

A Very Impressive and Fancy Title for a Thesis

by

Giovanni Pederiva

THESIS

for the degree of

MASTER OF SCIENCE



Faculty of Mathematics and Natural Sciences
University of Oslo

May 2018

Contents

I	Introduction to QCD and Lattice Field Theories	5
1	A Primer on QCD in the Continuum	7
1.1	The QCD Lagrangian	8
1.1.1	Feynman Rules of QCD	9
1.1.2	Gauge Symmetry of the Lagrangian	10
1.2	General Properties of QCD	11
1.2.1	Running Coupling	12
1.3	Methods and Regimes of Chromodynamics	14
1.4	Experimental Tests of QCD	15
2	Lattice Field Theories and Lattice QCD	17
2.1	Discretizing QCD on the Lattice	17
2.1.1	Naive Discretization of Fermions	18
2.1.2	The Gauge Transporter and the Wilson Loop	19
2.1.3	Lattice Fermions	20
2.2	Path Integrals on the Lattice	20
2.2.1	Pure Gauge Field Theory	22
2.2.2	Observables	22
3	The Gradient Flow Method	25

3.1	Perturbative Analysis of the Wilson Flow	25
4	Discretization Effects in Lattice Yang-Mills Theories and Scale Fixing	27
II	Implementation	29
5	Designing a Lattice $SU(3)$ Yang-Mills Theory Code	31
5.1	Generating Pure Gauge Fields	31
5.1.1	The Metropolis Algorithm	32
5.1.2	Sampling the Configuration Space	33
5.1.3	Updates Strategies	34
5.1.4	Parallelization Scheme	34
5.1.5	Summary of the Parameters	35
5.2	Wilson Flow of Gauge Configurations	35
5.2.1	The Action Derivative	35
5.2.2	Exponential of a $\mathfrak{su}(3)$ Element	36
5.2.3	Parallelization Scheme	36
5.3	Structure and Tools	36
6	Tests and Runs Description	37
6.1	Generated Ensembles	37
6.2	Test Runs	38
6.2.1	Strong and Weak Scaling	38
6.2.2	Autocorrelation of Observables	39
6.3	Production Runs and Timing	39

III	Data Analysis and Results	41
7	Raw Observables	43
8	Running Coupling and Scale Fixing	45
IV	Conclusion and Discussion	47
9	Summary and Conclusion	49
10	Future Developements	51
V	Appendices	53

Abstract

some abstract

Acknowledgements

Part I

Introduction to QCD and Lattice Field Theories

Chapter 1

A Primer on QCD in the Continuum

The Standard Model of particle physics (SM) is the theory of fundamental particles and their interactions. Three of the four known fundamental forces are described by it, the exception is gravity, so all phenomena regarding the Electromagnetic, Weak and Strong forces are included within the theory. The SM describes particles (fermions) which can be quarks or leptons, that interact through the action mediated by gauge fields (bosons) that carry the forces, the photon, gluons, W and Z are such bosons.

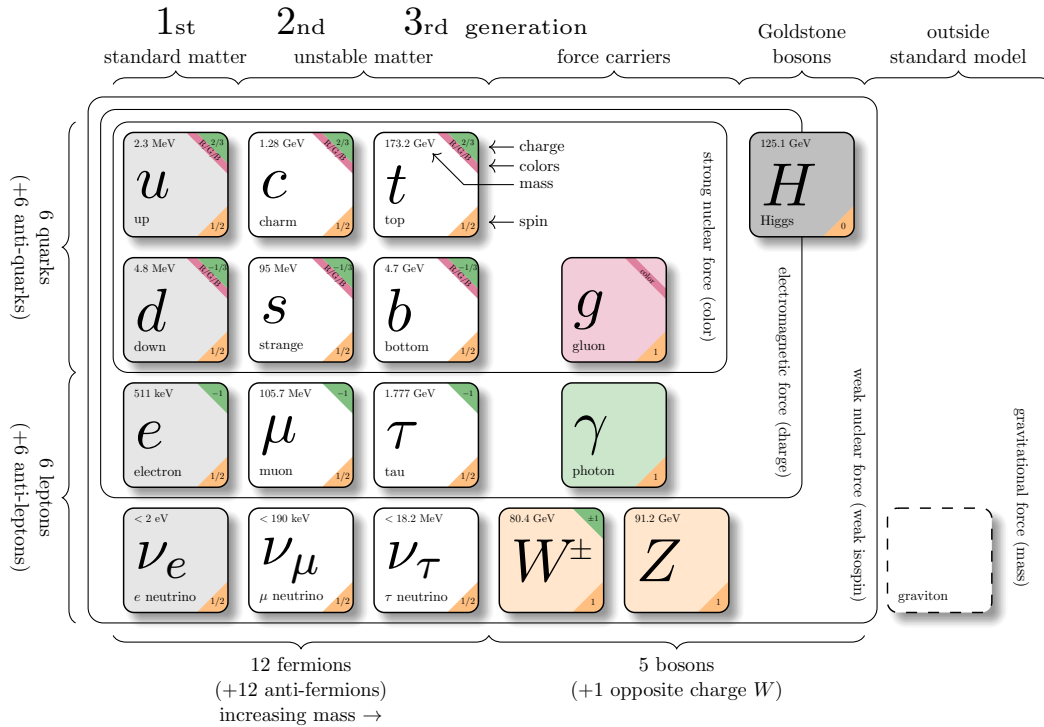


Figure 1.1: Summary of the particles on the Standard Model

From a group theory point of view, the SM is the composition of three different symmetry groups, each associated with a fundamental force:

$$SU(3)_C \times \underbrace{SU(2)_L \times U(1)_Y}_{\text{broken to } SU(2)_W \times U(1)_{EM}} \quad (1.1)$$

The $U(1)_{EM}$ symmetry group is associated with the electromagnetic interaction and derives from the spontaneously broken symmetry $U(1)_Y$, hyper-charge, that together with the $SU(2)_L$ through the Higgs mechanism recover also the term is associated with the weak force.

Quantum Chromodynamics, commonly referred to as QCD, is the quantum field theory, contained in the SM, that describes the behavior of strongly interacting matter, that is quarks and gluons. It is a non-abelian gauge theory based on a $SU(3)$ symmetry group. In this chapter we will discuss the QCD lagrangian density, its properties and some of the major results of the theory. Some intermediate knowledge of Quantum Field Theory is assumed and derivations and proofs mainly follow the reasoning found in [1].

1.1 The QCD Lagrangian

In quantum field theory the characterizing equation of a theory is its lagrangian density, because it contains all the information about the fields that are involved, their properties and most importantly, their interactions. For QCD the simplest form is based on the Yang-Mills Lagrangian, with a $SU(3)$ symmetry group:

$$\mathcal{L}_{QCD} = -\frac{1}{4}(G_{\mu\nu}^a)^2 + \sum_{f=1}^{n_f} \bar{\psi}_f(i\not{D} - m_f)\psi_f \quad (1.2)$$

Here ψ_f represents the complex-valued fermion field of flavor f , with mass m_f . These fields are the quark fields that come in six flavors: u (up), d (down), s (strange), c (charm), b (bottom) and t (top). The second element in the lagrangian is the Gluon Field Strength Tensor, $G_{\mu\nu}^a$. The two indices μ and ν are Lorentz indices and a is the index of the generators of the gauge group, $SU(3)$ in this case. Note that Einstein summing convention on repeated indices is implicit, for example when taking the square of the field strength tensor three sums are applied. The definition of $G_{\mu\nu}^a$ is:

$$G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c \quad (1.3)$$

in this equation A_μ^a is the gluon field, that carries a Lorentz index and a group generator index and the f^{abc} are the structure constants of $SU(3)$, which satisfy:

$$[t^a, t^b] = if^{abc}t^c \quad (1.4)$$

with t^a being the generators of the algebra $\mathfrak{su}(3)$. The covariant derivative \not{D} is defined then as:

$$\not{D} = \gamma^\mu \partial_\mu - ig\gamma^\mu t_a A_\mu^a \quad (1.5)$$

With the information contained in the lagrangian density the behavior and the interactions of all particles are set.

1.1.1 Feynman Rules of QCD

A key element that is needed to perform perturbative calculations in a quantum field theory are Feynman Rules. These are a set of equations and rules that represent the propagation of fields and the interaction vertices of the theory. For the case of QCD, being a non-abelian gauge theory, some vertices represent interactions between gauge bosons only, as opposed to Quantum Electrodynamics, QED, that forbids photon-photon interactions.

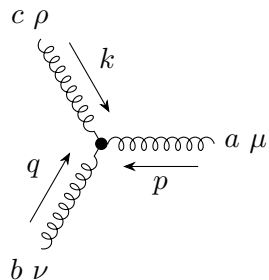
To begin with, we need to write out all of the terms of the lagrangian density individually. We assume only one quark flavor as no term in the lagrangian can change this quantum number.

$$\begin{aligned}
\mathcal{L}_{QCD} &= -\frac{1}{4}(G_{\mu\nu}^a)^2 + \bar{\psi}(i\not{D} - m)\psi \\
&= -\frac{1}{4}\left(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c\right)\left(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c\right) \\
&\quad + \bar{\psi}(i\gamma^\mu\partial_\mu + g\gamma^\mu t_a A_\mu^a - m)\psi \\
&= -\frac{1}{4}(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a)^2 \\
&\quad + \frac{1}{2}f^{abc}(\partial_\nu A_\mu^a - \partial_\mu A_\nu^a)[A^{b\mu}, A^{c\nu}] \\
&\quad - \frac{1}{4}g^2 f^{abc}f^{ade}A_\mu^b A_\nu^c A^{d\mu} A^{e\nu} \\
&\quad + g\gamma^\mu t_a A_\mu^a \bar{\psi}\psi \\
&\quad + \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi
\end{aligned} \tag{1.6}$$

First we define the gluon and quark propagators, which can be obtained from the first and last terms respectively:

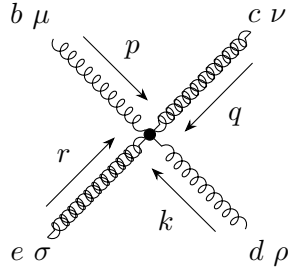
$$\begin{aligned}
a \mu \xrightarrow{k} b \nu &= -\frac{i\delta_{ab}g^{\mu\nu}}{k^2} \\
i \xrightarrow{k} j &= \frac{i\delta_{ij}}{\not{k} - m}
\end{aligned}$$

The third term in 1.6 represents a three gluon vertex:



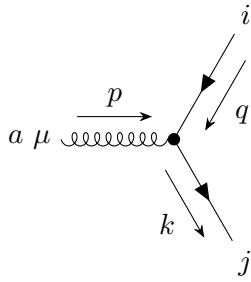
$$= gf^{abc}[g^{\mu\nu}(p-q)^\rho + g^{\nu\rho}(q-k)^\mu + g^{\rho\mu}(k-p)^\nu]$$

Then the four-gluon vertex:



$$= -ig^2 [f^{abe} f^{acd} (g^{\mu\nu} g^{\sigma\rho} - g^{\mu\rho} g^{\nu\sigma}) \\ + f^{abd} f^{ace} (g^{\mu\nu} g^{\sigma\rho} - g^{\mu\sigma} g^{\nu\rho}) \\ + f^{abc} f^{aed} (g^{\mu\sigma} g^{\nu\rho} - g^{\mu\rho} g^{\nu\sigma})]$$

Finally we have the gluon-quark interaction vertex:



$$= 2igt^a \gamma^\mu$$

Note that the lagrangian we considered, and the resulting Feynman rules, is over-simplified: the “full” lagrangian contains terms from Faddeev-Popov ghosts and counter-terms from the renormalization procedure. Nevertheless, interesting qualitative features can be inferred from the Feynman rules: the fact that gluons interact with each other through the three- and four-gluon vertices; the non flavor-changing interaction between gluons and quarks, which decouples completely different quark flavors; the “color-changing nature” of the quark-gluon interaction, given by the t^a matrix in the vertex term that shows how the color state of a fermion is changed by the absorption or emission of a gluon.

1.1.2 Gauge Symmetry of the Lagrangian

The QCD lagrangian must be gauge invariant to be physical. The concept of gauge invariance is crucial in the construction of a discretized lattice theory from the continuum one. Let’s consider a quark field $\psi(x)$ and a local gauge transformation in the internal space of $SU(3)$ applied to it:

$$\psi(x) \rightarrow \psi'(x) = U(x)\psi(x) \quad : \quad U(x) = \exp(i(\alpha^a(x)t^a)) \quad (1.7)$$

here $U(x)$ is an element of the $SU(3)$ gauge group. This transformation if applied to the simple Dirac free-field Lagrangian would generate an additional term:

$$\mathcal{L}_{Dirac} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \rightarrow \mathcal{L}'_{Dirac} = \bar{\psi}U^\dagger(i\gamma^\mu \partial_\mu - m)(U\psi) \\ = \bar{\psi}'(i\gamma^\mu \partial_\mu - m)\psi' + i\bar{\psi}'\gamma^\mu \psi(\partial_\mu U)$$

In order to fix this problem a new field $A_\mu(x)$, the gauge field, is introduced and it enters the definition of the covariant derivative, so that ∂_μ becomes $D_\mu = \partial_\mu - igA_\mu(x)$ as previously

stated. The transformation rule for $A_\mu(x)$ is fixed in order to cancel the extra term in 1.8 exactly, such that:

$$D_\mu\psi \rightarrow (D_\mu\psi)' = (\partial_\mu - igA'_\mu(x))\psi' = U(D_\mu\psi) \quad (1.9)$$

and this fixes the transformation for $A_\mu(x)$ to be:

$$A_\mu(x) \rightarrow A'_\mu(x) = U \left[A_\mu(x) - \frac{i}{g} U^\dagger \partial_\mu U \right] U^\dagger \quad (1.10)$$

we can see that the last term contributes effectively to the lagrangian as $-i\bar{\psi}'\gamma^\mu\psi(\partial_\mu U)$, which is what we want to cancel. For an infinitesimal transformation we can expand the matrix U and get the following for the gauge field:

$$A_\mu^a(x) \rightarrow A'^a_\mu(x) = \alpha^a(x) - \frac{i}{g} \partial_\mu \alpha^a(x) + f^{abc} \alpha^b(x) \alpha^c(x) \quad (1.11)$$

which is the last element needed. One can now look at all the possible gauge invariant objects that can be constructed with the fields ψ and A of order 4, the same of the lagrangian. Apart from the one already present in the Dirac lagrangian with the covariant derivative, there are only two additional terms that are gauge invariants and of order 4 and they both can be taken by considering the gauge field tensor:

$$G_{\mu\nu}^a \equiv \frac{i}{g} [D_\mu, D_\nu] = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f^{abc} A_\mu^b A_\nu^c \quad (1.12)$$

this is clearly gauge invariant since it is commutator of covariant derivatives and it has dimension 2. One can now construct the gauge field kinetic term, the well known $-\frac{1}{4}(G_{\mu\nu}^a)^2$ and reconstruct 1.2. However, there is an additional term, not included in the QCD lagrangian, that is the "dual term", or "theta term", which will be interesting further in the work. It is defined as $\theta G_{\mu\nu}^a \tilde{G}^{a\mu\nu}$ where $\tilde{G}^{a\mu\nu} = \epsilon^{\mu\nu\rho\sigma} G_{\rho\sigma}^a$ is the dual of the field tensor. With it the lagrangian becomes:

$$\mathcal{L}_{QCD} = -\frac{1}{4}(G_{\mu\nu}^a)^2 + \bar{\psi}(i\not{D} - m)\psi + \theta G_{\mu\nu}^a \tilde{G}^{a\mu\nu} \quad (1.13)$$

The theta term is usually neglected because there is no experimental evidence of it, but in principle it cannot be excluded. It is the simplest CP violating term that can be added to the QCD lagrangian and for this is of particular interest in the study of CP phenomena, like the nucleon electric dipole moment (EDM).

1.2 General Properties of QCD

Quantum Chromodynamics exhibits a set of features as a theory that are common to all non-abelian gauge theories. We have already seen one, that is the direct interaction of the gauge bosons, something that is not allowed in abelian theories as QED. Other interesting properties emerge when trying to renormalize the theory and are in general linked to the fixing of the scale, which leads to the concept of running coupling. In the particular case of QCD the coupling constant at in a low-energy regime leads to Confinement, while in the high-energy limit Asymptotic Freedom emerges. The discovery of asymptotic freedom by Gross, Wilczek and Politzer [2][3], was used as an indication that QCD is indeed the correct theory of the strong interaction.

1.2.1 Running Coupling

The beta function of the coupling constant, $\beta(g)$ defines the rate at which the renormalized coupling varies as the renormalization scale Q changes.

$$\beta(g) = \frac{d}{d \log(\mu)} g(\mu) \quad (1.14)$$

For a generic abelian theory $SU(N)$ one can expand the β function in orders of g as:

$$\beta(g) = \beta_0 g^3 + \beta_1 g^5 + \beta_2 g^7 + \dots \quad (1.15)$$

One can then integrate up to arbitrary order 1.14 and get an expression for $g(\mu)$. The coefficients β_i are obtained from computing contribution of higher and higher diagrams to the coupling. The value of β_0 , from 1-loop corrections, is:

$$\beta_0 = -\frac{g^3}{(4\pi)^2} \left(\frac{11}{3}N - \frac{2}{3}n_f \right) \quad (1.16)$$

the minus sign in front implies that any abelian theory with a sufficiently small number of fermions, less than $\frac{33}{2}$ (that is 16 in practice), is asymptotically free.

The renormalization group equation is usually expressed in terms of the analogous of the fine structure-constant for the strong force, $\alpha_s = g^2/4\pi$. The first order solution is given by plugging 1.16 into 1.14. In terms of α_s at a scale μ we get:

$$\alpha_s(\mu) = \frac{\alpha_s}{1 + \frac{\beta_0 \alpha_s}{4\pi} \log(\mu^2/M^2)} \quad (1.17)$$

The typical choice for μ is the mass of the Z boson, where QCD can be compared relatively simply to the other forces. The coupling depends on an arbitrarily chosen renormalization point M . A convenient choice is to define a mass scale Λ that satisfies:

$$1 = g^2 \left(\frac{\beta_0}{4\pi^2} \right) \log(M^2/\Lambda^2) \quad (1.18)$$

check constants... This simplifies 1.17 to its well-known form, correct up to one loop corrections:

$$\alpha_s(Q^2) = \frac{4\pi}{\beta_0 \log(\mu^2/\Lambda^2)} \quad (1.19)$$

In a later section (**REFLINK NEEDED**) we will show a more precise result, correct up to order 4, of this result. Our main goal for this thesis is to find a simple and not expensive way to estimate the mass scale parameter Λ from lattice calculations. In 1.2 a higher order approximation of $\alpha_s(\mu)$ is plotted against some experimental values. Now we will talk briefly of what this result implies.

Asymptotic Freedom

Equation 1.19 is a clear indication that QCD at high energies has a small coupling. Asymptotic freedom is the property of gauge theories, QCD is usually the example for it, that causes the

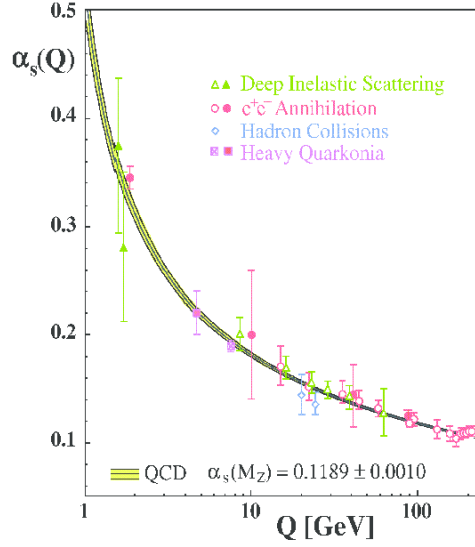


Figure 1.2: The strong coupling as a function of the energy scale. Image from [4]

interactions between the fields do become weaker as the energy scale increases. This is for example one of the basis on which unification theories are based on, the fact that there exists an energy scale at which the strong force has a coupling equivalent to the one of the electroweak interaction.

Asymptotically free theories, can be analyzed perturbatively at sufficiently large energies and are believed to be consistent up to any energy scale.

Confinement

At low energy scales from 1.19 we can infer that the coupling constant increases exponentially and approaches 1. This is the reason for the fundamentally different nature of the strong force compared to the other forces, no perturbative expansion can be made at low energies. The exact proof of how this links with the color confinement of QCD is yet not known, but qualitatively it can be explained by the fact that the gauge bosons of the theory, the gluons, carry color charge just like the quarks. In general color charged particles cannot be isolated, so that quarks and gluons are not detectable alone, but always in the form of hadrons: colorless objects formed of multiple quarks, mesons (quark-anti-quark) and baryons (three quarks or three anti-quarks). Also glueballs, combinations of gluons such that the total is colorless, are in principle allowed, but have not been observed yet. Confinement is the phenomenon for which it is not possible to isolate a color charge, single quarks or gluons, from a hadron without producing other new hadrons. Single colored particle in a very small time scale undergo hadronization, the process of spawning new quarks or anti-quarks from the vacuum to balance the total color charge and produce colorless matter.

The other usual picture is to consider the gluons being exchanged by two quarks to form flux-

tubes that, if stretched by separating the quark sources, eventually store enough energy to make a quark-anti-quark pair energetically favorable, as depicted in figure 1.3

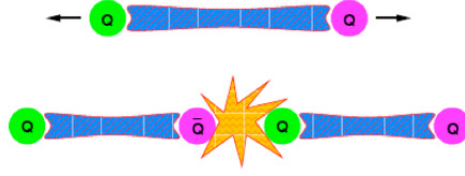


Figure 1.3: Representation of confinement using flux-tubes. As two colour sources are pulled apart, the energy stored in the colorfield between the sources increases so much that a $q\bar{q}$ is formed at some point.

1.3 Methods and Regimes of Chromodynamics

Given the very different behaviors of the strong force at different energy scales, QCD needs to be dealt with in various ways depending on the scale of interest. At high energies perturbation theory can be applied safely, but at low energies no expansion in the coupling constant can be made.

Perturbative QCD

At high energies, like the scale of the large particle accelerators that are colliders currently available, the QCD coupling constant is sufficiently small to allow analytical calculations of Feynman diagrams to be meaningful. Hadronization is neglected as the time scale is small enough to consider it a post-collision process. For example, at a scale where the strong force is comparable to the electroweak interaction, the cross section of $e^+e^- \rightarrow \text{hadrons}$ compared to that of $e^+e^- \rightarrow \mu^+\mu^-$ can be used to measure the number of quark flavors that are active below that energy.

Lattice Methods

To deal with the non-perturbative sector of the strong interaction the most widespread approach is to use numerical simulations, in particular lattice methods, so it is common to refer to it as Lattice QCD. The main idea, which will be expressed more in detail in Chapter 2, is to discretize space-time and evaluate the field only at fixed sites on a hyper-cubic lattice, physical quantities are then computed stochastically on ensembles of such gauge fields. This approach however is very expensive from a computational point of view and so far no calculation at physical quark masses with a sufficiently small lattice spacing, in principle closer to the continuum theory, have been performed.

Effective Field Theories

An interesting problem is to link QCD directly with nuclear forces, that are long range remnants of the strong interaction at a hadron level. The problem is of high interest because there is yet no fundamental theory of nuclear interactions from first principles. The most common approach is to define an Effective Field Theory, starting from a low energy approximation of chromodynamics, that preserves most of the symmetries of the underlying theory. Some common approaches start from considering nucleons as a fundamental $SU(2)$ group in isospin and recognizing pions and kaons as the goldstone bosons of the nuclear forces. **probably need more...**

1.4 Experimental Tests of QCD

SOME NICE PLOTS...

Chapter 2

Lattice Field Theories and Lattice QCD

Lattice QCD is one of the main Lattice Field Theories. It deals with the strong force in a numerical way. The idea of discretizing space-time in a lattice and perform calculations of field theories in such a model is attributed to Wilson (CITATION NEEDED). It has proven to be the most systematic approach to non-perturbative theories like Yang-Mills theory and QCD. In this chapter we will describe briefly how this discretization procedure is performed and some of the main computational strategies that are involved.

2.1 Discretizing QCD on the Lattice

The starting point for Lattice QCD is Feynman's path-integral formalism, but expressed in Euclidean space-time, through a Wick rotation. An observable of some field ϕ is then given by:

$$\langle O[\phi] \rangle = \frac{1}{Z[\phi]} \int \mathcal{D}[\phi] O[\phi] e^{-S[\phi]} \quad (2.1)$$

where the partition function $Z[\phi]$ is defined as:

$$Z = \int \mathcal{D}[\phi] e^{-S[\phi]} \quad (2.2)$$

and $S[\phi]$ is the classical action of the field. Evaluating path-integrals is not possible in general with analytical tools so, in order to allow numerical computations, the Euclidean space-time is discretized on a hyper-cubic lattice $L = (L_x, L_y, L_z, L_t)$. The choice of the lattice spacing, usually denoted a , is arbitrary, but most often it is chosen to be equal for all dimensions. if we then define a lattice site $n = (n_x, n_y, n_z, n_t)$ where all the n s represent the coordinates of a point in the lattice Λ , our fields are constrained to have values on the points an instead of on a continuum space-time x^μ .

$$\phi(x) \xrightarrow{\text{discretization}} \phi(an) \quad (2.3)$$

In the case of QCD there are two types of field at play, the gluon gauge field A and the n_f fermionic quark fields ψ .

2.1.1 Naive Discretization of Fermions

We will now follow the fermion discretization procedure found in (CITATION NEEDED)gatringer, quoting the main intermediate steps that lead to the formulation of Lattice QCD. The starting point is the fermionic euclidean action in the continuum:

$$S_F[\psi, \bar{\psi}] = \int d^4x \bar{\psi}(x) (\gamma_\mu \partial^\mu + m) \psi(x) \quad (2.4)$$

now we discretize the euclidean space-time on a lattice Λ of spacing a , each point will be denoted with n . The partial derivative can be turned into the central finite difference between neighborintg points along the direction of the derivative:

$$\partial_\mu \psi(x) \rightarrow \frac{\psi(n + \hat{\mu}) - \psi(n - \hat{\mu})}{2a} \quad (2.5)$$

The discretized fermion action is then:

$$S_F[\psi, \bar{\psi}] = a^4 \sum_{n \in \Lambda} \bar{\psi}(n) \left[\sum_{\mu=1}^4 \gamma_\mu \frac{\psi(n + \hat{\mu}) - \psi(n - \hat{\mu})}{2a} + m \psi(n) \right] \quad (2.6)$$

As we did in section 1.1.2 we try to apply a local gauge transformation $\Omega(n)$ to the field. It is simple to show that the terms of the derivative in the action are not gauge invariant:

$$\bar{\psi}(n) \psi(n \pm \hat{\mu}) \rightarrow \bar{\psi}(n) \Omega^\dagger(n) \Omega(n \pm \hat{\mu}) \psi(n \pm \hat{\mu}) \quad (2.7)$$

The problem is in the way the derivative was discretized, it also removed one Lorentz index so it had to be wrong, and the solution is the introduction of an additional field as in 1.1.2 that has the correct transformation laws. This field must connect the values of the fermion field at two different lattice sites and because of this it is denoted "link variables" and by convention it is denoted as $U(n, n + \hat{\mu}) = U_\mu(n)$. Furthermore, it must depend on the direction along the dimension μ in a simple way, $U(n, n - \hat{\mu}) \equiv U_\mu(n - \hat{\mu})^\dagger = U_{-\mu}(n)$. In particular, we want it to transform as:

$$\begin{aligned} \bar{\psi}(n) U_\mu(n) \psi(n + \hat{\mu}) &\rightarrow \bar{\psi}(n) \Omega^\dagger(n) U'_\mu(n) \Omega(n + \hat{\mu}) \psi(n + \hat{\mu}) \\ \bar{\psi}(n) U_{-\mu}(n) \psi(n - \hat{\mu}) &\rightarrow \bar{\psi}(n) \Omega^\dagger(n) U'_{-\mu}(n) \Omega(n - \hat{\mu}) \psi(n - \hat{\mu}) \end{aligned} \quad (2.8)$$

from which we infer that:

$$\begin{aligned} U_\mu(n) &\rightarrow U'_\mu(n) = \Omega(n) U_\mu(n) \Omega^\dagger(n + \hat{\mu}) \\ U_{-\mu}(n) &\rightarrow U'_{-\mu}(n) = \Omega(n) U_{-\mu}(n) \Omega^\dagger(n - \hat{\mu}) \end{aligned} \quad (2.9)$$

the field $U_\mu(n)$ which we identify with the link variables depends on the direction of the move along the μ direction. In particular the relation is simply . **FIG: links...** With this results we can write a gauge invariant lattice fermion action as:

$$S_F[\psi, \bar{\psi}] = a^4 \sum_{n \in \Lambda} \bar{\psi}(n) \left[\sum_{\mu=1}^4 \gamma_\mu \frac{U_\mu(n) \psi(n + \hat{\mu}) - U_{-\mu}(n) \psi(n - \hat{\mu})}{2a} + m \psi(n) \right] \quad (2.10)$$

2.1.2 The Gauge Transporter and the Wilson Loop

A more formal definition of the link variables we can look at the gauge transporter. This is the path-ordered product, denoted with \mathcal{P} of a gauge field $A(x)$ along some curve \mathcal{C} between two points in space-time:

$$G(x, y) = \mathcal{P} \exp \left(i \int_x^y A(x') dx' \right) \quad (2.11)$$

An important thing to note now is that link variables belong to the gauge group, $SU(3)$ for QCD, and not to the algebra, $\mathfrak{su}(3)$. With this definition is easy to see the relation we just stated earlier on the direction change:

$$U(n, n - \hat{\mu}) = U_{-\mu}(n) = U_{\mu}^{\dagger}(n) \quad (2.12)$$

wilson loop from peskin... On the lattice the minimal Wilson loop is just a square:

$$\begin{aligned} P_{\mu\nu}(n) &= U_{\mu}(n) U_{\nu}(n + \hat{\mu}) U_{-\mu}(n + \hat{\mu} + \hat{\nu}) U_{-\nu}(n + \hat{\nu}) \\ &= U_{\mu}(n) U_{\nu}(n + \hat{\mu}) U_{\mu}^{\dagger}(n + \hat{\nu}) U_{\nu}^{\dagger}(n) \end{aligned} \quad (2.13)$$

a pictorial representation is given in figure **FIG: PLAQUETTE** Up to order $\mathcal{O}(a)$ the integral on the straight line connecting two points on the lattice can be approximated with $aA_{\mu}(n)$, giving us $U_{\mu}(n) = \exp(iaA_{\mu}(n))$. This allows us to write using the Baker-Campbell-Hausdorff formula:

$$\begin{aligned} P_{\mu\nu}(n) &= \exp \left[iaA_{\mu}(n) + iaA_{\nu}(n + \hat{\mu}) - iaA_{\mu}(n + \hat{\nu}) - iaA_{\nu}^{\dagger}(n) - \frac{a^2}{2} [A_{\mu}(n), A_{\mu}(n + \hat{\mu})] \right. \\ &\quad - \frac{a^2}{2} [A_{\nu}(n + \hat{\nu}), A_{\nu}(n)] + \frac{a^2}{2} [A_{\mu}(n), A_{\nu}(n)] + \frac{a^2}{2} [A_{\nu}(n + \hat{\mu}), A_{\mu}(n + \hat{\nu})] \\ &\quad \left. + \frac{a^2}{2} [A_{\mu}(n), A_{\mu}(n + \hat{\nu})] + \frac{a^2}{2} [A_{\nu}(n + \hat{\mu}), A_{\nu}(n)] + \mathcal{O}(a^3) \right] \end{aligned} \quad (2.14)$$

Now the terms that are shifted from the site n are expanded as:

$$A_{\mu}(n + \hat{\nu}) = A_{\mu}(n) + a\partial_{\nu}A_{\mu}(n) + \mathcal{O}(a^2) \quad (2.15)$$

and with this substitution most terms cancel and we are left with:

$$\begin{aligned} P_{\mu\nu}(n) &= \exp [ia^2(\partial_{\mu}A_{\nu}(n) - \partial_{\nu}A_{\mu}(n) + i[A_{\mu}(n), A_{\nu}(n)]) + \mathcal{O}(a^3)] \\ &= \exp [ia^2F_{\mu\nu}(n) + \mathcal{O}(a^3)] \end{aligned} \quad (2.16)$$

This term can be used to build the euclidean lattice action term for the gluons. In particular, we would like a term of the form:

$$S_G[U] = \frac{a^4}{2g^2} \sum_{n \in \Lambda} \sum_{\mu\nu} \text{Tr}(F_{\mu,\nu}(n)^2) \quad (2.17)$$

Up to order $\mathcal{O}(a^2)$ this can be obtained by:

$$S_G[U] = \frac{2}{g^2} \sum_{n \in \Lambda} \sum_{\mu < \nu} \Re \text{Tr}(\mathbf{1} - P_{\mu\nu}(n)) \quad (2.18)$$

Higher order corrections to this action can be computed analytically by considering higher orders in the BCH expansion and in the exponential expansion when constructing 2.18.

2.1.3 Lattice Fermions

The discretization of fermions as discussed in the previous section is incomplete. It becomes evident if one considers the Fourier transform of the propagator. Let's first rewrite 2.10 in a more compact way, introducing the lattice Dirac propagator $M_{xy}[U]$.

$$S_F[\psi, \bar{\psi}] = \sum_{n \in \Lambda} \bar{\psi}(x) M_{xy}[U] \psi(y) \quad (2.19)$$

with

$$M_{xy}[U] = \sum_{\mu=1}^4 \gamma_{\mu} \frac{U_{\mu}(x) \delta_{x, (y-\hat{\mu})} - U_{-\mu}^{\dagger}(x) \delta_{x, (y+\hat{\mu})}}{2a} + m \delta_{x,y} \quad (2.20)$$

in momentum space the propagator becomes:

$$\tilde{M}_{pq} = \delta(p-q) \tilde{M}(p) \quad : \quad \tilde{M}(p) = m \mathbb{1} + \frac{i}{a} \sum_{\mu=1}^4 \gamma_{\mu} \sin(p_{\mu} a) \quad (2.21)$$

In order to calculate the inverse of the propagator in real space we need to invert the one in momentum space and perform an inverse Fourier transform. However, the inverse of the propagator in Fourier space has multiple poles:

$$\tilde{M}(p)^{-1} \Big|_{m=0} = \frac{-ia \sum_{\mu} \gamma_{\mu} \sin(p_{\mu} a)}{\sum_{\mu} \sin^2(p_{\mu} a)} \quad (2.22)$$

the problem vanishes for $a \rightarrow 0$, the continuum case, returning just one fermion type, but on the lattice multiple fermions. This is known as the "doubling problem". The solution, proposed by Wilson, is to modify the propagator adding some terms that make the poles in the inverse of the Fourier transformed propagator vanish. The final form of the Wilson Fermion Action is:

$$S_F[\psi, \bar{\psi}] = \sum_{n \in \Lambda} \bar{\psi}(x) M_{xy}^W[U] \psi(y) \quad (2.23)$$

with $m_{xy}^W[U]$ being the Wilson propagator:

$$M_{xy}^W[U] = \frac{1}{2a} \sum_{\mu=\pm 1}^{\pm 4} (\mathbb{1} - \gamma_{\mu}) U_{\mu}(x) \delta_{x, (y-\hat{\mu})} + \left(m + \frac{4}{a}\right) \delta_{x,y} \quad (2.24)$$

the shorthand notation $\gamma_{-\mu} = -\gamma_{\mu}$ has been introduced.

Now that all the needed information about the action and the fields is set, mainly through equations 2.10 and 2.23, the picture of how to discretize QCD from the continuum Minkowski space-time to an euclidean space-time lattice.

2.2 Path Integrals on the Lattice

To express expectation values and correlators on the lattice, path integral formalism is used. The partition function, as we have seen earlier, is the path integral of the fields over the whole

space of the action. For the case of QCD the fields are U , ψ and $\bar{\psi}$:

$$Z = \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}U e^{-S[\psi, \bar{\psi}, U]} \quad (2.25)$$

with the action being the sum of the gluonic and fermionic parts:

$$S[\psi, \bar{\psi}, U] = S_G[U] + S_F[\psi, \bar{\psi}, U] = S_G[U] + \sum_f \bar{\psi} M \psi \quad (2.26)$$

The immediate simplification is to integrate out the fermion fields. As in the continuum case one can perform an integration on the Grassmann-valued fields, in general for an integral over some Grassmann numbers θ_i and their complex conjugates θ_i^* , and a Hermitean matrix K :

$$\begin{aligned} \int \mathcal{D}\theta^* \mathcal{D}\theta e^{-\Theta K \Theta} &= \left(\prod_i \int d\theta_i^* d\theta_i \right) e^{-\theta_i^* K_{ij} \theta_j} = \left(\prod_i \int d\theta_i^* d\theta_i \right) e^{-\sum_i \theta_i^* k_i \theta_i} \\ &= \prod_i b_i = \det B \end{aligned} \quad (2.27)$$

This result is very different from what one would get in the real case, $(2\pi)^n / \det B$. It can also be shown that:

$$\begin{aligned} \int \mathcal{D}\theta^* \mathcal{D}\theta \theta_a^* \theta_b e^{-\Theta K \Theta} &= \left(\prod_i \int d\theta_i^* d\theta_i \right) \theta_a^* \theta_b e^{-\theta_i^* K_{ij} \theta_j} = \left(\prod_i \int d\theta_i^* d\theta_i \right) \theta_a^* \theta_b e^{-\sum_i \theta_i^* k_i \theta_i} \\ &= (\det B) (B^{-1})_{ab} \end{aligned} \quad (2.28)$$

This last result is crucial for computing fermion propagators for example, for a given "source" $\bar{\psi}(x)$ and a "sink" $\psi(y)$ the propagator between the two can be computed via path integrals, but it requires inverting the fermion action matrix.

With the result of 2.27 we can simplify greatly the partition function:

$$Z = \int \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}U e^{-S[\psi, \bar{\psi}, U]} = \int \mathcal{D}U e^{-S_G[U]} \det M[U] \quad (2.29)$$

In a similar fashion as in statistical mechanics, the expectation value of an observable on the lattice can be computed as:

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}U O(\psi, \bar{\psi}, U) e^{-S_G[U]} \det M[U] \quad (2.30)$$

The above expression cannot be evaluated or simplified analytically any further, so the usual approach is to approximate the path integral numerically. The main idea is to create an ensemble of field configurations to reproduce the integral $\int \mathcal{D}$, on such set $\mathcal{U} = \{U_1, U_2, \dots, U_N\}$ one computes the observable and the average value is the expectation value of the observable:

$$\langle O \rangle \approx \frac{1}{N} \sum_{i=1}^N O(\psi, \bar{\psi}, U_i) \quad (2.31)$$

The choice of the set \mathcal{U} is the interesting part, and it starts from identifying parts of 2.30 as a probability distribution, in particular, the probability of a configuration U is identified with:

$$P[U] = \frac{e^{-S_G[U]} \det M}{Z} \quad (2.32)$$

The integral is then evaluated using Monte Carlo integration, where the different configurations are chosen in the most widely accepted solution via a Markov chains. Given a field configuration U_i one chooses the following configuration U_{i+1} only based on properties of U_i . Later we will see how the Metropolis algorithm, one of the simplest Markov Chain Monte Carlo methods has been implemented.

2.2.1 Pure Gauge Field Theory

Computing full QCD on the lattice is computationally expensive, mainly due to the integration of fermions via the determinant. From a numerical point of view, the determinant need to be computed at every step of the Markov chain that is used to evaluate the path-integral and this operation affects the time cost of sampling the configuration space dramatically. A first approach is to neglect the determinant completely, considering it constant. This is effectively removing dynamical fermions, freezing them to the lattice sites. This approximation is usually referred to as "quenched QCD", or QQCD. The properties of this theory, that is then reduced to a simple Yang-Mills theory are still interesting to study and have played historically a very important role, being the only accessible simulation until sufficient computing power became available.

2.2.2 Observables

On the lattice, given the transformation 2.9, any product of link variables that starts and end at the same lattice-site, a closed loop, is gauge invariant. The average values of these objects over the whole lattice can be linked to physical observables, for example the field tensor. In a more general form any observable $L[U]$ defined as

$$L[U] = \text{Tr} \left[\prod_{(n,\mu) \in \mathcal{L}} U_\mu(n) \right] \quad (2.33)$$

where \mathcal{L} is a closed loop of links on the lattice is a gauge invariant object and a candidate observable.

Plaquette

The simplest observable, which we have already encountered upon defining the Wilson action in 2.10, is the plaquette. This is the minimal closed loop on the lattice and its value is related to the coupling constant of the action. For each lattice site there are 12 possible plaquettes

to be computed, given all the combinations of euclidean indeces. A proper definition of the observable is:

$$P[U] = \frac{1}{6|\Lambda|} \sum_{n \in \Lambda} \sum_{\mu < \nu} P_{\mu\nu}(n) \quad (2.34)$$

where $P_{\mu\nu}(n)$ is the one defined in 2.14. As a side note, in actual calculations at every lattice site only the positive sign link variables are stored, so the second line of 2.14 is the one that is actually used in simulations.

Energy Density

The energy of the field is proportional to the square of the field tensor, in particular:

$$E[U] = -\frac{1}{4|\Lambda|} G_{\mu\nu} G^{\mu\nu} \quad (2.35)$$

In order to estimate this quantity, one has to compute the field tensor at every lattice site, square it and sum over the whole space. It is then usually normalized by the lattice volume (the number of sites), to get the density. The simplest definition of the field tensor is $E_p = \mathbb{1} - P_{\mu\nu}$ but this is not very accurate. A more symmetric definition can be obtained by the "clover", that is summing all the plaquettes of a same plane that start from a given lattice site. **Expression for the clover lattice calc** **FIG: clover**

Topological Charge

The gauge fields in QCD exhibit particular topological properties that are believed to have important physical implications. One example, which will be discussed more thoroughly in section **(REFLINK NEEDED)**, is the relation to the mass of the η' meson, the flavor-singlet meson **(CITATION NEEDED)**. The topological charge is an integer quantum number of the field in the continuum, but on the lattice certain definitions can be used to reproduce the continuum properties, especially for the so called "topological susceptibility", that is the second moment of the distribution of the topological charge, which seems to be independent from the definitions of the base observable **(CITATION NEEDED)**. In the continuum topological sectors, regions of space with same charge, are separated from each other, on the lattice through discretization effects the behavior is different, with instantons that allow tunneling between sectors **(CITATION NEEDED)**.

The topological charge is the integral over all space-time of the topological charge density:

$$Q = \int d^4x q(x) \quad (2.36)$$

where

$$q(x) = \frac{1}{32\pi^2} \epsilon_{\mu\nu\rho\sigma} \text{Tr}(F_{\mu\nu} F_{\rho\sigma}) \quad (2.37)$$

this can be estimated on the lattice, in the simplest way, by using the same definition of the energy tensor we used before, the clover:

$$Q[U] = \frac{a^4}{32\pi^2} \sum_{n \in \Lambda} \sum_{\mu < \nu} \epsilon_{\mu\nu\rho\sigma} \text{Tr}[G_{\mu\nu}(n)G_{\rho\sigma}(n)] \quad (2.38)$$

As the lattice spacing is reduced, approaching the continuum, the topological sectors get more and more separated, preventing tunneling between them. This makes the Markov chain used to generate the ensemble less efficient in terms of growing autocorrelation times, as will be shown in (REFLINK NEEDED).

Chapter 3

The Gradient Flow Method

In the case of highly non-linear quantum field theories it can be useful to study the properties of flows in field space. The key idea is to see how the theory evolves as it becomes less local, by driving it to the stationary points of the action.

For the $SU(3)$ gauge field case, starting from a field $A(x)$ the characterizing equations are:

$$\begin{aligned}\dot{B}_\mu &= D_\mu G_{\mu\nu} \\ G_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu + [B_\mu, B_\nu]\end{aligned}\tag{3.1}$$

where the $B(t_f, x)$ field is the flowed version of the original field at flow-time t_f . This is imposed by fixing:

$$B_\mu|_{t_f=0} = A_\mu\tag{3.2}$$

The covariant derivative is extended to represent the derivative of the field B instead of A , this leaves the simple definition for the non-flowed field from the boundary condition. It is generalized straightforwardly as:

$$D_\mu = \partial_\mu + [B_\mu, \cdot]\tag{3.3}$$

In a discretized theory taking the derivatives is not a trivial procedure, but we can use instead the intuitive idea of the flow equations, that is to use the steepest descent method on the action of a field. We then introduce the flowed lattice gauge field $V_{t_f}(x, \mu)$ as the flowed version of a field $U(x, \mu)$. Our flow equation then becomes:

$$\dot{V}_{t_f}(x, \mu) = -g_0^2 [\partial_{x,\mu} S_G(V_{t_f})] V_{t_f}(x, \mu)\tag{3.4}$$

where S_G is the Wilson action, as defined in 2.18, generalized to flowed fields. The generalization is really straightforward, as it only requires to compute the Wilson loops on the flowed gauge field.

3.1 Perturbative Analysis of the Wilson Flow

An important thing to consider when applying the Wilson flow to a field is the renormalization of the observables: one has to check that expectation values of observables at non-zero

flow-time are renormalized quantities. Following the calculations performed in (CITATION NEEDED)luscher, we will consider the energy as our base observable.

First, we note that the flow equation is invariant under flow-time independent gauge transformations, this prevents a detailed study of the renormalization. So we consider a modified version of the flow equation by adding one term:

$$\dot{B}_\mu = D_\mu G_{\mu\nu} + \lambda D_\mu \partial_\nu B_\nu \quad (3.5)$$

the original case is obtained again by setting $\lambda = 0$ and considering:

$$B_\mu = \Lambda B_\mu|_{\lambda=0} \Lambda^{-1} + \Lambda \partial_\mu \Lambda^{-1} \quad (3.6)$$

and with $\Lambda(t, x)$ now being a flow-time dependent gauge transformation, set by:

$$\dot{\Lambda}_\mu = -\lambda \partial_\nu B_\nu \Lambda \quad \text{with} \quad \Lambda|_{t_f=0} = 1 \quad (3.7)$$

Chapter 4

Discretization Effects in Lattice Yang-Mills Theories and Scale Fixing

Part II

Implementation

Chapter 5

Designing a Lattice $SU(3)$ Yang-Mills Theory Code

One of the major focuses of this work has been a completely new implementation of a program to generate and analyze $SU(3)$ gauge fields. As it is common practice in Lattice QCD numerical implementations, the program is separated in two parts that are computationally intensive and one that is easier in that sense:

- *generation of gauge fields*: in this case it is done through a simple Metropolis algorithm using the standard Wilson action;
- *computation of observables*: this includes applying the gradient flow as well as computing the energy density and the topological charge at every flow-time;
- *computation of derived observables*: mainly post analysis, error analysis and model fits to data.

Here we will present the main features of the first two steps, which are the most interesting ones. The programming language of choice is `C++` because of its high efficiency, high abstraction capabilities (the code-base is highly object oriented) and for the easiness of the `MPI` integration. The analysis of data has been performed using `python` and in particular relying heavily on its standard data science packages such as `numpy` and `pandas`.

5.1 Generating Pure Gauge Fields

The task of generating lattice field configurations is extremely demanding in terms of computation requirements. The case of QCD is much more demanding than that of a pure Yang-mills theory, but overall the latter calculation is still challenging. The main, and perhaps overwhelmingly simple, reason for this problem is the dimensionality. Dealing with a discretized space-time

lattice, things tend to scale with powers of 4, a trivial example is cutting in half the lattice spacing keeping a fixed total volume: this requires 2^4 more points in the global lattice.

To get a better feeling of the algorithm we first have to look at what the basic object of the program is: the lattice. The number of double precision floating point numbers to be stored for a field configuration is given by

$$\underbrace{N^3}_{\text{spatial dimension}} \times \underbrace{N_t}_{\text{time dimension}} \times \underbrace{4}_{\text{links per site}} \times \underbrace{9}_{\text{SU(3) matrix size}} \times \underbrace{2}_{\text{real and imaginary part}} \quad (5.1)$$

this implies that, for example, if we choose $N = 48$, $N_t = 96$ the resulting configuration is 6115295232 *bytes* large, that is 6 *GBytes*. This limits greatly the possibility of simulating large systems

The basic element at each lattice site is a set of 4 different $SU(3)$ matrices, one for each dimension. These links are to be intended as the integral of the gauge field from one site to the adjacent one along each dimension. A more formal description has been given in chapter 3.

5.1.1 The Metropolis Algorithm

The main algorithm that has been used to generate an ensemble of gauge field configurations is the Metropolis Algorithm. It is a widely popular Markov Chain Monte Carlo Method to generate a sequence of random samples from a probability distribution. In our case the probability distribution is the gluon Wilson Action and our random samples are the gauge configurations themselves. A simplified version of the algorithm is described below:

Algorithm 1 Metropolis Algorithm

- 1: $configuration \leftarrow$ initial configuration
 - 2: **for** $i < MonteCarloCycles$ **do**
 - 3: $newConfiguration \leftarrow$ random move + $configuration$
 - 4: $\Delta S \leftarrow action(newConfiguration) - action(configuration)$
 - 5: **if** $\Delta S > random(0, 1)$ **then**
 - 6: $configuration \leftarrow newConfiguration$
-

The actual implementation of this algorithm on a lattice is however not as trivial as it seems, because we are dealing with a lattice gauge configuration, so we need to define what is a random move and what is the action difference for our specific case.

One configuration is saved, in the form of a simple binary containing all the data of the lattice, as an intermediate result every N_C Monte Carlo updates. We need this in order to apply the gradient flow afterwards to the configurations and compute the observables we want. The choice of N_C turned out to be crucial for the autocorrelation of some observables, in particular the topological charge.

5.1.2 Sampling the Configuration Space

To use Metropolis' algorithm we need to define what a random move is. Using the Wilson Action (REFLINK NEEDED), which is defined on plaquettes, an update on a single link variable can be seen as a small unitary transformation. In order to generate such transformation use three random $SU(2)$ matrices "close to unity". By this expression we mean that the real diagonal components are the dominant terms of the matrix. (CITATION NEEDED)

$$R_2 = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \quad S_2 = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} \quad T_2 = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} \quad (5.2)$$

The elements of the matrix are chosen at random by choosing four random numbers r_0, \mathbf{r} (a three component vector) between $(-\frac{1}{2}, \frac{1}{2})$. We then introduce a "spread parameter" ϵ that controls how much the off-diagonal terms will weight, so we scale our random variables by:

$$x_0 = \text{sign}(r_0)\sqrt{1 - \epsilon^2} \quad \mathbf{x} = \epsilon \frac{\mathbf{r}}{|\mathbf{r}|} \quad (5.3)$$

and we use these coefficients together with the generators of the $SU(2)$ group (the Pauli matrices) to build an element of the group:

$$U = x_0 \mathbb{1} + i\mathbf{x} \cdot \boldsymbol{\sigma} = \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \quad (5.4)$$

We then embed these $SU(2)$ matrices in three $SU(3)$ matrices by mapping them as:

$$R = \begin{pmatrix} r_{11} & r_{12} & 0 \\ r_{21} & r_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad S = \begin{pmatrix} s_{11} & 0 & s_{12} \\ 0 & 1 & 0 \\ s_{21} & 0 & s_{22} \end{pmatrix} \quad T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & t_{11} & t_{12} \\ 0 & t_{21} & t_{22} \end{pmatrix} \quad (5.5)$$

These three matrices are clearly members of $SU(3)$ and so is their product $X = RST$. We thus have defined a recipe for numerically generating random group transformations, the key element for our algorithm.

An additional element that we need to define is the action difference ΔS . On a single link $U_\mu(x)$ we apply a random transformation X and get $U'_\mu(x) = XU_\mu(x)$. The total change in the action only depends on those plaquettes that contain the considered link variable. In four dimensions there are 12 such elements:

$$\Delta S = S[U'_\mu(x)] - S[U_\mu(x)] = -\frac{\beta}{N} \Re \text{Tr}[U'_\mu(x) - U_\mu(x)]A \quad (5.6)$$

where A is the sum of the "staples" of the link U . They are the constant three sides of the plaquettes that contain U :

$$A = \sum_{\nu \neq \mu} [U_\nu(x + \mu)U_{-\mu}(x + \mu + \nu)U_{-\nu}(x + \nu) + U_{-\nu}(x + \mu)U_{-\mu}(x + \mu - \nu)U_\nu(x - \nu)] \quad (5.7)$$

add simpler form

FIG: STAPLES AND LINKS

5.1.3 Updates Strategies

There is now some arbitrariness in what is defined as an update. In this work we call an update the following procedure:

Algorithm 2 Metropolis Update

```

1: for  $x, \mu$  do
2:    $A \leftarrow \text{computeStaples}(x, \mu)$ 
3:   for  $i < N_H$  do
4:      $U_{\text{new}}(x, \mu) \leftarrow X \cdot U(x, \mu)$ 
5:      $\Delta S \leftarrow (U_{\text{new}}(x, \mu) - U(x, \mu)) \cdot A$ 
6:     if  $\text{realTrace}(\Delta S) > \text{random}(0, 1)$  then
7:        $U(x, \mu) \leftarrow U_{\text{new}}(x, \mu)$ 

```

These updates are the ones we consider when we refer to Monte Carlo cycles, autocorrelation times and so on. Note that each update includes a loop over all links on the lattice and that every link is "hit" N_H times before moving to the next one. This is done for computational efficiency, because computing the staples is the most expensive part of the algorithm, once A is computed for a link, the result is used to attempt multiple updates on the link. It can also be shown that if N_H is sufficiently large the algorithm becomes equivalent to the heatbath algorithm. (CITATION NEEDED)

The order in which the links in the lattice are visited is also arbitrary. The simplest way, the ordered one, was adopted. This however might have impacted the autocorrelation of the system, not allowing significant modification to the system as an update depends on the neighbors. An alternative choice is a checkerboard pattern, which has potential benefits to the autocorrelation time of the observables as well as on the parallelization scheme.see discussion .

5.1.4 Parallelization Scheme

Given the size of the lattice (from $V \approx 10^5$ to $V \approx 10^7$, the total number of lattice sites), it is necessary to split the computation of the updates on more than one processors. The most direct way is to divide the lattice into sub-blocks and have each process handle its portion of the field alone. FIG: blocks? However, because of the dependence of the action difference of a link on its neighbors, when updating a link on the edge of a sub-block information about another block is needed. Here is where the Message Passing Interface (MPI) comes in use. For this particular problem we decided to use a point to point communication scheme between the processors, the use of periodic boundary conditions allowed also for non-blocking communications to be used, in particular the collective geometry based scheme shown in (REFLINK NEEDED) has been implemented. FIG: sendrecv Note that the communication is relevant only for the computation of the staples, this is another argument in favor of performing multiple hits on a link at every update. It is clear to see that the algorithm can be affected by large communication overhead problems. If the sub-blocks are too small, then most of the time would be spent on sending and receiving links from the neighbors. This causes the execution time to depend not linearly on the

number of processors, instead the relation flattens at some value given by the communication overhead.

5.1.5 Summary of the Parameters

In total the algorithm need four parameters as inputs. Of these, only one defines the physics of the system, the others have to be set and optimized in order to improve the acceptance ratio of the metropolis test and the autocorrelation of the observables computed on the generated configurations.

Parameter	Description
β	Coupling parameter of the Wilson Action
ϵ	"Spread" of the random $SU(3)$ elements for the updates
N_C	Number of updates between one saved configuration (observable measurement)
N_H	Number of of "hits" per link at every update, with constant staple

Table 5.1: Parameters for the Monte Carlo generation of Yang-Mills gauge field configurations via the Metropolis Algorithm

5.2 Wilson Flow of Gauge Configurations

To study the flow-time dependence of the observables a numeric implementation of (REFLINK NEEDED) is needed. The main challenges here are to define the derivative of the action at a given lattice site and the numerical exponentiation of an $\mathfrak{su}(3)$ element, which returns an $SU(3)$ element.

5.2.1 The Action Derivative

Following the procedure in (CITATION NEEDED) we define the derivative of the action at a lattice site $U_\mu(x)$ to be:

$$\partial_\mu S[U_\mu(x)] = \frac{i}{2} \left(\Omega_\mu(x) - \frac{1}{3} \mathbb{1} \Im \text{Tr}[\Omega_\mu(x)] \right) \quad (5.8)$$

with

$$\Omega_\mu(x) = U_\mu(x) A_\mu(x) - A_\mu^\dagger(x) U_\mu^\dagger(x) \quad (5.9)$$

where $A_\mu(x)$ are the staples of the link. We can note that the derivative is always traceless anti-hermitean matrix **check...**, thus an element of $\mathfrak{su}(3)$ as expected.

5.2.2 Exponential of a $\mathfrak{su}(3)$ Element

Again following (CITATION NEEDED) we provide a numerical recipe for taking the exponential function of a traceless anti-hermitean 3×3 matrix. The key idea is to use the Cayley-Hamilton theorem, da fare...

5.2.3 Parallelization Scheme

Also this problem is very expensive in computational terms, so it is worth designing a parallelization scheme for it. The idea is still to split the lattice into sub-blocks and have each processor handle one of them. By looking at the (REFLINK NEEDED) we see that all operations can be defined on a lattice-wise scale and that they all depend on one previous state of the field, not on an intermediate one as was the case for the generation of gauge fields. This allows us to define a "shift" operation, that effectively creates an additional lattice that is the translation of the original along an axis. The huge advantage is that in order to perform this operation only one communication instance is needed (following the previous scheme in (REFLINK NEEDED)), but the message is now a whole shared cube between to processors. FIG: cubes being shared ADD ALGO... This has to be implemented for the action derivative lattice and the gauge field itself. The massive reduction in communication overhead improves the scaling of the algorithm greatly, making this part of the problem much more efficient.

5.3 Structure and Tools

The full code can be found on the web under the link [github...](#), where both the code for the generation and the flow of gauge fields is hosted. The technical documentation is found at [documentation...](#). As already mentioned the language of choice was C++, mainly because of the high-performance and abstraction level it provides. An object-oriented structure has been used, as can be seen in FIG: [scheme class](#). Notable tools that have been used that deserve a mention are MPI, nlhoman json (used for easy input parameters handling via json files), cmake (for building the project) and valgrind for function and memory profiling.

Chapter 6

Tests and Runs Description

6.1 Generated Ensembles

In order to study the scale parameter in the continuum limit, it has been necessary to choose a set of decreasing lattice spacings to then be able to take the continuum limit. The lattice spacing is linked to the coupling g and β by the known relation: **SOMMER E LOG** We chose 4 values of β that span lattice spacings from approximately 0.1 fm to 0.05 fm in approximately equal steps.

For the calculation to be consistent however, the total volume of the lattice should be kept constant, so the choice of the lattice spacings also determined the number of lattice sites per dimension, having $L = aN \approx \text{const.}$ The time dimension has been take to be twice as big as the space dimension.

Table **(REFLINK NEEDED)** summarizes the physical properties of the ensembles that were generated. For each value of β a statistical ensemble was needed. Ideally, one would take

β	a	$N^3 \times T$	aL [fm]
6.00	0.98	$24^3 \times 48$	2.2
6.10	0.98	$28^3 \times 56$	2.2
6.20	0.98	$32^3 \times 64$	2.2
6.45	0.98	$48^3 \times 96$	2.2

Table 6.1: Physical properties of the ensembles used for this work.

as many configurations as possible for each value and in principle one would have the same number of configurations for each of them. However it is clear from the discussion in chapter **(REFLINK NEEDED)** that the number of lattice sites affects computation times and the capability of storing the configurations dramatically. Moreover, as we will discuss in chapter **(CITATION NEEDED)** the autocorrelation time of the observables has a non-trivial, power-law or exponential, behavior with the lattice spacing, making the generation of the larger β ensembles even more time consuming.

autocorr for topc at different beta The following table summarizes the final values for the pa-

rameters of the Metropolis algorithm for the different ensembles. These values have been chosen

β	N_{conf}	N_{corr}	MC Steps	N_{hit}
6.0	1000	200	200000	30

Table 6.2: Physical properties of the ensembles used for this work.

after many tests, checks of the autocorrelation time and mainly in an empirical way. There are some parameters that are free in principle and no real reference study on their impact on the resulting ensemble. In the next section the trial and error approach that led to this decision will be briefly discussed.

6.2 Test Runs

Running some test calculations with a completely new code base is obviously necessary. First some benchmarks to check the expectation values of the observables, both on raw configurations as on flowed configurations. These benchmarks were made using 2 configurations generated with the CHROMA (CITATION NEEDED) code base from USQCD for zero flow-time observables and one configuration flowed using an extension of CHROMA called FlowOps (CITATION NEEDED), built on QDP++, that applies the Wilson flow to configurations. Both types of test, once all the parameters were made equal, gave results equal to machine precision with the ones generated with the new code base.

Checking the validity of expectation values of observables is a solid indication that the overall back-end of the new code-base has been implemented well, as the results are deterministic. Testing the Metropolis algorithm and assessing the quality of the generated ensembles is much harder, because it involves stochastic computations, hence no numerical check can be easily defined, one can only look at average properties of the ensemble. Unfortunately the generation of gauge field configuration as we saw in (REFLINK NEEDED) TABLE has 3 parameters that need to be set that do affect the statistical properties of the ensemble, mainly the autocorrelation.

6.2.1 Strong and Weak Scaling

First we can look at the scaling properties of the two sections of the code. We will distinguish the analysis in the two usual quantities used in High Performance Computing for parallel programs: strong and weak scaling.

Strong scaling is the performance of a program measured in execution time as a function of the number of processors used. One would obviously expect the relation to be ideally inverse, with execution time dropping as $1/N_{procs}$, however given the overhead caused by parallelization this is seldomly the case.

Weak scaling is the measure of the performance of a program as the size of the system increases but by keeping the total workload assigned to each processor constant. This measure is

important to assess the quality of the parallelization scheme as it gives insights on how much time is spent in communication as more inter-processor messages are being sent. **FIG: PLOT SCALING COMMENTI...**

6.2.2 Autocorrelation of Observables

The most important test has been the assessment of the autocorrelation time for different observables at various lattice spacings varying the parameters of the generation algorithm. The first test shows the integrated autocorrelation time for **asdasd** at fixed N_{corr} for different lattice spacings:

FIG: PLOT autocorr spacing COMMENTI...

Next we looked at a single lattice spacing and tried to vary the parameters N_{corr} and N_{hits} . The choice of the lattice spacing, motivated by the previous analysis, has been $\beta = 6.45$.

FIG: PLOT autocorr params COMMENTI...

6.3 Production Runs and Timing

All production runs were carried out on the High Performance Computing Center at Michigan State University (MSU), with the support of the Institute for Cyber-Enabled Research (iCER). Development was performed on local machines and on the small cluster SMAUG located at the Department of Physics of the University of Oslo (UiO) and some larger benchmarks were run on the Abel Computer Cluster also at UiO.

The final ensembles were generated using the parameters in table **(REFLINK NEEDED)**, shows the computing resources used for the generation of all four ensembles.

Part III

Data Analysis and Results

Chapter 7

Raw Observables

Chapter 8

Running Coupling and Scale Fixing

Part IV

Conclusion and Discussion

Chapter 9

Summary and Conclusion

Chapter 10

Future Developements

Part V

Appendices

Bibliography

- [1] Schroder Peskin. *An Introduction to Quantum Field Theory*. Cambridge University Press, 1989.
- [2] David J. Gross and Frank Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.*, 30:1343–1346, Jun 1973.
- [3] H. David Politzer. Reliable perturbative results for strong interactions? *Phys. Rev. Lett.*, 30:1346–1349, Jun 1973.
- [4] A Chaudhuri. A short course on relativistic heavy ion collisions. 07 2012.