocated for the event requests the ROBs themselves.

The EF receives the LVL2 accepted events, which are here subject for a more precise and computationally extensive analysis that takes 1 s on average. These events are fully built by the SFI, containing all the information that was available for the LVL2 decision and detailed object informations like p_T, E_T and isolation variables. It performs more sophisticated object reconstruction and calibration to provides a more complex decision.
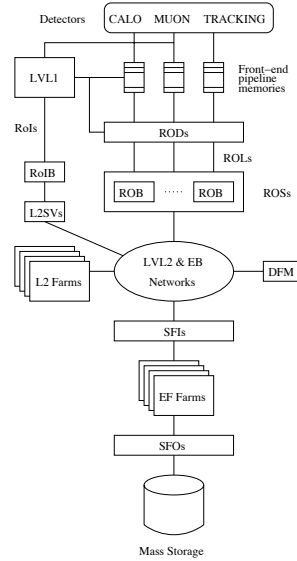
Figure 9.1: The main components of the Data Flow system and HLT

9.1.3 The online software system

The Online Software system is an architecture that is responsible for the configuring, controlling and monitoring of the TDAQ system. Furthermore, it defines an interface between the DAQ, HLT and Data Control System (DCS), and it is responsible for some information distribution service and gives access to configuration and other databases of meta-data, but does not handle information about individual events.

It plays a role of a supervisor for the entire Data Flow system, responsible for the synchronisations and for the proper start up and shut down of the TDAQ. This system also helps in the verification and diagnostics of the DCS, and acts as an early error detector of the occurring problems. It also stores the topology of the system and provides the information about the sources of errors, diagnostic messages and monitoring histograms from the TDAQ.

9.1.4 The detector control system

The DCS communicates with the LHC infrastructure, the CERN services and supervises the experimental setup, the detector infrastructure of ATLAS. It is also responsible for monitoring and controlling operations that have safety requirements, and executes the low level commands of controlling commands in their proper sequence. It also gathers event-by-event data, that can be used to evaluate the conditions of the detectors and is necessary for offline analysis. Unlike the DAQ, DCS operates outside data taking periods too.

9.2 Data types and data flow

The event data together with data about detector control, configuration, event and detector conditions, online statistics and monitoring are stored together, and is produced with a rate of 10 petabytes per year. To face this demand, this data is processed in a computing grid distributed around the world, the LHC Computing Grid. The data is stored in four layers on the grid with different complexity. The four layers

are:

- Tier-0 site keeps a permanent copy of the raw data. Currently it is at CERN, but is going to be expanded with a Hungary facility in 2013.

- Tier-1 sites do the full event reconstruction and data reduction, and they also serve as analysis facilities.

- Tier-2 sites belong to larger institutions, and have a copy of a data part of certain interest and they are responsible to generate the Monte Carlo simulations. They also accept analysis jobs from other institutions.

- Tier-3 sites are formally similar to Tier-2 sites, but they only serve their own institution.

The offline software of ATLAS is Athena, and is based on the Gaudi framework, that is shared with the LHCb experiment. Gaudi serves as an abstract layer that interprets detector architecture schemes and handles Monte Carlo simulations by Geant4 [66], whereas Athena handles the event generation, interface to simulation, reconstruction, analysis and visualisation.

Athena handles different abstractions of data, each are derived from the combination of raw detector output and calibration information. These abstractions are

- **RAW data** contains every event that passes the EF in a byte-stream, unprocessed format. An event is about $1.6\,\text{MB}$, and they are expected to come with a $200\,\text{Hz}$ frequency, accompanied with low frequency trigger calibration data.

- **ESD** or Event Summary Data contains a noise filtered event information, as an output of the reconstruction process. Its size is intended to be smaller than of the RAW events, and it is already in an object-oriented format. Its main purpose is to be used for calibration.

- **AOD** or Analysis Object Data is a reduced version of the ESD, and general enough for physics analyses. It is an object oriented ROOT file, and contains several calibration schemes and necessary meta-data. The size of an event is approximately one tenth of the RAW data, and AODs are subject to be duplicated at various sites. ATLAS recommends that any physics analysis should be reproducible by running on AODs, but the popular use cases dictated the standardisation of DPDs.

- **DPD** or Derived Physics Data are flat ROOT files, n-tuples. They come with a set of tools that help to personalise the content for individual groups or for certain physics goals. Refining iteration of analyses are typically faster on DPDs than on AODs; they occupy less space and it is easier to use them on local, non grid-enabled clusters. DPDs come in three level of abstractions, from the regular DPD to D3PD.

- **TAG** These files contain event-level meta-data, some basic information that can be used for event preselection, intending to speed up queries and event selection on AODs.

- **SIM** files are simulation files that contain the generator-level hard process information or any higher levels of the Geant simulation.

9.3 Local analysis codes

As AODs are rather large and new calibration schemes may appear rather often, their regular download is inefficient. A more convenient way is creating DPDs from the AODs on the computing grid by the process of *dumping*. The local analysis code used in this thesis is based on the SFrame framework [67], which offers several lightweight tools and eases the compilation of ROOT related codes. Uncompiled ROOT macros are generally considered to be prone to errors and recommended only for very small pilot analyses.

The infrastructure of SFrame was augmented for the local needs with the SFM code, offering features like multi-jet estimation through the fake lepton methods, simplified event selection and automatic control of histogram creation.

Common object selection

This chapter describes how the physical objects are identified within an event, following the recommendations of the various expert groups within ATLAS. Recognising the patterns of electrons, muons, jets or compound objects like missing transverse energy might seem straightforward, simply separating a pattern correlating with the objects from the detector noise. Most of the difficulties arise by defining patterns for the various objects that do not interfere each other; patterns that are independent of the abundances of the particles or at least well known for certain event topologies. In order to introduce the minimum possible systematic errors, all the object definitions were justified on well known physical processes by looking at control plots. The chapter follows the structure of the ATLAS note directed to define the object and event selection recommendations for 2012 top-quark analyses [68].

10.1 Muons

The recommendation of the Muon Combined Performance (MCP) group [69][70] are the following:

- `muid` algorithm for muon identification. This requires the reconstruction of track standalone track segments which are extrapolated to the vertex and refitted with the vertex constraint [71].

- Muons are required to be `combined`. The algorithm first a matches the `muid` track with a track from the inner detector and makes a global refit.

- The muons have to be within the acceptance range of the detector, $|\eta| < 2.5$.

- Muons are required to have $p_T > 25\,\text{GeV}$, in order to be on the plateau of the single muon trigger efficiency.

- Muons are required to pass the `MCP ID` track quality cuts.

- A b-layer hit for the muon track if the extrapolated track passes this region.

- The muon track should have at least one pixel hit or cross a dead pixel.

- The number of SCT hits of the muon track plus the number of cross dead SCT sensors must be greater or equal to 5.

- The number of pixel and SCT holes must be smaller than 3.

- With the n number of TRT hits plus TRT outliers, for $|\eta|$ from 0.1 to 1.9 n must be greater than 5 and the fraction of outliers smaller than 90%. For $|\eta| < 0.1$ and $|\eta| > 1.9$ require the outlier fraction to be smaller than 90% only if $n > 5$.

  Rejection of cosmic muons:

- The z projection closest approach of the track to the primary vertex must be smaller than 2 mm.

  Isolation from other objects:

- The requirement for the sum of the transverse energy around the muon in $0.1 < \Delta R < 0.2$ has to be $E_{\mathrm{T}}^{0.2} < 4$ GeV.

- The sum of transverse momentum of tracks around the moun within $0.2 < \Delta R < 0.3$ has to be $p_{\mathrm{T}}^{0.3} > 2.5$ GeV.

- The angular distance from a jet with $p_{\mathrm{T}} > 25$ GeV and $|\mathrm{JVF}| > 0.75$ has to be $\Delta R > 0.4$.

The jet vertex fraction *JVF* is a variable that determines if a jet is coming from the primary vertex or not. With this, the last three isolation requirements separate prompt muons from heavy flavour decays and typically from multi-jet backgrounds, which are expected to increase with pile-up, following the increase of the luminosity. The misidentification rate is estimated with the so-called matrix method, a fractional fitting method based on four or more bins. One of the bins should contain a relatively clear signal, dominated by $Z \to \mu\mu$ and selected by an invariant mass window around the Z mass. There is another bin with a high content of multi-jet events, the most dominating background, and the rest of the bins are overlapping regions of signal and background. The MCP group also provides efficiency scale factors, an estimate for the ratio of moun identification efficiency in MC and data, typically determined with a tag-and-probe method. The same scale factor is estimated for the muon trigger efficiencies for different run periods.

10.2 Electrons

The core of electron identification is associating clusters in the EM calorimeter to tracks in the Inner Detector. The complete list of requirements for the electron candidates are the following [72]:

- Pre-cluster seeds are reconstructed with the `SlidingWindowClustering` algorithm. This requires *towers*[1] to be built from calorimeter hits from longitudinal layers. The algorithm then checks all possible rectangular regions with a predefined $\Delta\eta$ and $\Delta\phi$ sizes, and keeps those clusters which pass a certain $E_T = E/\cosh(\eta)$ threshold. In case of overlap, the cluster with higher E_T is kept.

- EM cluster search is performed for different cluster sizes and differently calibrated for electron and photon candidates. The size in number of calorimeter cells in $\Delta\eta \times \Delta\phi$ are 3×7 and 5×5 in the barrel and in the end-caps respectively.

- The EM clusters have to be within the calorimeters acceptance range $|\eta| < 2.47$, excluding the transition region $1.37 < |\eta| < 1.52$.

---

[1] The two main clustering algorithms are the tower grid and the topological clusters. Their difference is detailed at the description of the jet reconstruction.

- Tracks used for association with clusters must pass good quality cuts: at least two hits in the pixel detector from which one should be in the b-layer and at least seven hits in pixel and SCT together (so called precision hits).

- The fraction of high threshold hits in TRT must be higher than 0.08 in order to reject pions.

- Transverse impact parameter of a track must be $|d_0| < 0.1$ cm

- Track $p_T > 2$ GeV

- The cluster and the track must be within $|\Delta\eta| < 0.2$ and $|\Delta\phi| < 0.1$.

- Electrons with low transverse energy, calculated from the cluster energy and track η with $E_T = E_{EM}/\cosh(\eta_{track}) < 25$ GeV are excluded

- Isolation cuts are made on multivariate likelihoods, based on sum of calorimeter momenta and energy in certain cones around the track to differentiate electrons from photons and jets.

Forward electrons, with $2.5 < |\eta| < 4.9$ are not used in this analysis, as the reconstruction of jets in this region are less certain, making jet-electron overlap removal impossible. Similarly to the muonic case, energy scale corrections and efficiency scales are determined with using the tag-and-probe method on $Z \rightarrow ee$ and $J/\Psi \rightarrow ee$ event candidates or E/P studies, because these latter ones are sensitive to the mass of the charged particle. The electron isolation was optimised on $W \rightarrow e\nu$ samples. The energy scale corrections are a function of $\phi_{cluster}$, $\eta_{cluster}$ and E_T, and are within $\pm 1.5\%$.

10.3 Jets

Jets are formed using the anti-k_t algorithm with a clustering parameter that corresponds to a radius of $R = 0.4$. The algorithm is seeded by *topological clusters* at the electromagnetic scale (EM) in the range of $|\eta| < 4.5$ and $p_T > 20$ GeV. The difference from the towering algorithm is, that the tower grid sums up the energies in the different calorimeter layers, resulting 2 dimensional (η, ϕ) grid, while the topological clusters are also taking into account the depth of the different calorimeter layers, hence creating 3-dimensional energy clusters [73]. The contributions of the two types of pile-up, in-time and out-of-time, can be subtracted using the knowledge of the number of primary vertices in the event and the average number of collisions per bunch crossing in a given luminosity block, respectively. This is typically done in bins of pseudorapidity. Due to physical and detector effects, jets must be *cleaned*. The important variables used to determine the quality of jets are the following:

- energy fraction in the electromagnetic calorimeter,

- maximum energy fraction in one calorimeter layer,

- energy fraction in the defective HEC layer,

- the fraction of good quality LAr cells.

Three categories of jets can be defined this way:

- Bad: jets that need to be removed, as they are likely to originate from background events or detector effects.

- Ugly: marks jets in problematic calorimeter regions, where energy is not well measured.

- Good: jets recommended for physics analysis.

The jets are calibrated to the hadronic scale, using p_T and η dependent factors derived from inclusive dijet Monte Carlo. The uncertainty of the energy scale for this method is much more homogeneous and smaller than for the data-driven Local Cluster Weighting (LCW) method at the moment. The uncertainty of the scale of the used EM+JES algorithm was derived from the early data periods from multi-jet events, taking flavour tagging information into account. The calorimeter jet reconstruction efficiency can be measured via the tag-and probe method, by reconstructing jets from charged tracks in the inner detector and matching them to calorimeter jets.

10.4 Missing transverse momentum

Missing transverse energy or missing transverse momentum can be reconstructed in two fundamentally different ways. The first is simply summing up the calibrated calorimeter energies vectorially, and taking the absolute value of the transverse component. This definition was used only in the early periods of LHC data taking. The present recommendation of the Jet/ETmiss Working Group is more sophisticated and uses more reconstructed information in the `MET_RefFinal` algorithm. This algorithm uses reconstructed and calibrated objects, based on the objects optimised for a certain physics channel. These objects are the electrons, photons, taus, muons, jets and soft jets (jets not passing the p_T threshold); all calibrated with their own scheme, but lacking out-of-cluster corrections. These latter corrections may appear for physics objects, one reason is that the clustering algorithms have threshold cuts and low energy hits might not be used. A second possibility is that charged particles in the showers may travel far from the centre of a cluster because of the strong magnetic field. The correction for the out-of-cluster effects should not be applied at this stage, since the missing transverse momentum calculation must be based on all of the calorimeter cells and otherwise the out-of-cluster energies should be subtracted again. Those calorimeter cells that were not associated with an object are calibrated to the EM scale and summed up as an additional term, E_T^{CellOut}. Considering all these contributions the $\not{E}_T$ can be expressed as:

$$\not{E}_{x,y} = E_{x,y}^{\text{RefElec}} + E_{x,y}^{\text{RefJet}} + E_{x,y}^{\text{RefSoftJet}} + E_{x,y}^{\text{RefMuon}} + E_{x,y}^{\text{CellOut}},$$

$$\not{E}_T = \sqrt{\not{E}_x^2 + \not{E}_y^2}.$$

The electrons in this collection have to fulfill the `Tight`++ criteria with transverse momentum $p_T > 10\,\text{GeV}$. The jets above $p_T > 25\,\text{GeV}$ are *refined jets*, and are calibrated to the EM+JES scale, while those with p_T between $7\,\text{GeV}$ to $25\,\text{GeV}$ are considered as *soft jets* and calibrated only to the EM scale. Muons are used in the full detector acceptance range, up to $|\eta| < 2.7$ and might be non-isolated muons. For the top-quark analyses the contributions of photon and tau are not included.

10.5 B-tagging

One of the most important types of flavour tagging is *b*-jet tagging, and ATLAS offers several algorithms to do this. The principles of *b*-identification can be the relatively long decay time and decay length of *B*-mesons, the large mass of such mesons or their large branching ratio into leptons. The three most mature algorithms are JetFitter, IP3D and SV1. The first one exploits the weak *b* and *c* decays, identifying

secondary decay vertices along the jet axis. The latter two methods look for secondary vertices around the primary vertex, but on different projections. This three methods are combined with a multivariate analysis, a TMVA-based neural network, to provide a superior discriminator: MV1.

Most of the systematic uncertainties for *b*-tagging can be traced back to the uncertainty of the *b*-jet contribution in the training sample, which are typically di-jet QCD events. Since its cross section has a large, approximately 100% uncertainty in the theoretical calculations, the experimental observations have an unknown factor of light-jet background. However, there are additional theoretical uncertainties in the proper description of *b*-hadron decays. Hence a certain *b*-tagging algorithm is only validated for a few discriminating values, where a *b*-jet selection efficiency with a light-jet rejection factor is provided.

10.6 Fake lepton estimations

Most of the objects faking the lepton identification algorithms originate from QCD multi-jet events. These have a large uncertainty in cross section and lack Monte Carlo simulations, therefore it was important to develop data driven methods to estimate the number of fake leptons in physically interesting channels. A more-or-less channel independent method was developed based on slightly loosening the lepton-identification criteria, and defining a new quality class which passes the lose cuts but fails the tight ones. The ratio of these objects can be measured in multi-jet dominated channels and then extrapolated to the phase space of interest, giving a tool to estimate the number of fake leptons from the number of loose-but-not-tight leptons in that given channel. This procedure of efficiency estimation is a form of matrix method. The multi-jet dominated control channel is selected by inverting some of the QCD suppression cuts. One of such cuts is the transverse *W*-mass $M_\mathrm{T}(W)$, and another one is the so-called inverted triangular cut:

$$M_\mathrm{T}(W) < 20\,\mathrm{GeV}\,,$$

$$\not{E}_\mathrm{T} + M_\mathrm{T}(W) < 20\,\mathrm{GeV}\,.$$

There are several methods based on these ideas, and the one that is used for this analysis is the so called matrix method B for muons [74].

CHAPTER **11**

Analysis of ATLAS data

The *Wt*-channel of the single top-quark productions has not yet been confirmed [75], and presents a challenge for multivariate techniques. In the following, the previously described density estimator and the polynomial regression method was applied to Monte Carlo simulations of the Standard Model processes, to search for phase spaces where the single top quark *Wt*-channel is dominant in order to extract this elusive signal. As fig. 11.1 shows, the ATLAS detector recorded $5.21\,\text{fb}^{-1}$ data in 2011. From this data $4.71\,\text{fb}^{-1}$ was available for this analysis, using periods B to H4 from the muon stream.
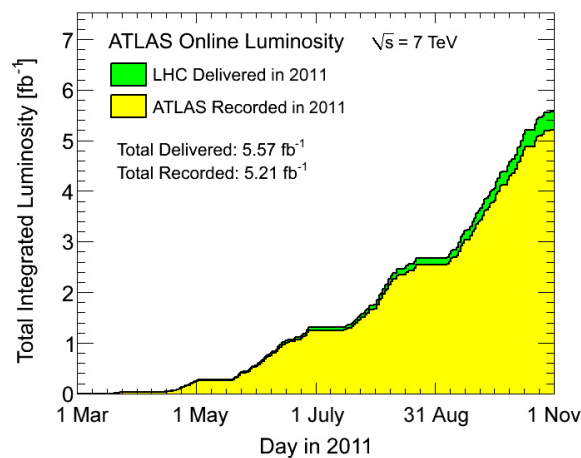
Figure 11.1: The collected integrated luminosity of ATLAS until the end of the year 2011

11.1 Preselection

The analysis strategy consist of two steps. First a preselection of events with loose cuts and second the multivariate analysis on the selected variables. The preselection follows the recommendations of the single top group, by selecting events with objects defined in chapter 10 passing the following thresholds:

- $p_\text{T} > 18\,\text{GeV}$ muon trigger;

- exactly three central jets with $p_T > 25\,\text{GeV}$ and $|\eta| < 2.5$, where one is identified as a *b*-quark jet using the MV1 tagger at 70% tagging efficiency;

- the jet vertex fraction must be $|\text{JVF}| > 0.75$;

- exactly one muon with $|\eta| < 2.47$ and $p_T > 25\,\text{GeV}$;

- $\not{E}_T > 25\,\text{GeV}$;

- $M_T(W) + \not{E}_T > 60\,\text{GeV}$ to suppress QCD events;

- the first vertex must be the primary vertex, which must contain at least 4 charged tracks;

- no jets tagged as bad with EM scale $p_T > 20\,\text{GeV}$.

The various Standard Model processes that contain prompt leptons, neutrinos and hard partons may appear as background to this single top quark channel, and were simulated by ATLAS groups. The list can be seen seen in table 11.1, along with the number of events remaining in the sample after the preselection, its corresponding integrated luminosity, the cross-section and the preselection efficiency.

11.2 Weights

Several of the processes appear more than once, as their production with multiple associated hard partons is not negligible and at least one event passed the selections. The NLO simulations are inherently weighted, with ±1 values, but all samples contain several weights to shape it to look like data. These all originate from accelerator- and detector-related variables, and can be thought of as calibrations. The global weight is a combination of the following:

- shaping the distribution of the number of simulated primary vertices to that of the data;

- scaling trigger efficiencies;

- scaling muon, electron and jet reconstruction efficiencies;

- scaling with the represented integral luminosity of the sample;

As fig. 11.2 shows that the distribution of weights in the Monte Carlo samples cover more than three orders of magnitude. Such a large weight fluctuation present a challenge to many multi-variate methods, since this may cause large uncertainty when investigating a small phase space. A way to decrease this effect is grouping together the samples with similar weights, train a classifier for the signal agains each background sample and combine these into a single classifier. The clustering of the different processes is listed in table 11.1. This technique was used for the density estimator only, as the polynomial regression method was proved to be robust against the effect.

11.3 The variables used in the likelihood analysis

To demonstrate the capabilities of the density estimator described in section 3.3 only kinematic variables were used. Although the composition of the variables is always allowed, it is only a transformation of the underlying probability density. Additional variables may improve the separation of signal and background, when they contain information that might help to improve the reconstruction or calibration

N_{sample}	$\mathcal{L}_{\text{int}}/[\text{pb}]$	Process	+partons	$\sigma[\text{pb}]$	ϵ_{presel}	Legend name
20	417	$W \to \mu\nu$	+p0	8 304	7.00×10^{-6}	
54	1 597		+p1	1 565	2.77×10^{-5}	
322	4 776	Wbb	+p1	43	0.001 70	
2	5 069	$t\bar{t} \to$ jets		76	-6.77×10^{-7}	Wc
813	6 642		+p1	312	0.000 392	
2459	6 734	Wc	+p2	77	0.004 69	
849	6 637		+p3	17	0.007 53	
168	7 049		+p4	4	0.005 94	
19	7 914		+p0	836	2.95×10^{-6}	
20	7 956		+p1	168	1.82×10^{-5}	
88	8 033	$Z \to \mu\mu$	+p2	50	0.000 260	
230	7 863		+p3	14	0.002 42	
87	8 726		+p4	3	0.003 32	
26	10 384		+p5	1	0.002 94	
3	12 703		+p0	836	3.26×10^{-7}	
37	19 787		+p1	169	1.07×10^{-5}	
187	19 918	$Z \to \tau\tau$	+p2	50	0.000 190	
286	34 849		+p3	15	0.000 601	
94	41 571		+p4	3	0.000 678	
888	8 290		+p2	453	0.000 360	
2363	8 246	$W \to \mu\nu$	+p3	122	0.003 87	W/Z+jets
908	8 251		+p4	31	0.005 95	
196	8 430		+p5	8	0.004 72	
51	8 266		+p2	454	2.05×10^{-5}	
134	8 292	$W \to \tau\nu$	+p3	122	0.000 208	
66	8 125		+p4	31	0.000 423	
44	8 366		+p0	57	0.000 101	
1332	8 385	Wbb	+p2	21	0.008 19	
657	8 798		+p3	8	0.0107	
137	8 986		+p0	153	0.000 107	
580	8 677	Wcc	+p1	126	0.000 564	
939	8 399		+p2	63	0.001 92	
563	8 354		+p3	20	0.003 98	
326689	127 936	$t\bar{t} \to l\nu +$ jets		91	0.0217	TTbar
3831	146 254	WW		17	0.001 56	
1844	180 389	WZ		6	0.001 89	Diboson
209	197 784	ZZ		1	0.000 803	
267	25 310	t-chan $\tau\nu$		7	0.000 965	SgTop t-chan
5618	25 366	t-chan $\mu\nu$		7	0.0200	
419	506 981	s-chan $\tau\nu$		1	0.001 30	SgTop s-chan
7483	507 147	s-chan $\mu\nu$		1	0.0246	
12427	50 637	Wt-chan		16	0.0131	SgTop Wt

Table 11.1: List of the used samples. The table shows the unweighted number of events passing the cuts, the integrated luminosity of the sample, cross-section corresponding to the process and the efficiency of the preselection cuts. The different samples have to be weighted with the inverse of their luminosity to be properly represented in the training. To avoid large weight fluctuations the samples with similar weights were grouped together. Only two negative weighted events from hadronic $t\bar{t}$ decay passed the preselection cuts, resulting in negative efficiency. The multijet sample is not listed, as it was obtained in data-driven way.
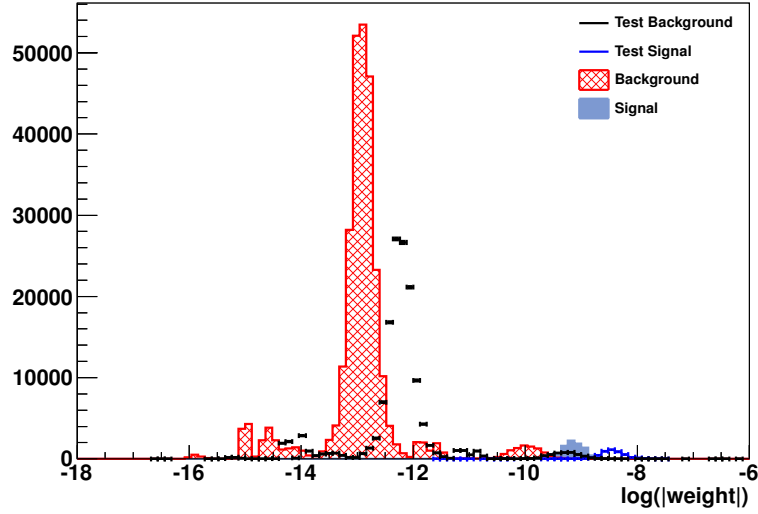
Figure 11.2: The distribution of weights used in the Monte Carlo simulations on logarithmic scale. As the ratio of weights can be thousandfold, it is necessary to form groups of samples with smaller weight fluctuations for the density estimator. The polynomial regression uses test sample that is half the size of the training sample. The test samples were also normalised to unity, which is reflected by the upward shift of their weight distributions with a decrease in the amplitudes corresponding to the smaller sample sizes.

of the kinematical variables, such as flavour tagging. Nevertheless, their systematical uncertainty is believed to be high at the moment, and only *b*-tagging is used. The charge of the lepton was also considered as an input variable, but since it is a discrete value, it eventually halves the useful statistics for the training so it is left out.

The remaining variables are then defined by the selected channel topology, the four vectors of the three jets, the lepton and the known components of the missing momentum. Altogether these are eighteen variables:

- p_T, E, ϕ and η of the three jets;

- p_T, E, ϕ and η of the muon;

- $\not{E}_T$ and $\phi_{\not{E}_T}$.

In principle it is possible to neglect some of these. This is because the detector respects the cylindrical symmetries of the possible scattering distributions, the origin of the rotational parameter ϕ can be chosen for every event, it could be set for example to the ϕ_{hj} of the hardest jet. One possible parametrisation is then using the p_T, E, $\phi - \phi_{hj}$ and η of the jets and the lepton and finally $\not{E}_T$ and $\phi_{\not{E}_T} - \phi_{hj}$. Nevertheless, due to detector effects neither the jet nor the muon reconstruction efficiency is uniform in ϕ. These defects would propagate into the $\Delta\phi$ distributions causing unnecessary smearing, therefore all the ϕ variables listed above were used.

11.4 Preprocessing

The used density estimator has two caveats, which must be noted:

- The applied regularisation scheme introduces a resolution with Euclidean measure, hence it is worthwhile to use variables with approximately the same variance.

- The density estimation is uncertain around edges, and so around cut values. However, if the variables are transformed, it is better to be a continuous transformation. This way one can partially avoid to introduce an artificial edge in interesting phase space, where the signal can be found.

The selection of the $\phi_{\rm hj}$ as an origin for the ϕ coordinates would eliminate the uncertainty of the density estimation at $\phi = 0$ and $\phi = 2 \cdot \pi$. However, this reparameterization was rejected, because the used physics objects have different resolutions and efficiencies in function of ϕ. The η distribution has cut-off values where the detector acceptance ends, $|\eta| < 2.5$, but the $p_{\rm T}$ and $E_{\rm T}$ variables are approximately exponential distributions, and are much wider. To compensate this, the natural logarithm of these values, measured in GeV, was used.

Figure 11.3 shows the quality of the muon-related variables. The angular distributions in fig. 11.3c and 11.3d show good agreement between the geometrical variables of data and MC. The resolution of the muon reconstruction is good enough to reconstruct the positions of dead materials. The dips in the ϕ distribution are the results of the metal legs of ATLAS, the dips in $|\eta| \approx 1$ are caused by the transition region between the solenoidal and the toroidal magnetic fields, while at $|\eta| \approx 0$ is due to the gap between the muon spectrometers. Figure 11.3a and 11.3b demonstrates that the $p_{\rm T}$ and E distribution of the muons are well described and accurately calibrated.

The control plots for the hardest light-parton jet are shown in fig. 11.4. Contrary to the $p_{\rm T}$ distribution of the softer jet on fig 11.5a, the $p_{\rm T}$ of harder light-parton jet has a smoother turn-on on fig 11.4a. It is also visible, that the simulation systematically underestimates data in the low $p_{\rm T}$ region and overestimates it in the high $p_{\rm T}$ region. This effect is more pronounced for the $p_{\rm T}$ spectrum of the b-jet on fig. 11.6a, suggesting that this discrepancy it originates from a badly estimated b-tagging and mistagging rate, mis-estimating the contribution of the different processes depending on their b-jet content.

Figure 11.5 shows the control plots for the second hardest light-parton jet. The sharp rise of the $p_{\rm T}$ distribution in fig. 11.5a suggests that the real peak is below the 25 GeV cut value. Since this distribution has a strong slope around the cut value, it is expected that the analysis depends on the jet energy scale and the chosen $p_{\rm T}$ threshold. Although the sum of the backgrounds follow a featureless exponential $p_{\rm T}$ fall or a linear fall in $\log(p_{\rm T})$, the distribution of the individual background shapes slightly differ from each other. The multijet, the vector boson plus light jet and for the W-boson plus with associated charm production this distribution are quite similar, indicating that the soft light jet in these processes are indeed originating from a non-hard process. This jet can be either from a secondary interaction within the proton-proton collision, initial or finial state radiation or a randomly picked jet from a pile-up event. The $p_{\rm T}$ distribution of the $t\bar{t}$ is visibly different as it has a larger $p_{\rm T}$ on average than the other backgrounds, showing that these jets are originating from a high mass particle, a top quark. The softer light jet on average is softer than the b-tagged jet, hence in most of the events this is the softest jet. Overall, the $p_{\rm T}$ and the E distributions on fig. 11.5a and 11.5b describe the data more accurately than for the selected b-jets on fig 11.6. Despite the accuracy in the kinematic distributions, as fig. 11.5d shows, the η distribution of the softer light jet is worse than for the other jets, as it predicts more central jets than what is visible in the data.

The control plots for the b-tagged jet can be found on fig. 11.6. The $p_{\rm T}$ distribution of the b-tagged jet on fig. 11.6a shows the largest systematic discrepancy between data and simulation from all of the selected variables. This can mean that either the b-tagging efficiency together with the mis-tagging rates are off, the physics of b-jets is not well described or the true b-jets are mis-calibrated. In case it is a calibration problem, a $p_{\rm T}$ correction only on the true b-jets could move the $t\bar{t}$ peak, while leaving the

(a) p_T

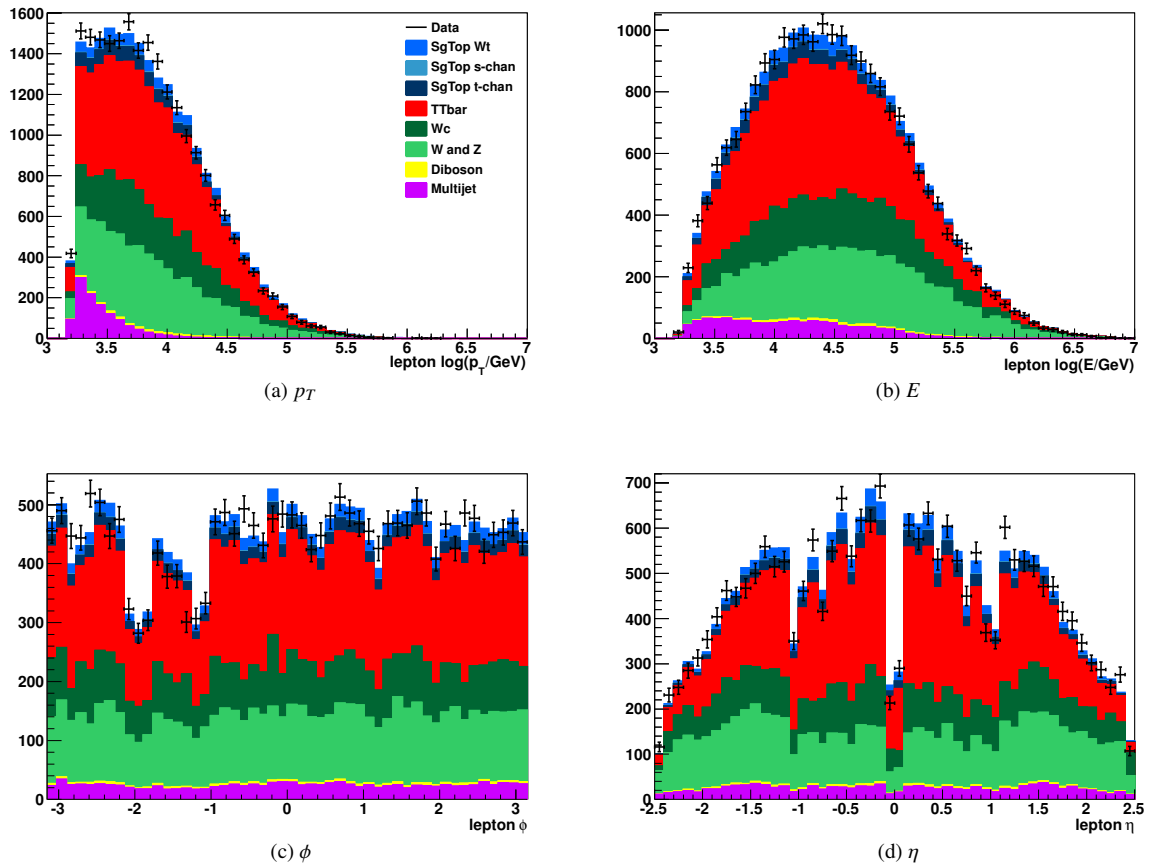(b) E

(c) ϕ

(d) η

Figure 11.3: Muon control plots. Data shown as points with error bars, simulation with filled areas as indicated by the legend. The simulations were normalised to 4.71 fb^{-1}, the luminosity of the data. The sub-constituents of the samples indicated in the legend are detailed in table 11.1. The statistical and the systematic uncertainty of the simulations are not shown.
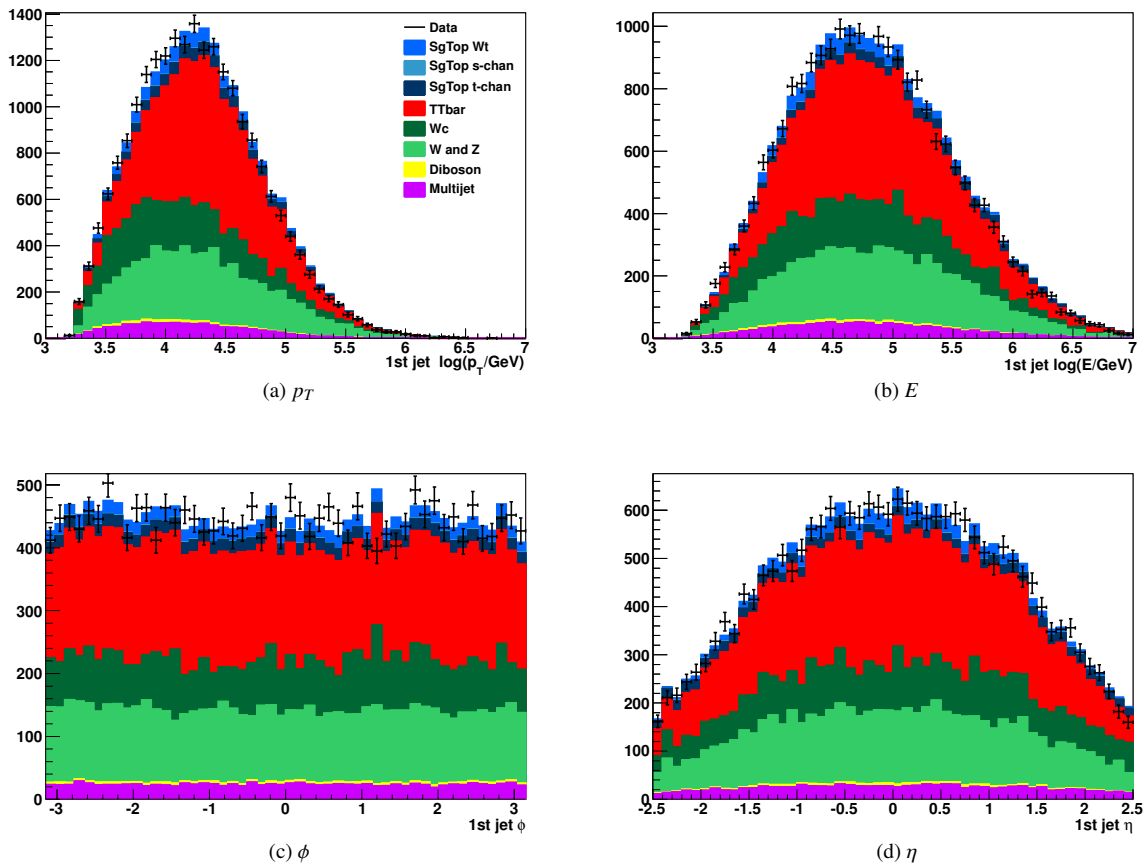
Figure 11.4: Hardest light jet control plots. The constituents of the samples indicated in the legend are listed in table 11.1.
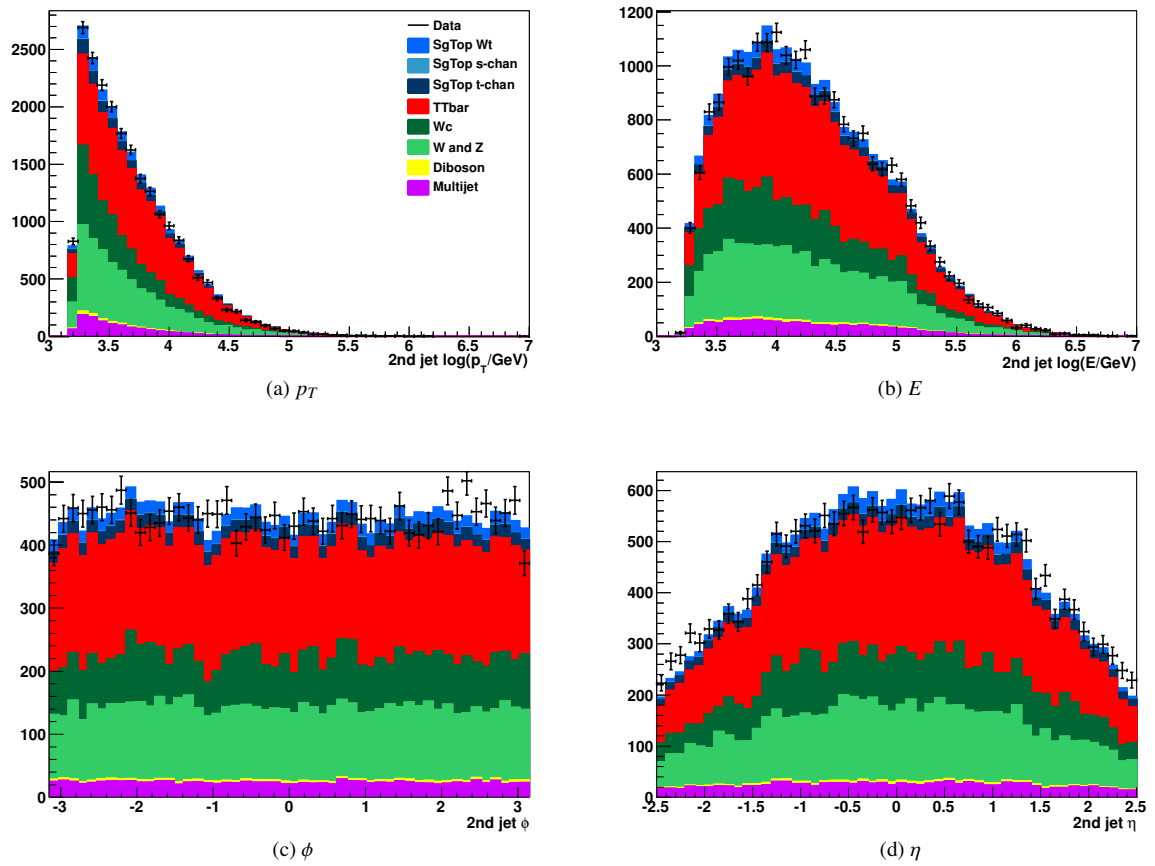
Figure 11.5: Second hardest light jet control plots. The constituents of the samples indicated in the legend are listed in table 11.1.

other processes with fake b-jets intact. On the other hand, if the difference between Monte Carlo and data comes from the mis-mesurement of b-tagging efficiency and mis-tagging rates, the nearly exponential distributions of the processes with fake b-jets should be rescaled with different factors than the $t\bar{t}$ with true b-jets. Apart from this discrepancy, the geometric description of the b-tagged jets is rather good, as it is visible on the ϕ and η distributions on fig. 11.6c and 11.6d.
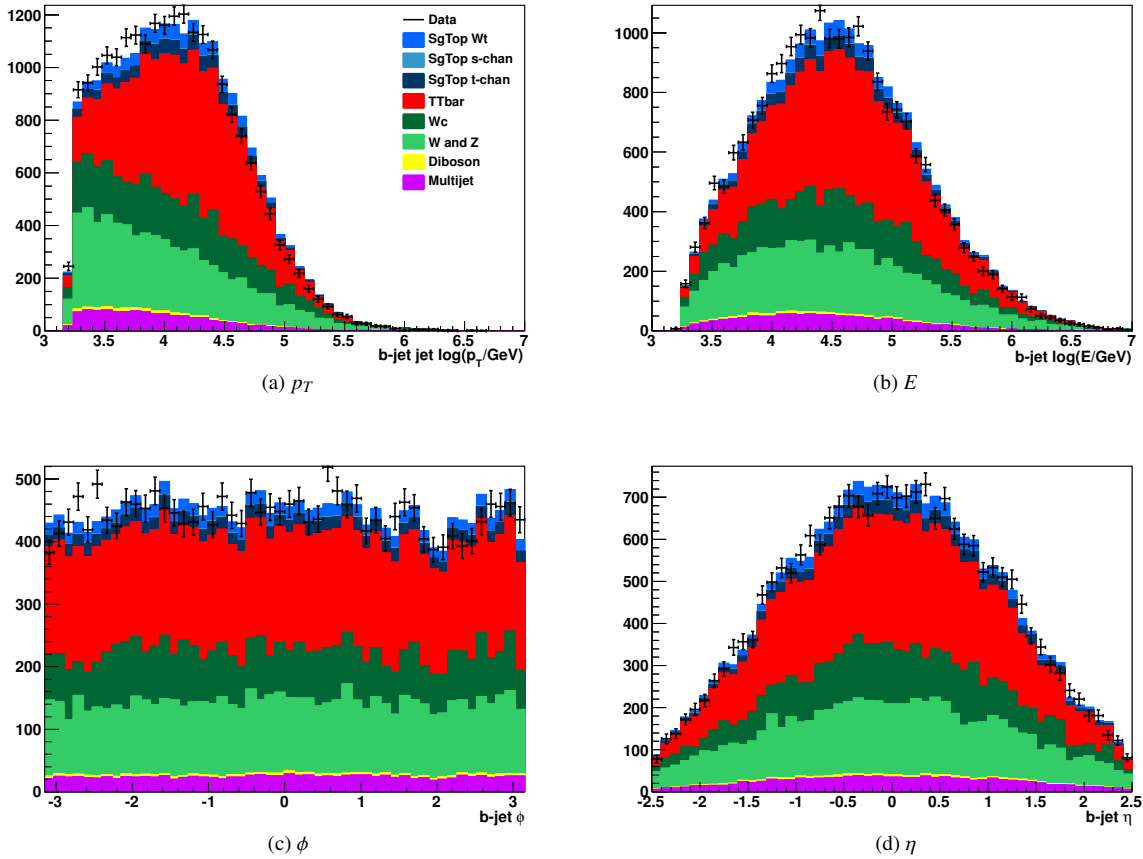
Figure 11.6: b-jet control plots. The constituents of the samples indicated in the legend are listed in table 11.1.

Despite being the most complex physics object used, the description of the missing transverse momentum is rather good as the shape estimation from simulation matches well with the shape of the data, as it is visible on fig. 11.7.

In summary, the ϕ and η distributions agree well with the data, proving that the detector geometry, the acceptance and efficiencies are well modelled. However, a discrepancy in the momentum related variables is clearly visible. The simulation shows a tendency to overestimate the number of events with high p_T and E, while underestimates it in the low region. It is especially strong for the p_T distribution of the b-jets, which suggests that it is not a simple calibration problem, but a mis-measurement of the b-jet tagging efficiency and light-jet rejection rate. Since the sample for the multijet background was obtained in a data-driven way, it can be assumed that it is calibrated along with the data. In order to investigate this in more detail, a classifier was trained to separate the multijet sample from the other Standard Model processes. Figure 11.8a shows that the multijet sample can be very effectively separated from the other SM model processes. Comparison with data is visible on fig. 11.8b, which shows surplus

(a) $\not{E}_T$

(b) ϕ
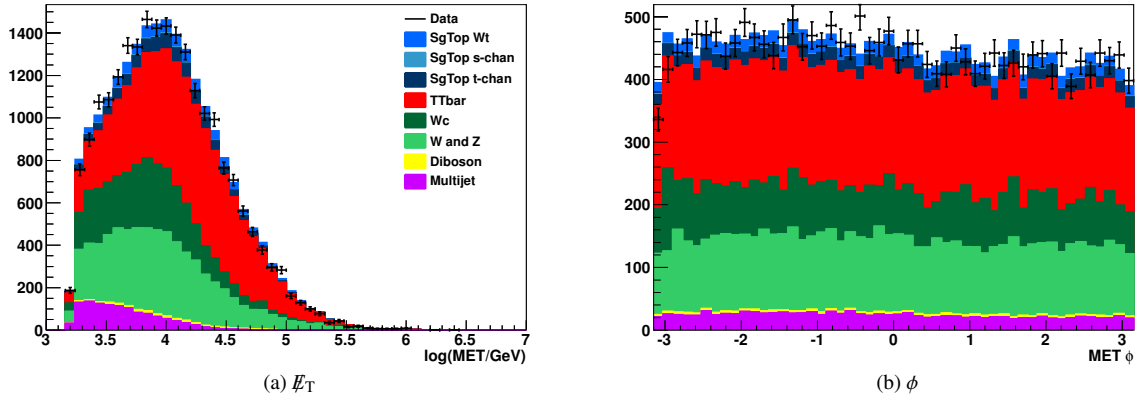
Figure 11.7: Missing transverse momentum control plots. The constituents of the samples indicated in the legend are listed in table 11.1.

at the multijet-dominated right side and deficit on the left side. This discrepancy indicates that the actual cross-section of the multijet background is somewhat larger than expected, and some other processes need to be scaled down in the meantime. This is consistent with the assumption that the *b*-tagging efficiency and the mis-tagging rate are not accurate.

(a) Training with multijet as a signal

(b) Data
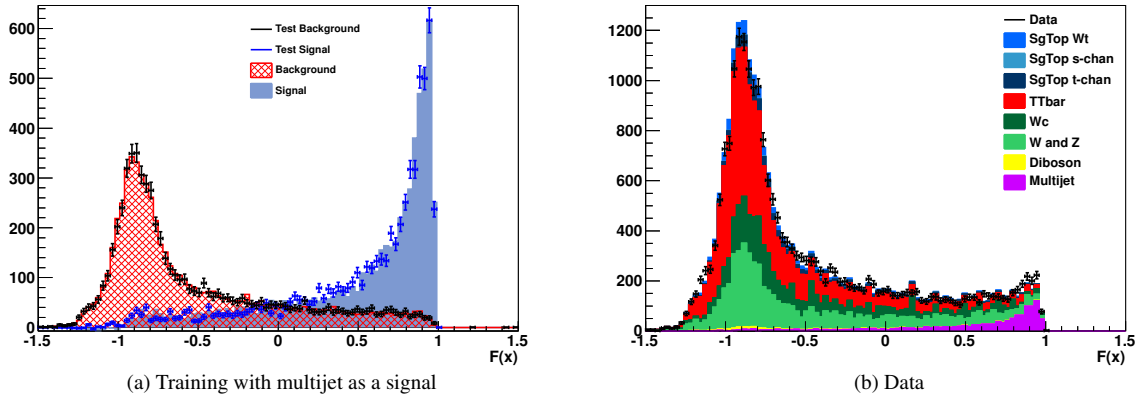
Figure 11.8: Multijet training and evaluation. The polynomial regression method was used to separate the data-driven multijet background from the other Standard Model processes with their expected cross-sections, including the unconfirmed single-top *Wt*-channel.

11.5 Training results and evaluation on data

Figure 11.9 shows the likelihood estimation on the simulation and the evaluation on data based on the density estimator. As the density estimator requires using a regularisation, there is no need for independent training and testing samples. Both signal and background were normalised to unity for the training, with the notion that the multijet samples were excluded from the training. The network is able to see a slight difference between the signal and the background, as it is shown on fig. 11.9a.

Indeed, as fig. 11.9c indicates, the network response correlates with the purity of the signal. Since it is off the ideal linear curve, there must be a bias present in the density estimations. Applying a cut on the network output and the signal efficiency, the background rejection and the signal purity can be calculated for the remaining sample. Their relation, as is shown on fig. 11.9d is actually independent of the parameterisation of the network output and reflects the separating power of the method. The better the separation is, the background rejection vs signal efficiency curve (red) approaches closer to the upper-right corner. The diagonal from the upper-left corner to the bottom right represents the worst case, when the separation is not better than a random choice. The Gini index of this training is 4 %, which is far from the maximal possible 50 %, also indicating that the separating power is poor, but not negligible. The blue curve on the same plot shows the relation between the signal purity after the cut on the classifier against the signal efficiency. At 100 % efficiency the purity is 50 %, representing the signal probability for a random choice. At lower signal efficiencies the purity slowly increases to its maximum around 65 %. As a comparison, the polynomial regression reaches this purity much earlier, around 60 % signal efficiency. Figure 11.9b contains the comparison of the simulations scaled to the expected cross-sections agains the data. This shows that although the data is well described in the projection of the network response, due to the small cross-section of the Wt-channel signal the detected slight shape difference can not be exploited for cross-section measurements with the available statistics. As it was indicated before, the cross-section of the various background processes are also uncertain. To measure them, the training should provide phase spaces where either the signal or the background is rather pure, but this is not the case.

Using the same variables the data was also fed to a regression method for comparison, its result is summarised in fig. 11.10 and 11.11. Its separation power can be seen on fig. 11.10a, overlaid with an evaluation of an independent test sample with half the size of the training sample. The similarity of the training and testing shapes indicates that no overtraining occurred. The separation power has greatly improved in comparison to the results of the density estimator on fig. 11.9a. As fig. 11.10c shows, the regression function $F(x)$ has a strong correlation with the purity of the signal, resulting in a good training quality. Figure 11.10d was produced on the testing samples in a similar manner to fig. 11.9d, with a Gini index of 17 %, quantifying the improved separation power. This number is calculated from the Lorenz curve showed in fig. 11.11 which plots the signal efficiency against the cut efficiency. The ratio of the area between the curve and the diagonal to the area below the this diagonal is called the Gini index. The larger this index is, the better is the separation between signal and background. The ideal classifier, which contours follows the signal and background density ratio maximises this index. As a comparison, the Gini index of the classifier based on dipole density estimator was only 4 %. Using variables other than the physically motivated four-vectors can result in simpler fit function and better fits. As an example, the polynomial regression can produce a Gini index of 18.5 % on variables that were selected by having strong linear correlation with the regression target.

Despite the good results of this method, the small expected cross of the single top Wt-channel is still practically undetectable with this technique with the current statistics as it is apparent on fig. 11.10b.

The separation between signal and background is much better than for the density estimator, as the shape of the signal and the background distributions are visibly different in this projection. With equal amounts of signal and background, the maximal signal purity is more than 70 %, but this drops down to nearly 5 % when applying the expected cross-sections. Nevertheless, the shape difference can still be exploited to define signal-rich and signal-depleted regions. Assuming that the relative cross-sections of the backgrounds are precise enough, cut and count methods, such as the matrix method can be applied to determine the posterior probability of the signal cross-section.
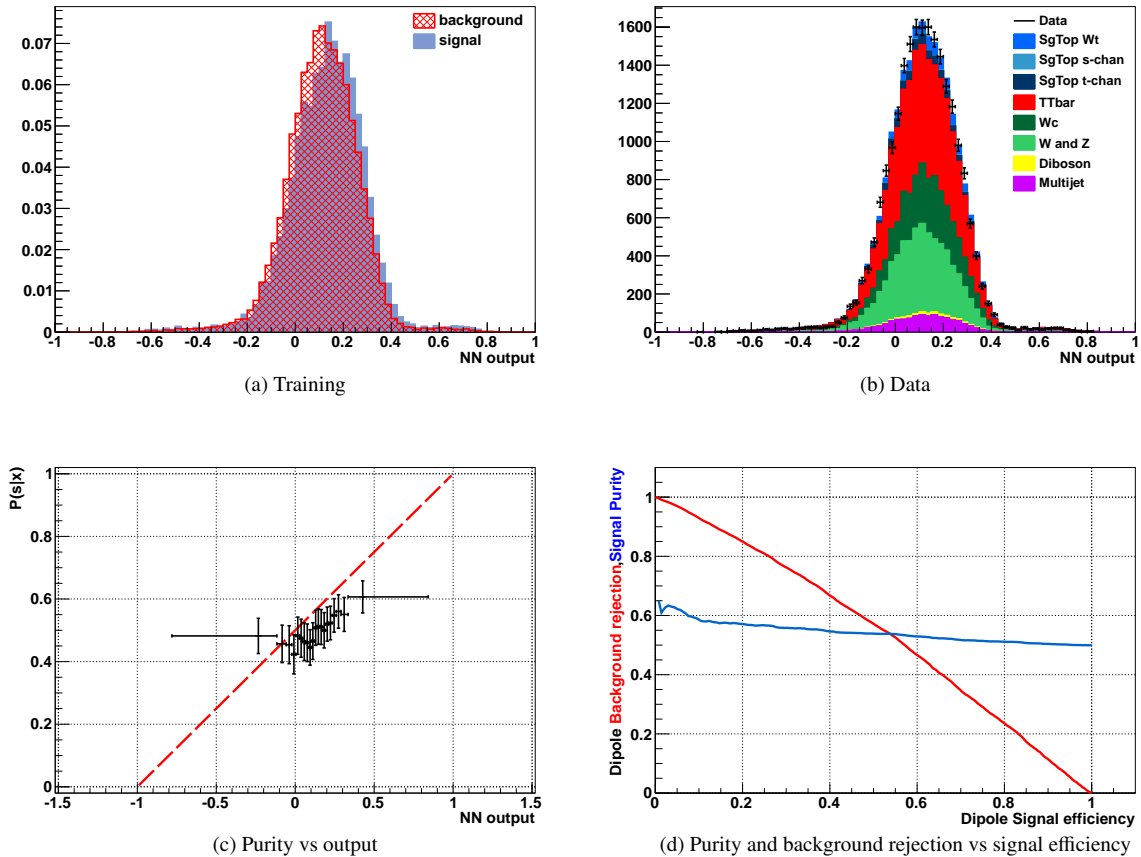
(a) Training

(b) Data

(c) Purity vs output

(d) Purity and background rejection vs signal efficiency

Figure 11.9: Traning and evaluation of the dipole network on data.

11.6 Conclusions

Since the input variables are only the known components of the four vectors of the physical objects, the different processes are embedded into various hyper-surfaces. They possibly have different dimensionality, defined mainly by the mass constraints. Although the density estimation is biased, as can be deduced from the non-linearity of fig. 11.9c, it is still able to see a small difference between the signal and the background processes. However, these hyper-surfaces are certainly curved, which may indicate that the Euclidean measure-driven regularisation and resolution may not be the optimal one. At a given phase space region the crucial variable for the optimal separation is the ratio of the signal to background densities, to be precise the contour surfaces of this ratio. The optimal signal and background density resolution should follow these contours, but these can not be calculated without the precise knowledge of the densities. To circumvent this circular dependency, one can use a regression technique which has access both to the signal and background samples in the same time. The presented regression method, based on the polynomial expansion of the ideal classification function indeed performs better than the density estimation based likelihood calculation.

Although the density estimator inherently applies a regularisation, or in other words a smoothing window, due to the large weight fluctuations it was necessary to apply further machinery for smoothing due to the large weight fluctuations present in the samples. This latter type of smoothing is technically

(a) Training

(b) Data

(c) Purity vs output

(d) Purity and background rejection vs signal efficiency
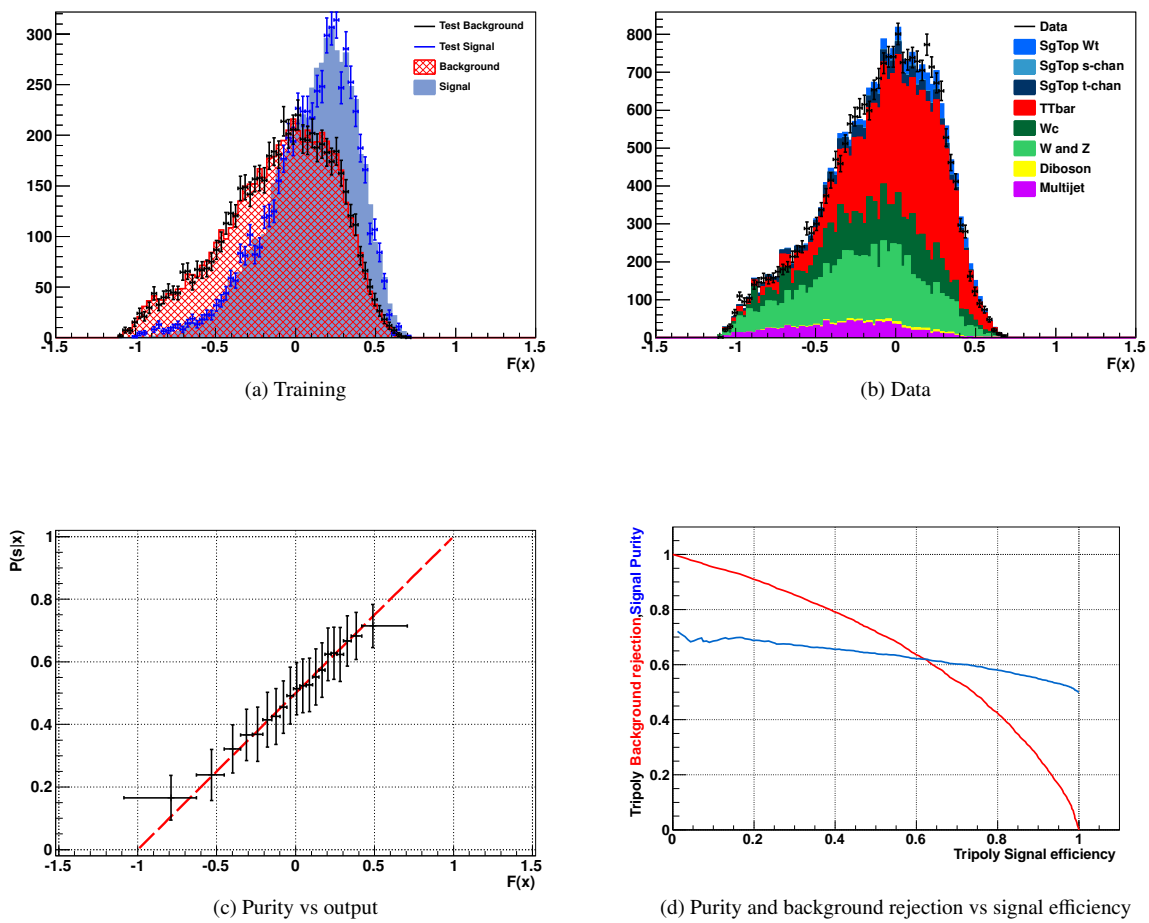
Figure 11.10: Training and evaluation of the polynomial regression method on data.

Gini index = 17.01%, max possible = 30.89%
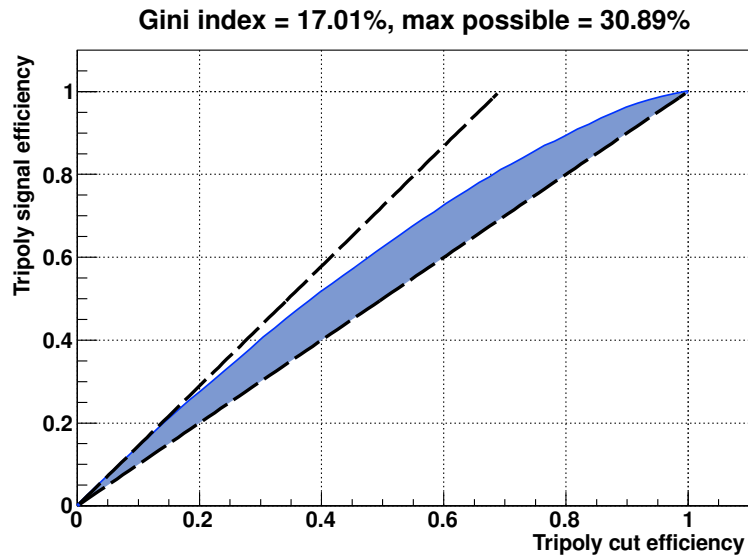
Figure 11.11: Lorenz curve (blue) of the polynomial regression. This curve uses similar informations as the ROC curve on fig. 11.10d, but instead the background rejection rate it plots the signal efficiency against the sample efficiency belonging to a certain cut value on the polynomial response. The diagonal line between $(0,0)$ and $(1,1)$ represents the worst-case classifier, which is equivalent to a random choice. The dashed line above the Lorenz curve is its estimated tangent at $(0,0)$, showing that a strong cut on the response value would produce a maximal Gini index of nearly 31 %.

equivalent to the regression of weights, in which regression function is used later in combination with the density estimation on the unweighted data points to produce a density estimate on the weighted sample. To do this efficiently without using unnecessarily large smoothing windows on the weights, the training sample had to be separated into subsamples with smaller weight fluctuations. Nor the clustered sampling and neither the explicit regression of weights was necessary when the classifier was calculated directly via regression, as it performed the same way with and without these.

In summary, the regression method is simpler, faster and easier to use than the density estimator and provides a better separation.

Summary

This thesis discussed various methods related to probability theory that can be used to analyse physical phenomena. It is well established that the physical laws must be a description of interacting quantised fields, and this was derived in Chapter 2. This chapter showed why it is important to look for correlations, either expected or unexpected ones, in the physical experiments. Three machine learning algorithms were described in the following chapters, two supervised and an unsupervised one.

The first supervised method is based on a novel density estimation technique that incorporates an unusual, vector-valued kernel derived from the Green's function of the Laplace operator. As Chapter 3 proves theoretically and practically, it is capable of estimating the probability density function from a data sample and can be used to estimate a likelihood of a class appearing on a known background.

The unsupervised algorithm in Section 4.3 approaches classification from the information theory viewpoint, and aims to classify data samples as belonging to certain hyper-surfaces. This criterium is the theoretical requirement to form parametric classes, consequently ideal to be used in finding physical objects.

In Chapter 11 the density estimator was applied to Monte Carlo simulations of a physics processes propagating through the ATLAS detector in comparison with recorded data. The used density estimator could see a slight difference between the simulated samples of the single top-quark Wt-channel and the expected background from the other processes of the Standard Model. The same chapter also showed that although the density estimator had inherent regularisation of the statistical noise, its preference on certain distance measures can be a disadvantage.

The simple regression method based on the polynomial approximation of the ideal classification function, being described in Chapter 5 can have a better classification power. It is mainly because prior to comparing the two distributions it is impossible to know which variables might turn out to be useful *locally* in a certain phase space region, and because it is hard to estimate densities reliably over orders of magnitude differences. A binary regression is not affected by this, as it has access to both the signal and the background sample and it is sensitive to the $s(x)/b(x)$ density ratio instead of the densities itself. This regression method is unique in many senses. It does not require computationally intensive parameter optimisations, only a solution of linear equations. This not only makes it fast, but also allows the propagation of the Gaussian uncertainties into the output of the regression function and to its χ^2, allowing the regularisation of the fit function with statistical confidence.

Bibliography

[1] P. Kövesárki, I. C. Brock, and A. E. N. Quiroz, "Green's function based unparameterised multi-dimensional kernel density and likelihood ratio estimator," *Journal of Physics: Conference Series* **368** no. 1, (2012) 012041, `http://arxiv.org/abs/1112.2093`. `http://stacks.iop.org/1742-6596/368/i=1/a=012041`.

[2] J. Arguin *et al.*, "Jet selection for top physics," Tech. Rep. ATL-PHYS-INT-2010-134, CERN, Geneva, Dec, 2010.

[3] P. Kövesárki, "Polynomial expansion of the binary classification function," *ArXiv e-prints* (Mar., 2012) , `arXiv:1203.5647 [stat.ML]`.

[4] E. Noether, "Invariant Variation Problems," *Gott.Nachr.* **1918** (1918) 235–257, `arXiv:physics/0503066 [physics]`.

[5] A. M. Gleason, "Measures on the closed subspaces of a Hilbert space," *J. Math. Mech.* **6** (1957) 885–893.

[6] J. V. Neumann, *Mathematical Foundations of Quantum Mechanics*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, 1996. `http://books.google.de/books?id=JLyCo3RO4qUC`.

[7] H. Everett, "Relative State Formulation of Quantum Mechanics," *Rev. Mod. Phys.* **29** (Jul, 1957) 454–462. `http://link.aps.org/doi/10.1103/RevModPhys.29.454`.

[8] S. Weinberg, *The Quantum Theory of Fields: Supersymmetry*. No. v. 3 in The Quantum Theory of Fields. Cambridge University Press, 2000. `http://books.google.de/books?id=OoR0lZdqxqIC`.

[9] Michael E. Peskin, Daniel V. Schroeder, *Introduction to quantum field theory*. Perseus Books Publishing, L.L.C., 1995. ISBN 0-201-50397-2.

[10] J. Goldstone, A. Salam, and S. Weinberg, "Broken Symmetries," *Phys.Rev.* **127** (1962) 965–970.

[11] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys.Rev.Lett.* **13** (1964) 508–509.

[12] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys.Rev.Lett.* **13** (1964) 321–323.

[13] **Particle Data Group** Collaboration, K. Nakamura *et al.*, "Review of particle physics," *J.Phys.G* **G37** (2010) 075021.

[14] S. R. Coleman and J. Mandula, "All possible symmetries of the S matrix," *Phys.Rev.* **159** (1967) 1251–1256.

[15] S. P. Martin, "A Supersymmetry primer," `arXiv:hep-ph/9709356 [hep-ph]`.

[16] A. Andai, "Mathematical basis of quantum mechanics," Master's thesis, ELTE Department of Physics, Budapest, 1998. Master thesis (in Hungarian).

[17] B. D. Ripley, *Pattern Recognition and Neural Networks*. Cambridge University Press, Jan, 1996. ISBN 978-0521-71770-0.

[18] R. Brun and F. Rademakers, "ROOT - An Object Oriented Data Analysis Framework," in *Proceedings AIHENP'96 Workshop, Lausanne*, vol. 389 (1997), pp. 81–86. Nucl. Inst. & Meth. in Phys. Res. A, Sep., 1996. See also http://root.cern.ch/.

[19] T. J. Hastie, *Principal curves and surfaces*. PhD thesis, Stanford University, 1984. Ph.D. Thesis.

[20] D. Huffman, "A method for the construction of minimum-redundancy codes," *Proceedings of the IRE* **40** no. 9, (Sept., 1952) 1098 –1101.

[21] G. Young and A. Householder, "A note on multidimensional psychophysical analysis," *Psychometrika* **6** (1941) 331–333. `http://dx.doi.org/10.1007/BF02288589`. 10.1007/BF02288589.

[22] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006. ISBN 0-387-31073-8.

[23] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London. Series* A **231** no. 694-706, (1933) 289–337, `http://rsta.royalsocietypublishing.org/content/231/694-706/289.full.pdf+html`. `http://rsta.royalsocietypublishing.org/content/231/694-706/289.short`.

[24] The European Mathematical Society, "Articles about the Hankel matrix and the Hamburger-Stieltjes moment problem.," *Encyclopedia of Mathematics* . `http://www.encyclopediaofmath.org/index.php?title=Hankel_matrix&oldid=23850`.

[25] G. Guennebaud, B. Jacob, *et al.*, "Eigen v3." Http://eigen.tuxfamily.org, 2010.

[26] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, third ed., 1999.

[27] M. Peskin and D. Schroeder, *An Introduction To Quantum Field Theory*. Advanced Book Program. Addison-Wesley Publishing Company, 1995. `http://books.google.de/books?id=i35LALN0GosC`.

[28] M. Seymour, "Introduction to monte carlo event generators," in *CTEQ-MCnet School 08*. Aug., 2008. `http://conference.ippp.dur.ac.uk/getFile.py/access?contribId=13&resId=0&materialId=slides&confId=156`.

[29] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e-Annihilation by Perturbation Theory in Quantum Chromodynamics.," *Sov.Phys.JETP* **46** (1977) 641–653.

[30] V. Gribov and L. Lipatov, "Deep inelastic e p scattering in perturbation theory," *Sov.J.Nucl.Phys.* **15** (1972) 438–450.

[31] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," *Nucl.Phys.* **B126** (1977) 298.

[32] A. Martin, W. Stirling, R. Thorne, and G. Watt, "Parton distributions for the LHC," *Eur.Phys.J.* **C63** (2009) 189–285, `arXiv:0901.0002 [hep-ph]`.

[33] M. Cacciari, "Heavy quarks," in *CTEQ-MCnet School 08*. Aug., 2008. `http://conference.ippp.dur.ac.uk/getFile.py/access?contribId=3&resId=0&materialId=slides&confId=156`.

[34] G. Salam, "Jets," in *CTEQ-MCnet School 08*. Aug., 2008. `http://conference.ippp.dur.ac.uk/materialDisplay.py?contribId=7&materialId=slides&confId=156`.

[35] G. Sterman and S. Weinberg, "Jets from Quantum Chromodynamics," *Phys. Rev. Lett.* **39** (Dec, 1977) 1436–1439. `http://link.aps.org/doi/10.1103/PhysRevLett.39.1436`.

[36] S. Bentvelsen and I. Meyer, "The Cambridge jet algorithm: features and applications," *The European Physical Journal C - Particles and Fields* **4** (1998) 623–629. `http://dx.doi.org/10.1007/s100520050232`. 10.1007/s100520050232.

[37] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm," *JHEP* **0804** (2008) 063, `arXiv:0802.1189 [hep-ph]`.

[38] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput.Phys.Commun.* **178** (2008) 852–867, `arXiv:0710.3820 [hep-ph]`.

[39] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, "ALPGEN, a generator for hard multiparton processes in hadronic collisions," *JHEP* **0307** (2003) 001, `arXiv:hep-ph/0206293 [hep-ph]`.

[40] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," *Journal of High Energy Physics* **2002** no. 06, (2002) 029. `http://stacks.iop.org/1126-6708/2002/i=06/a=029`.

[41] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms," *JHEP* **0411** (2004) 040, `arXiv:hep-ph/0409146 [hep-ph]`.

[42] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method," *JHEP* **0711** (2007) 070, `arXiv:0709.2092 [hep-ph]`.

[43] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," *JHEP* **1006** (2010) 043, `arXiv:1002.2581 [hep-ph]`.

[44] M. Kobayashi and T. Maskawa, "CP-Violation in the Renormalizable Theory of Weak Interaction," *Progress of Theoretical Physics* **49** (Feb., 1973) 652–657.

[45] S. L. Glashow, J. Iliopoulos, and L. Maiani, "Weak interactions with lepton-hadron symmetry," *Phys. Rev. D* **2** (Oct, 1970) 1285–1292. `http://link.aps.org/doi/10.1103/PhysRevD.2.1285`.

[46] P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 2003.

[47] **ATLAS** Collaboration, G. Aad *et al.*, "Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics," `arXiv:0901.0512 [hep-ex]`.

[48] J. Gunion, H. Haber, G. Kane, and S. Dawson, *The Higgs Hunter's Guide*. Frontiers in Physics. Perseus Publishing, 2000. `http://books.google.de/books?id=M5moXN_SA-MC`.

[49] **CDF** Collaboration, T. Aaltonen *et al.*, "First Observation of Electroweak Single Top Quark Production," *Phys.Rev.Lett.* **103** (2009) 092002, `arXiv:0903.0885 [hep-ex]`.

[50] **D0** Collaboration, V. Abazov *et al.*, "Observation of Single Top Quark Production," *Phys.Rev.Lett.* **103** (2009) 092001, `arXiv:0903.0850 [hep-ex]`.

[51] N. Kidonakis, "Top quark pair and single top production at Tevatron and LHC energies," *PoS* **ICHEP2010** (2010) 059, `arXiv:1008.2460 [hep-ph]`.

[52] CERN, "LHC Design Report." `http://lhc.web.cern.ch/lhc/LHC-DesignReport.html`.

[53] M. Syphers, "Overview of the LHC detector." `http://atlas.physics.arizona.edu/~johns/downloads/linac/UTclass2lhc.pdf`.

[54] **ATLAS** Collaboration, G. Aad *et al.*, "Luminosity Determination in pp Collisions at sqrt(s)=7 TeV Using the ATLAS Detector at the LHC," *Eur.Phys.J.* **C71** (2011) 1630, `arXiv:1101.2185 [hep-ex]`.

[55] **ATLAS** Collaboration, "ATLAS inner detector: Technical design report. Vol. 1,".

[56] *ATLAS calorimeter performance: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996.

[57] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996.

[58] *ATLAS level-1 trigger: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1998.

[59] S. Ask, D. Malon, T. Pauly, and M. Shapiro, "Report from the luminosity task force," Tech. Rep. ATL-GEN-PUB-2006-002. ATL-COM-GEN-2006-003. CERN-ATL-COM-GEN-2006-003, CERN, Geneva, Jul, 2006.

[60] T. Luminosity Group, "Improved luminosity determination in **pp** collisions at $\sqrt{s} = \mathbf{7}$ tev using the atlas detector at the lhc," Tech. Rep. ATLAS-COM-CONF-2012-086, CERN, Geneva, May, 2012.

[61] B. D. GIROLAMO, "ATLAS Roman Pots at IR1," in *EDR of the TOTEM Roman Pots*. Nov., 2005. `http://indico.cern.ch/conferenceDisplay.py?confId=330`.

[62] M. ORIUNNO, "Roman Pot Design," in *EDR of the TOTEM Roman Pots.* Nov., 2005.
`http://indico.cern.ch/conferenceDisplay.py?confId=330`.

[63] J. Jackson, *Classical Electrodynamics.* Wiley, 1999.
`http://books.google.de/books?id=U3LBQgAACAAJ`.

[64] V. Hedberg, "LUCID - The ATLAS Luminosity Monitor," in *LHC Experiment Accelerator Data
Exchange Working Group (LEADE).* March, 2005. `http://lhc-data-
exchange.web.cern.ch/lhc-data-exchange/mar05/hedberg_mar05.pdf`.

[65] *Atlas Computing: technical design report.* CERN, Geneva, 2005. `http://atlas-proj-
computing-tdr.web.cern.ch/atlas-proj-computing-tdr/Html/Computing-TDR.htm`.

[66] M. Asai, "Geant4-a simulation toolkit," *Trans.Amer.Nucl.Soc.* **95** (2006) 757.

[67] A. Krasznahorkay, D. Berge, J. Haller, S. Ask, N. Berger, T. Eifert, and A. Hoecker, "SFrame - A
ROOT Analysis Framework." `http://sframe.sourceforge.net/SFrame/SFrame_-
_A_ROOT_analysis_framework.html`.

[68] B. Acharya *et al.*, "Object selection and calibration, background estimations and MC samples for
the Winter 2012 Top Quark analyses with 2011 data," Tech. Rep. ATL-COM-PHYS-2012-224,
CERN, Geneva, Feb, 2012.

[69] Single top group of ATLAS, "Event Challenge Twiki page for Single Top." `https://twiki.
cern.ch/twiki/bin/viewauth/AtlasProtected/SgTopD3PDAccChallenge2012`.

[70] MCP group, ATLAS, "Recommendations on muon identification from the MCP group."
`https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/
MCPAnalysisGuidelinesRel17MC11a`.

[71] ATLAS Muon Workgroup, "MuidMuonCollection twiki Page." `https:
//twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/MuidMuonCollection`.

[72] E/gamma group of ATLAS, "E/gamma twiki page on electron reconstruction."
`https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/
ElectronReconstruction#Standard_electron_cluster_based`.

[73] Jet/EtMiss group of ATLAS, "Jet inputs twiki page."
`https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/JetInputs`.

[74] Top workgroup of ATLAS, "Recommendations on estimation of fake muons from multijet events
for top analyses for the Moriond 2012 conference." `https://twiki.cern.ch/twiki/bin/
viewauth/AtlasProtected/TopFakesMorind2012MmMujetsB`.

[75] B. Alvarez *et al.*, "Measurement of Single Top-Quark Production in the Lepton+Jets Channel in
pp Collisions at sqrt(s)=7 TeV," Tech. Rep. ATL-PHYS-INT-2011-049, CERN, Geneva, May,
2011.

List of Figures

List of Tables