

Stochastic methods in physics (unfinished, 2011-11-16)



Torquil Macdonald Sørensen

Centre of Mathematics for Applications
and
Department of Physics
University of Oslo

Nov 2011

Dissertation presented for the degree of
Philosophiae Doctor

Contents

| | | |
|----------|---|-----------|
| I | Introduction and summary of results | 7 |
| 1 | Introduction | 9 |
| 2 | Stochastic processes, path integrals and quantum physics | 13 |
| 2.1 | Introduction | 13 |
| 2.2 | Probability spaces | 14 |
| 2.3 | Random variables | 14 |
| 2.4 | Markov chains and stochastic processes | 15 |
| 2.5 | Random walks and Brownian motion | 16 |
| 2.6 | Brownian bridge | 18 |
| 2.7 | Path integrals | 18 |
| 2.8 | Path integrals in physics | 21 |
| 2.9 | Function spaces (in progress) | 21 |
| 2.10 | Equilibrium quantum statistical physics (in progress) | 22 |
| 2.11 | Metropolis/Hastings sampling algorithm | 23 |
| 3 | Levy process simulation by stochastic step functions | 25 |
| 3.1 | Introduction | 25 |
| 3.2 | Jump-diffusions | 27 |
| 3.3 | Simulation of a probability measure | 28 |
| 3.3.1 | Metropolis/Hastings (MH) | 28 |
| 3.3.2 | Stochastic step function (SF) | 29 |
| 3.3.3 | Simulation comparison | 31 |
| 3.4 | Probability space subdivision and adaptability | 31 |
| 3.4.1 | Adaptive Independent MH (AIMH) | 32 |
| 3.4.2 | Adaptive Independent SF (AISF) | 32 |
| 3.4.3 | Simulation test | 33 |
| 3.4.3.1 | Local MH | 34 |
| 3.4.3.2 | AIMH | 34 |
| 3.4.3.3 | Local SF | 34 |
| 3.4.3.4 | AISF | 34 |
| 3.5 | Jump-diffusion simulation | 34 |
| 3.6 | Application to NIG and CGMY infinite activity pure jump processes | 36 |
| 3.6.1 | Jump-diffusion approximation | 36 |
| 3.6.2 | Simulation of NIG | 37 |

| | | |
|----------|---|-----------|
| 3.6.3 | Simulation of CGMY | 39 |
| 3.7 | Discussion | 40 |
| 4 | Worm PIMC for Canonical Ensembles | 43 |
| 4.1 | Introduction | 43 |
| 4.2 | Physical system | 44 |
| 4.3 | Mathematical description | 45 |
| 4.4 | Discretisation | 49 |
| 4.5 | Pair action approximation | 51 |
| 4.6 | Sampling methods | 54 |
| 4.6.1 | Metropolis/Hastings | 55 |
| 4.6.2 | Brownian bridge sampling | 55 |
| 4.6.3 | Translation | 57 |
| 4.7 | Exchange effects | 57 |
| 5 | Simplicial gauge theory action | 59 |
| 5.1 | Overview | 59 |
| 5.2 | Motivation | 59 |
| 5.2.1 | General | 59 |
| 5.2.2 | Discretisation | 60 |
| 5.2.3 | Simplicial mesh | 60 |
| 5.3 | The continuous Yang-Mills action | 62 |
| 5.4 | Wilson action | 63 |
| 5.5 | Simplicial gauge theory action | 64 |
| 5.6 | Consistency | 66 |
| 5.7 | Numerical convergence tests | 67 |
| 5.8 | Discussion | 69 |
| 6 | Simplicial gauge theory and quantum simulation | 71 |
| 6.1 | Overview | 71 |
| 6.2 | Background | 71 |
| 6.3 | Parameters and continuum limit | 72 |
| 6.4 | Computer simulation | 74 |
| 6.5 | Discussion | 76 |
| 6.6 | Program details | 77 |
| 7 | Stochastic Gross-Pitaevskii equation | 79 |
| 7.1 | Introduction | 79 |
| 7.2 | Regarding the proof of Theorem 4.1 | 81 |
| 7.3 | Regarding the proof of Lemma 4.3 | 82 |
| 7.4 | Discussion | 82 |

| | |
|---|------------|
| II Papers | 87 |
| 8 Levy process simulation by stochastic step functions | 89 |
| 9 Simplicial gauge theory on spacetime | 111 |
| 10 Simplicial gauge theory and quantum simulation | 137 |
| 11 Comments on “The Stochastic Nonlinear Schrödinger Equation in H^1” | 157 |

Part I

Introduction and summary of results

Chapter 1

Introduction

The main topic of this thesis is the use of stochastic methods within physics. Stochastic methods have many important applications throughout science. Our observations of the universe has revealed the existence of randomness from the tiniest to the largest scales. One example of this is the randomness inherent in quantum physics which causes significant randomness on the smallest observable scales. On a larger scale, the phenomenon of Brownian motion was observed by considering the erratic movement of small particles suspended in a liquid. Human and society involves a myriad of phenomena that are too complicated to describe in accurate detail, which are described well probabilistically. On the largest scales, the structure of the non-isotropic temperature fluctuations in the cosmic background radiation seems to be caused by randomness occurring during an early period after the big bang. Interestingly, in this last example, the randomness is conventionally explained as being the resulting fingerprint of random quantum fluctuations in the early universe, so in a way the circle is closed with regard to randomness that appears on the smallest and largest scales.

Observed randomness can roughly be classified into three types:

1. Inherent randomness in the laws of nature, as in quantum physics. An example is the random and isotropically distributed, direction of a moving particle from the decay of some spin-less atomic nucleus.
2. Apparent random outcomes of chaotic physical processes, even though the governing laws are well understood. Our inability to detect or prepare initial conditions accurately enough make it impossible for us to predict the outcome with any confidence.
3. Simple lack of knowledge about a physical process. We are unable to understand the pattern in a series of outcomes, and must rely on a probabilistic description. This is the mechanism behind mathematical pseudo random number generators. These are simple algorithms that generate seemingly random sequences of numbers.

Stochastic methods involve random quantities. it does not matter if the randomness is inherent in the laws of nature or caused by human ignorance. The modern mathematical description of randomness in terms of probability spaces is always applicable for a well-defined probabilistic description. For this, it is necessary to have a well-defined set of states, events and probabilities for each possible event. As long as the problem is well-posed, any probabilistic prediction can be calculated by means of this theory.

Note the we will not be careful in distinguishing the concepts of stochastic methods, probabilistic methods and statistics in this discussion. In addition to descriptions of phenomena such as mentioned above, stochastic methods can be used as mathematical tools for calculation of quantities which are not random. Perhaps the most important application of stochastic methods is the application to multi-dimensional integration. This is usually called Monte Carlo integration, and is advantageous for calculation of many quantities in physics, particularly ones related to systems of many particles.

Random number are of course the key ingredient in the subject of statistics. This has to do with analysis of random systems. This might involve calculation of probabilistic properties of a particular random system, or the analysis of the validity of a statement made concerning a random system. Often one must try to deduce the correct probability distribution for the outcomes of a random system. This is done using statistical methods. In addition, confidence levels for such results can also be calculated. In physics, this often has to do with experimental errors. One must analyse their distribution in order to understand if a particular disagreement between experimental and theoretical results is likely the result of a “fluke” in the experiment, or an experimental or theoretical error or inaccuracy. But we will not discuss such matters much in this thesis.

This thesis involves the following topics that are related to stochastic methods and physics:

- Some mathematical and physical preliminaries. This involves the relationship between Brownian motion, Wiener path integrals and quantum statistical physics in thermal equilibrium. This is discussed in chapter 2. We find it unnecessary to go into too much detail. For example, on the topic of Wiener path integrals, which are central in the correspondence between quantum statistical mechanical systems in physics and the mathematical subjects of stochastic analysis and mathematical finance, we will only briefly summarise some basic facts that underlie the theory of Wiener integrals, which includes for example topology and measurable spaces. For a rigorous treatment, we will provide the necessary references to dedicated works.
- Simulating from a probability distribution and simulation of Levy processes in chapter 3. This involves a simulation algorithm based on stochastic step functions (SF), which we compare to the ordinary Metropolis/Hastings (MH) method.
- Canonical ensemble simulation of quantum many-particle bosonic systems in thermal equilibrium based on a path integral formulation, by means of a worm path integral Monte Carlo (WPIMC) method in chapter 4. This is a variant of the WPIMC algorithm for grand canonical ensemble simulations. We discuss the application to simulation of low-temperature Helium-4, related to Bose-Einstein condensates (BEC), superfluidity and the famous λ phase transition.
- The simplicial gauge theory (SGT) action, a discretisation of the Yang-Mills gauge theory action on a simplicial mesh in chapter 5. This alternative discretised action is inspired by the finite element method (FEM) for numerical solutions of partial differential equations (PDE). This involves theoretical estimates for consistency of the approximation, as well as numerical tests of convergence towards exact results are made for a few different gauge field configurations.

- Quantum Yang-Mills simulation on a simplicial mesh in chapter 6. This involves Monte Carlo simulations of the traditional type used in QCD, but using the SGT action. Measurements of the average action and differently sized Wilson loops are made.
- Finally, a discussion on the validity of a mathematical proof of solution existence for a stochastic Gross-Pitaevskii (GP) equation in chapter 7. This equation is related to the stochastic GP equation that describes a BEC interacting with a thermal reservoir.

I would like to thank...

Chapter 2

Stochastic processes, path integrals and quantum physics

2.1 Introduction

This chapter contains some mathematical preliminaries for the following chapters. It includes brief discussions of probability spaces and random variables, stochastic processes, measurable spaces, function spaces, Wiener path integrals and their use in the physics of many-particle systems and quantum field theory in equilibrium at some temperature.

In order to understand the concept of stochastic processes, we provide an informal introduction to the subject. We will only provide enough details to understand the basics of how randomness is treated mathematically, and what kind of mathematical objects are involved in the description [61].

We then go on to describe the connection to Wiener path integrals. In keeping with the usual physics terminology, a path integral is defined as an integral over a space of paths, or functions. That is, an integral over a subset of an infinite-dimensional space that can be represented as a space of functions. So it is not to be confused with an integral along a curve. This will also be used in this thesis, but no confusion should arise. Path integrals were introduced by the mathematician Norbert Wiener [58] in the mathematical study of Brownian motion, and independently in physics by Richard Feynman in the study of quantum mechanics [34].

The theories of quantum mechanics, quantum statistical mechanics and relativistic quantum field theory can be fundamentally formulated in terms of path integrals. In the latter case, convergence troubles arise due to the Minkowskian signature of spacetime. These integral no longer converge due to rapid exponential decrease of the integrand at large distances as is typical for Euclidean path integrals, but convergence must instead occur due to interference effects. However, rotations in a complex time domain are usually employed to obtain exponential suppression of the integrand. For a more comprehensive introduction to path integrals in physics, see [52, 16, 17].

Stochastic processes are also the basic mathematical tool for mathematical study of finance, and connections can be made between the stock market and the behaviour of a thermal quantum many-particle system. Of course, any system that includes randomness may be at least partly described by the use of stochastic analysis, and randomness is something that can occur in a

wide range of phenomena, especially complicated systems, as discussed in chapter 1. An interesting aspect of physics is that randomness occurs even for very simple systems, for example a single electron particle.

Regarding notation, it is customary to let the Wiener stochastic process denote a unit volatility Brownian motion process. However, we will not make any distinction, and use both Wiener process and Brownian motion independently of the volatility parameter value.

2.2 Probability spaces

We start by defining a set Ω of *states*, whose elements ω are in a one-to-one correspondence with all the possible states of the system. If we are discussing a mechanical system, this would correspond to all the different mechanical configurations this system might attain.

Next, we would like to define a family \mathcal{F} of possible *events*. Events are mathematically formulated as subsets of Ω . E.g. the event that the system is “in the state ω_1 ” corresponds to the event $\{\omega_1\}$. The event that the system is “in the states ω_1 or ω_2 , but in none of the states $\mathcal{U} \subset \Omega$ ”, corresponds to the set $\{\omega_1, \omega_2\} \cap (\mathcal{F} - \mathcal{U}) \in \mathcal{F}$. The event that “anything happens” corresponds to $\Omega \in \mathcal{F}$, and “nothing happens” corresponds to $\emptyset \in \mathcal{F}$. The object \mathcal{F} is called a σ -algebra. It will not be rigorously defined here, but its defining properties are in keeping with our usual notion of probabilistic events. Elements of \mathcal{F} are the *measurable subsets* of Ω .

Having defined events, we must define probabilities $P : \mathcal{F} \rightarrow [0, 1]$. A probability $P(\mathcal{U}) \in [0, 1]$ is defined for each event $\mathcal{U} \subset \mathcal{F}$. This formalism is defined to behave in keeping with the usual notions of probabilities, as regards for example expressions for the probabilities of compound and independent events. Consider two events $\mathcal{U}, \mathcal{V} \in \mathcal{F}$. In that case, the probability for the event “ \mathcal{U} or \mathcal{V} ” is given by the probability $P(\mathcal{U} \cup \mathcal{V})$. If \mathcal{U} and \mathcal{V} are disjoint, this becomes $P(\mathcal{U}) + P(\mathcal{V})$. The probability for the event “ \mathcal{U} and \mathcal{V} ” is given by $P(\mathcal{U} \cap \mathcal{V})$. In the case of disjoint events, this becomes $P(\emptyset) = 0$. The action of P upon such unions and intersections is always well-defined, due to the definition of \mathcal{F} . Being a probability measure, we have $P(\Omega) = 1$ and $P(\emptyset) = 0$, and it defines an integration measure on Ω .

2.3 Random variables

A *random variable* X is simply a function

$$X : \Omega \rightarrow \mathbb{R}^d, \tag{2.1}$$

for which the inverse image of measurable subsets in \mathbb{R}^d are measurable subsets of Ω . We are now assuming that a σ -algebra of measurable sets have been defined on \mathbb{R}^d , usually the Borel σ -algebra \mathcal{B} . This means that a family of possible events are defined as subsets of \mathbb{R}^d , and that the probability of each event can be determined by its inverse image in Ω and the probability measure on Ω . Note that the concept of a random variable can be generalised to other target spaces than \mathbb{R}^d . We write that $X(\omega)$ is a *realisation* of X , i.e. a possible value of the random variable X .

As ω traverses Ω , $X(\omega)$ will traverse some subset of \mathbb{R}^d . The image $X(\mathcal{U})$ of some subset

$\mathcal{U} \in \mathcal{F}$ might correspond to more probable events than others. Thus the probability measure P on Ω will induce a corresponding probability measure μ_X on \mathbb{R}^d , by defining

$$\mu_X : \mathcal{B} \ni U \mapsto P(X^{-1}(U)) \in [0, 1]. \quad (2.2)$$

Note that we will usually work with a corresponding probability density ρ_X and the usual Lebesgue measure dx on \mathbb{R}^d (for lack of a better notation). For some integrable function f on \mathbb{R}^d , the integral using the measure induced from P is defined as

$$\int_{\mathbb{R}^d} f \mu_X = \int_{\mathbb{R}^d} f(x) \rho_X(x) dx, \quad (2.3)$$

where the density

$$\rho_X(x) := \frac{d\mu_X}{dx} \quad (2.4)$$

is the Radon-Nikodym derivative of μ_X with respect to the ordinary Lebesgue measure dx .

For a random variable X , averages and variances are defined as

$$\text{Avg}[X] = \int_{\Omega} X(\omega) dP(\omega) = \int_{\mathbb{R}^d} x \rho_X(x) dx, \quad (2.5)$$

$$\text{Var}[X] = \int_{\Omega} (X(\omega) - \text{Avg}[X])^2 dP(\omega) = \int_{\mathbb{R}^d} (x - \text{Avg}[X])^2 \rho_X(x) dx. \quad (2.6)$$

Therefore, these quantities can be calculated either by integration over the state space Ω , or the target space \mathbb{R}^d .

2.4 Markov chains and stochastic processes

We are now ready to consider time-dependent random variables. First, consider a countable family, a *chain*, of \mathbb{R}^d -valued random variables $\{X_i\}$. For a given $\omega \in \Omega$, the sequence $\{X_i(\omega)\} \subset \mathbb{R}^d$ is a realisation of the chain. We would like to mention an important class of such chains, namely *Markov chains*. They are characterised by the fact that they have independent increments. The random variable X_{i+1} only depends on X_i , not on X_{i-1}, X_{i-2} , etc.

We can generalise from such a stochastic chain to a stochastic process, which depends continuously on time. It is therefore an uncountable family $\{X_t\}$ of random variables, enumerated by continuous time. Also here we can define a corresponding *Markov process*, for which the increments are independent of the previous history of the process.

In this case, we can speak of realisation paths $t \mapsto X_t(\omega)$ for each $\omega \in \Omega$. These are also called *sample paths*. Stochastic processes are usually denoted X_t , omitting the brackets. When the target space for the sample paths of a process X_t is some space M , we write that X is a stochastic process in M .

2.5 Random walks and Brownian motion

Consider an infinite cubic grid in \mathbb{R}^d with grid constant $h > 0$, and a homogeneous time discretisation with a time interval $\tau > 0$. On this grid, a point particle is moving discontinuously a distance h between neighbouring vertices in a random direction between each discrete time value, starting at the origin at $t = 0$. Since the particle moves randomly, its position at a given time t is a random variable governed by some induced probability measure on the grid. This probability measure can be calculated from these assumptions, but we are only interested in a probability density in the continuum limit. The position at time $t + \tau$ only depends on the previous position. Thus the particle position defines a Markov chain. Note that the average velocity of the particle is $v = h/\tau$.

It turns out that the Markov process continuum limit $h, \tau \rightarrow 0$ of this Markov chain is only well-defined if we keep h^2/τ constant. Instead of a Markov chain, this leads us to a stochastic process, depending continuously on time. Its properties are characterised only by the limiting value of h^2/τ and the number of dimensions d . We encode the first value into a parameter called the *volatility*, $\sigma := \sqrt{h^2/\tau d}$. This stochastic process is called a *Brownian process*, or *Wiener process*, and is denoted B_t . Its sample paths are continuous. Note that the continuum velocity of the particle is no longer well-defined, since $h/\tau \rightarrow \infty$ because σ is kept finite in this limit.

In the continuum limit, it can be shown that the continuum probability density $\rho_t(x)$ of the particle position satisfies the heat equation,

$$\partial_t \rho_t(x) = \frac{\sigma^2}{2} \Delta \rho_t, \quad \rho_t(x) = \delta(x), \quad (2.7)$$

where δ is the Dirac delta distribution. The use of a distributional initial condition simply means that we required that the limit $\lim_{t \rightarrow 0^+} \rho_t = \delta$ within a suitable space of distributions. Thus the continuum probability density is given by the Gaussian distribution with variance $\sigma^2 t$,

$$\rho(x, t) = \frac{1}{(2\pi\sigma^2 t)^{d/2}} \exp\left(-\frac{x^2}{2\sigma^2 t}\right). \quad (2.8)$$

It is easy to determine the transition probability density $k(x_2, t_2; x_1, t_1)$ for this process. This function is the probability density of B_{t_2} , given that $B_{t_1} = x_1$, for $t_2 > t_1 \geq 0$. Due to the space and time translation invariance of the laws that govern the defining discrete random walker, it must be a translationally invariant function of space and time,

$$k(x_2, t_2; x_1, t_1) = k(\Delta x, \Delta t; 0, 0), \quad (2.9)$$

where $\Delta x := x_2 - x_1$ and $\Delta t := t_2 - t_1$. Therefore,

$$k(x_2, t_2; x_1, t_1) = \frac{1}{(2\pi\sigma^2 \Delta t)^{d/2}} \exp\left(-\frac{\Delta x^2}{2\sigma^2 \Delta t}\right), \quad t_2 > t_1 \geq 0. \quad (2.10)$$

We can easily understand that for $t_3 > t_2 > t_1 \geq 0$, we have the Chapman-Kolmogorov

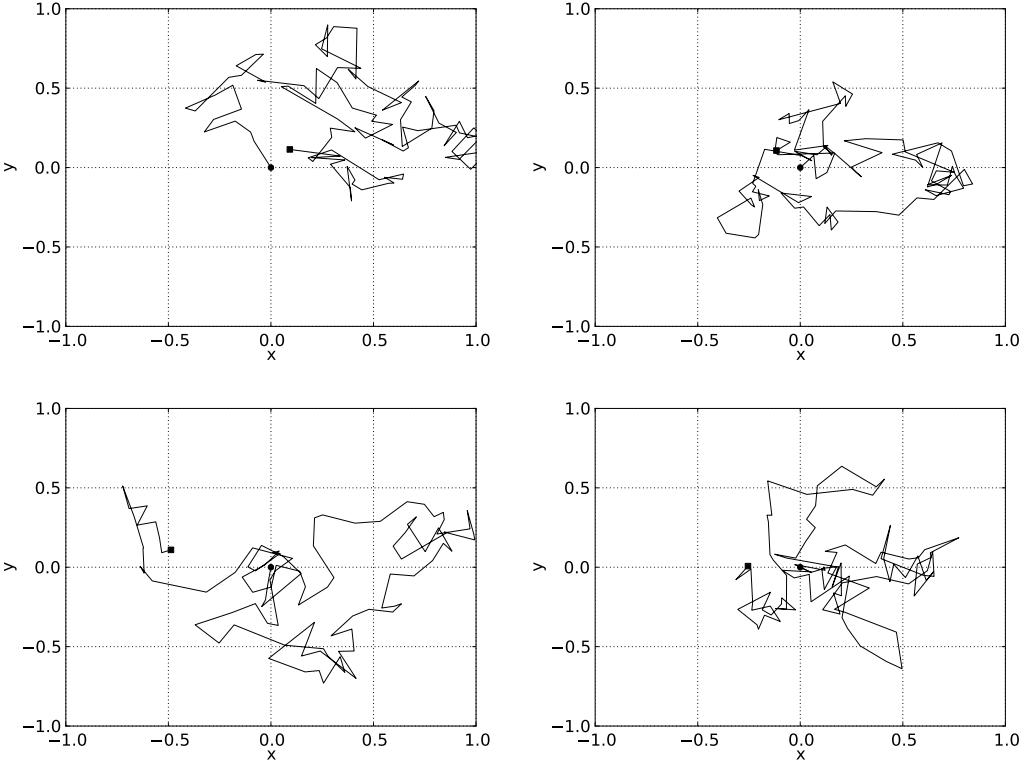


Figure 2.1: Four sample paths of discretised unit volatility Brownian motion in two dimensions on the time interval $[0, 1]$, starting at the origin. The discrete time step is $\Delta t = 0.01$. The starting point is indicated by a small disk, and the ending point by a small square.

equation

$$k(x_3, t_3; x_1, t_1) = \int k(x_3, t_3; x_2, t_2)k(x_2, t_2; x_1, t_1)d^d x_2. \quad (2.11)$$

It has a simple interpretation: the probability density for a transition $(x_1, t_1) \rightarrow (x_3, t_3)$ can be expressed as the sum over the probability densities for the probabilistically exclusive transition events $(x_1, t_1) \rightarrow (x_2, t_2) \rightarrow (x_3, t_3)$ for all possible intermediate positions x_2 at the intermediate time t_2 .

Using these facts, it follows that the average, variance and covariance of B_t are given by

$$\text{Avg}[B_t] = 0, \quad \text{Var}[B_t] = \sigma^2 t, \quad (2.12)$$

$$\text{Avg}[B_t B_s] = \frac{\sigma^2}{2} (t + s - |t - s|). \quad (2.13)$$

In figure 2.1 we show a sample path of discretised Brownian motion in two dimensions.

2.6 Brownian bridge

It is possible to define a stochastic process BB_t between two fixed endpoints, called a Brownian bridge. Given fixed endpoints $x_a, x_b \in \mathbb{R}^d$, the Chapman-Kolmogorov equation 2.11 determines the induced probability density on \mathbb{R}^d of BB_t as

$$p_t(x) = k(x_b, t_b; x, t)k(x, t; x_a, t_a), \quad t_b > t > t_a \geq 0. \quad (2.14)$$

This probability density is a Gaussian distribution with average and variance

$$\text{Avg}[BB_t] = x_a + \frac{t - t_a}{t_b - t_a}(x_b - x_a), \quad (2.15)$$

$$\text{Var}[BB_t] = \sigma^2 \frac{(t - t_a)(t_b - t)}{t_b - t_a}. \quad (2.16)$$

The average value is simply a linear path between x_a and x_b , while the variance is maximal at the middle time value, and approaches zero towards both ends. Given two endpoints, any intermediate position of the Brownian bridge can be sampled from a Gaussian distribution with the above data. In particular, it is possible to define a recursive bisection algorithm that samples a midpoint value, and goes on to sample the midpoint values on each of the smaller intervals. This bisection algorithm is used in PIMC simulations as described in chapter 4.

The process BB_t can be explicitly expressed in terms of B_t using a stochastic Ito integral, however, this is an unnecessary complication for our context [61]. Figure 2.2 includes four sample paths from a Brownian bridge process.

2.7 Path integrals

Our goal is now to relate the Brownian process on \mathbb{R}^d to path integrals over a space of \mathbb{R}^d -valued paths. We have defined a stochastic process as a continuum of random variables. Consider the Brownian/Wiener process for a time interval $[0, \beta]$. Its sample paths are continuous, and the continuum of mappings $B_t : \Omega \rightarrow \mathbb{R}^d$ can be “unified” to define a mapping $B : \Omega \rightarrow C$, where C is the space of continuous paths $[0, \beta] \rightarrow \mathbb{R}^d$, starting at the origin,

$$\mathcal{C} := \{q \in C([0, \beta], \mathbb{R}^d) | q(0) = 0\}. \quad (2.17)$$

If we can define a σ -algebra \mathcal{F}_C of measurable sets on \mathcal{C} , we can use the properties of the Brownian process to also define a probability measure on \mathcal{C} . This probability measure will assign a volume to each measurable subset of continuous functions. Therefore, we will be able to integrate over this space. This is the theory of Wiener path integrals, and the Wiener path integral measure[58, 52, 16].

We avoid a rigorous treatment, and go on to define our notation for a subset of \mathcal{C} that will

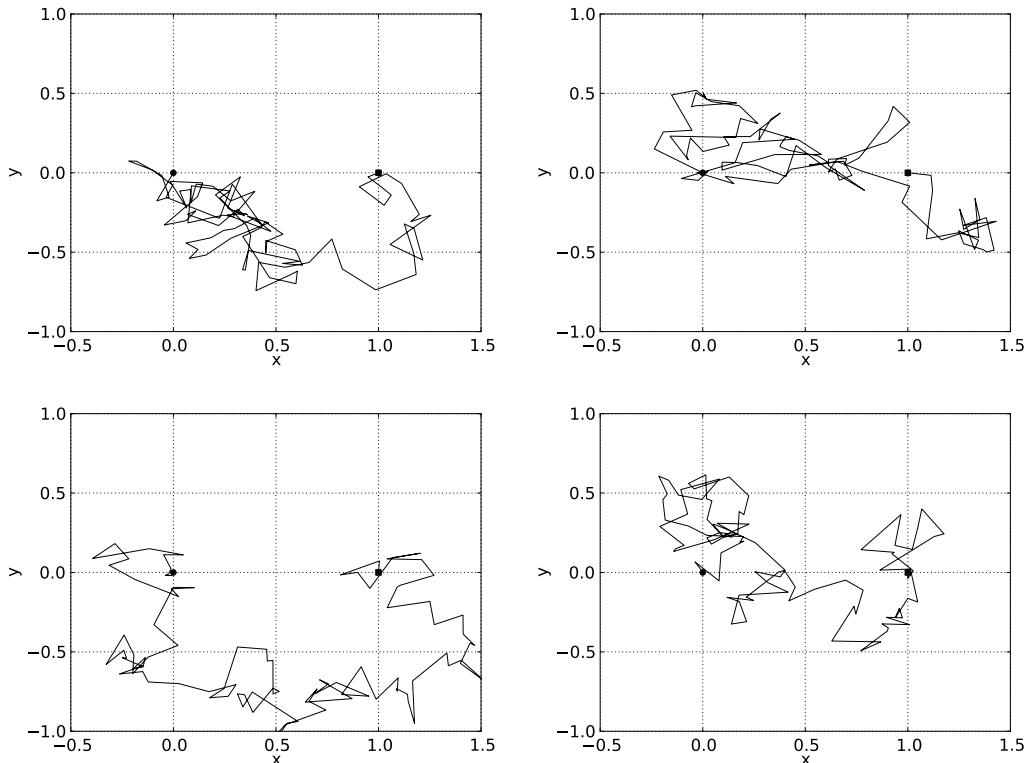


Figure 2.2: Four sample paths of a discretised unit volatility Brownian bridge in two dimensions on the time interval $[0, 1]$, starting at $(0, 0)$ and ending at $(1, 0)$. The discrete time step is $\Delta t = 0.01$. The starting point is indicated by a small disk, and the ending point by a small square.

be used as a kind of building block,

$$\mathcal{U}_1(t, D) := \{q \in \mathcal{C} | q(t) \in D\}. \quad (2.18)$$

This subset contains all paths in \mathcal{C} that intersect the measurable subset $D \subset \mathbb{R}^d$ at time $t \in (0, \beta]$. This subset is one element of the σ -algebra $\mathcal{F}_{\mathcal{C}}$, which also contains intersections of such sets,

$$\mathcal{U}_M(t_1, \dots, t_M; D_1, \dots, D_M) := \mathcal{U}_{t_1}(D_1) \cap \dots \cap \mathcal{U}_{t_M}(D_M), \quad (2.19)$$

containing all paths in \mathcal{C} that intersect a series of subsets $\{D_i\}$ at the times $0 < t_1 < \dots < t_M \leq \beta$. Such sets are used to define the σ -algebra $\mathcal{F}_{\mathcal{C}}$, and therefore also to define integration over \mathcal{C} . We are only missing one essential component, the definition of the volume for each measurable subset of \mathcal{C} , which is defined by the Wiener measure. This is deduced by constructing an induced probability measure on \mathcal{C} from the known probability measures of B_t on \mathbb{R}^d . The resulting volume for a measurable set of the above form, in the particular case of $t_m := m\Delta t$ for convenience, is

$$\mu_W(\mathcal{U}_M) := \int \dots \int \exp \left(- \sum_{m=0}^{M-1} \frac{\Delta x_m^2}{2\sigma^2 \Delta t} \right) \prod_{m=1}^M \frac{d^d x_m}{(2\pi\sigma^2 \Delta t)^{d/2}}, \quad (2.20)$$

where each x_m (corresponding to the position $q(t_m)$ of a path) is integrated over D_m , $t_0 := 0$ and $\Delta x_m := x_{m+1} - x_m$. This measure is also a probability measure on \mathcal{C} , since $\mu_W(\mathcal{C}) = 1$. Intuitively, one can see that subsets mainly consisting of paths for which the square differences Δx_m^2 are very large will have a small volume.

One can also go through the same procedure, based on the Brownian bridge process instead of the unrestricted Brownian process. One then ends up with a restricted Wiener path integral measure

$$\mu_W^r(\mathcal{U}_{M-1}) := \frac{\exp(\frac{x_M^2}{2\sigma^2 t_M})}{(2\pi\sigma^2 t_M)^{-d/2}} \int \dots \int \exp \left(- \sum_{m=0}^{M-1} \frac{\Delta x_m^2}{2\sigma^2 \Delta t} \right) \prod_{m=1}^{M-1} \frac{d^d x_m}{(2\pi\sigma^2 \Delta t)^{d/2}}, \quad (2.21)$$

where x_m is the fixed endpoint at time t_M , and is not integrated. This is also a probability density on the space of paths \mathcal{C}^r accordingly restricted at both endpoints, since $\mu_W^r(\mathcal{C}^r) = 1$ due to the prefactor.

In physics, it is common to consider the limit $M \rightarrow \infty$ and $\Delta t \rightarrow 0$, and to write the unrestricted Wiener density as the formal expression

$$d\mu_W(q) = \exp \left(- \int_0^\beta \frac{1}{2\sigma^2} \dot{q}(t)^2 dt \right) \prod_{t=0}^\beta \frac{d^d q(t)}{(2\pi\sigma^2 dt)^{d/2}}. \quad (2.22)$$

This is a misuse of notation for several reasons. The paths $q \in \mathcal{C}$ are not assumed to be differentiable, the meaning of an infinitesimal quantity in a denominator is undefined, and it contains an uncountable product of differential measures. However, the meaning of such expressions is

usually clear from the context, and calculations always involve a division of the time domain into a finite number of pieces, just as we have done above in the definition of the Wiener measures. We will also misuse notation in this way later in the thesis. When we do, the integrand will always include an exponential damping factor, so it is really just a rewriting of a Wiener measure.

2.8 Path integrals in physics

Path integrals are used in physics to calculate quantum amplitudes both in quantum mechanics of finite and infinite number of degrees of freedom. The method of variational calculus is a powerful tool in many areas of physics. Consider e.g. classical mechanics. The equations of motion of a particle moving in a potential can be deduced by minimising a quantity called the action S . The action is a mapping that assigns a real number to each possible path that the system can take. It is given as an integral over time of a quantity called the Lagrangian function. Given an expression for the Lagrangian in terms of the particle path and its time-derivative, one can deduce the equations of motion by minimising the action integral. This leads to the Euler-Lagrange equations.

The early workers with quantum mechanics discovered that the variational formalism was intimately related to quantum mechanics. P. A. M. Dirac speculated that for a transition of a quantum system between two states A and B , e.g. a two possible positions of a point particle, one could assign to each path from A to B a “quantum amplitude” $\exp(iS)$, where S depends on the particular path taken.

Later, this lead R. P. Feynman to consider an infinite sum over all such paths. It was known that the most fundamental property of quantum mechanics is the superposition principle, i.e. the effect of interference between different possible quantum events. Feynman considered an integral over all paths going from A to B , using Dirac’s amplitude $\exp(iS)$ as the integrand. This leads to highly oscillatory path integrals that are difficult to define rigorously. We will consider Wiener integrals, since they are used in physics in the context of quantum statistical mechanics. A quantum many-particle system, or even a quantum field theory of an uncountably infinite number of degrees of freedom, is governed by a Wiener path integral, using paths that are cyclic in the time dimension.

Such path integrals form the basis of many Monte Carlo simulation techniques, and we will employ this in the context of Monte Carlo simulation of a many-particle system of bosonic particles in chapter 4, and for lattice simulations of Yang-Mills theory discretised on a simplicial lattice in chapters 6.

2.9 Function spaces (in progress)

We define the space $L^p(\mathbb{R}^d, \mathbb{C})$ with the norm $\|\cdot\|_{L^p(\mathbb{R}^d, \mathbb{C})}$, of all mappings $f : \mathbb{R}^d \rightarrow \mathbb{C}$ for which

$$\|f\|_{L^p(\mathbb{R}^d, \mathbb{C})} := \left(\int_{\mathbb{R}^d} \|f(x)\|_{\mathbb{C}}^p d^d x \right)^{1/p} < \infty, \quad (2.23)$$

where the norm of a complex number is defined as usual as

$$\|z\|_{\mathbb{C}} := \sqrt{c\bar{c}}.$$

Note that we identify mappings for which differ only on sets of measure zero. The space $L^2(\mathbb{R}^d, \mathbb{C})$ is a Hilbert space, with the inner product

$$\langle f, g \rangle := \operatorname{Re} \int f(x) \overline{g(x)} d^d x,$$

Without going deeply into the concept of weak derivatives, we also define the space $H^1(\mathbb{R}^d, \mathbb{C}) \subset L^2(\mathbb{R}^d, \mathbb{C})$ for which for the first order weak derivatives ∂f exist and are elements of $L^2(\mathbb{R}^d, \mathbb{C}^d)$. Essentially, g is the weak derivative of f if the following equation involving partial integration holds for all smooth Schwartz functions φ ,

$$\int_{\mathbb{R}^d} g \phi d^d x = - \int_{\mathbb{R}^d} f \partial \phi d^d x.$$

This means that g acts like the derivative of f inside integrals, even though the ordinary derivative ∂f is not well-defined. But we still denote the weak derivative of f by ∂f .

All this is similarly defined for mappings into the target spaces \mathbb{R}^n and \mathbb{C}^n for any $n \in \mathbb{N}$. When the target space is omitted in the notation, it is \mathbb{C} .

2.10 Equilibrium quantum statistical physics (in progress)

A quantum physical system in thermodynamic equilibrium at some temperature T can be described by means of the density operator

$$\rho(\beta) := \exp(-\beta H), \tag{2.24}$$

where $\beta := 1/T$ and H is the Hamiltonian operator, depending on the canonical position operators Q , and accompanying generalised momentum operators P . Real-time dynamical quantum physics without thermal fluctuations can be related to this statistical description of a quantum system with thermal fluctuations, by means of a complex rotation of the time variable onto the imaginary axis. This comes about because the unitary time-propagator of real-time quantum mechanics is given by

$$U(t) := \exp(iHt), \tag{2.25}$$

where $t \in \mathbb{R}$ is time. The correspondence is obtained by rotating t into the complex plane, obtaining $t = i\beta$. Therefore, β is often described as an imaginary time variable, and it is often said that quantum statistical systems are propagating along the imaginary time axis. Similarly, the word time will often be misused as a term indicating the coordinate along the inverse temperature dimension, even though it does not really describe anything that has to do with physical time-evolution in a thermodynamic system.

Note that our units of temperature and energy coincide, so Boltzmann's constant $k_B = 1$.

Our unit for both temperature and energy will be the Kelvin (K). For time-dependent quantum mechanics without a thermal influence, the density operator is normalised, due to its probability interpretation. In this case, (2.24) is only normalised for the case of a free theory. However, since we will mostly be concerned about relative probabilities, this is not a problem. The common numerical simulation techniques are independent of the exact normalisation. It would perhaps be more appropriate to describe ρ as an imaginary time propagator. However, we will stick with the usual terminology, so as not to cause unnecessary confusion. For a derivation of (2.24) from basic principle, see [45].

Note that the density matrix satisfies the composition property

$$\rho(\beta_1 + \beta_2) = \rho(\beta_1)\rho(\beta_2), \quad (2.26)$$

as a simple consequence of the trivial identity $[H, H] = 0$.

The density operator completely describes both the quantum and thermal randomness of the system. In a basis of energy eigenstates $\{|i\rangle\}$ of H , it is diagonal and given by

$$\rho_{ij}(\beta) := \langle i | \rho(\beta) | j \rangle = \delta_{ij} \exp(-\beta E_i),$$

where E_i is the energy eigenvalue of state $|i\rangle$. Within the statistical ensemble of many such systems, the probability of observing a randomly chosen system to be in the state $|i\rangle$ is given by the Gibbs weight

$$w_i := \exp(-\beta E_i)/Z(\beta), \quad (2.27)$$

where the partition function responsible for the normalisation is

$$Z(\beta) := \text{tr}[\rho(\beta)] = \sum_i \exp(-\beta E_i). \quad (2.28)$$

So we see that the density operator formalism reproduces the well-known Gibbs distribution for the distribution of states of a statistical system in equilibrium at a finite temperature. Any observable thermodynamic quantity of such a system can be calculated by means of the partition function. As an example, the total energy of the system is given by

$$E(\beta) = -\frac{d}{d\beta} \log Z(\beta). \quad (2.29)$$

2.11 Metropolis/Hastings sampling algorithm

Consider the real line \mathbb{R} , one which a probability density π is defined. We wish to draw numbers from \mathbb{R} that are distributed according to π . Actually, we go one step further and say that our goal is to define a Markov chain on \mathbb{R} that has π as its invariant distribution. This is done by defining a computer algorithm that will create realisations of such a Markov chain.

The Metropolis/Hastings (MH) algorithm works as follows [48, 41]. Since we are generating a Markov chain, the MH algorithm is inductive, generating each value only based on the knowledge of the previous one. Assume that the i 'th value x_i has been selected. Then, draw a

value y from \mathbb{R} using some probability density $t(x_i, y)$. Since \mathbb{R} is unbounded, we cannot use a uniform distribution here. However, in the compact case, this would be possible. The MH algorithm decrees that we should calculate the probability

$$p := \min \left(1, \frac{\pi(y) t(x_i, y)}{\pi(x_i) t(y, x_i)} \right) \in [0, 1], \quad (2.30)$$

and then demand

$$\begin{aligned} \text{Prob}(x_{i+1} = y) &= p \\ \text{Prob}(x_{i+1} = x_i) &= 1 - p. \end{aligned} \quad (2.31)$$

In other words, y is accepted as the subsequent value x_{i+1} with probability p . Otherwise, the preceding value x_i is repeated. Any initial value $x_0 \in \mathbb{R}$ may be used.

Given some quite weak demands on the properties of t which we will not go into, it is possible to prove that this algorithm satisfies the properties of detailed balance and ergodicity, which implies that π is the invariant distribution of the resulting Markov chain. Note that even if π is not normalised, that is $\pi(\mathbb{R}) \neq 1$, this algorithm will nevertheless do the sensible thing, and simulate from a normalised version $\tilde{\pi} := \pi / \pi(\mathbb{R})$. In many cases, we will have a symmetric transition probability density t , which simplifies the probability calculation (2.30).

Chapter 3

Levy process simulation by stochastic step functions

3.1 Introduction

We study a Monte Carlo algorithm for simulation of probability distributions, based on stochastic step functions (SF), and compare to the traditional Metropolis/Hastings (MH) method [48, 41]. Unlike MH, the SF approach can produce an uncorrelated Markov chain.

Algorithms for simulation from arbitrary probability distributions have many areas of application throughout just about all science that involves numerical analysis. Typically, repeated measurements might show that some quantity involved in a natural phenomenon appears to be random and distributed according to some probability distribution. For numerical simulations of that phenomenon, the ability to draw random numbers from an arbitrary distribution is paramount.

As mentioned above, in this chapter we study such an algorithm based on step functions (SF), and compare to an MH algorithm. As an application, we choose the simulation of Levy processes, for which simulation of uncorrelated jumps are essential.

Levy processes are a type of stochastic process whose paths can behave erratically like a Brownian motion, as well as include discontinuities, i.e. jumps. Apart from the usefulness of the probability distribution simulation method itself, simulation of Levy processes are useful in many areas, since many observed phenomena in nature and human society display continuous erratic motion similar to Brownian motion, while also having random jumps at random times. They have applications within the subjects of laser cooling, animal movement patterns, earth crust movements during earthquakes, and financial markets. In some cases, the Levy process might describe position variables (e.g. the ocean crust elevation during an earthquake), and in other cases momentum variables (e.g. laser cooling).

Time-dependent probability distributions occur many places in physics, for example in the case of the Schrödinger equation. Such probability distributions can be described as distributions of a stochastic process as a function of time, although in this case the paths of the stochastic process do not represent physical particle paths or particle momentum. In the case of Levy processes, this is the case for the distribution generated by relativistic Klein-Gordon wave equation. See [36] for details.

To increase efficiency of the SF method, and to decrease correlations in the MH method, we introduce adaptive hybrid algorithms which employ uncorrelated draws from an adaptive discrete distribution defined on a space of subdivisions of the Levy measure space.

The nonzero correlations in MH simulations result in heavy tails for the Levy process distribution at any fixed time. This problem is eliminated in the step function approach. In each case of the Gaussian, NIG and CGMY processes, we compare the distribution at $t = 1$ with exact results and note the superiority of the step function approach.

In our case, will will employ a jump-diffusion approximation when we apply our method to infinite activity Levy processes. Jump-diffusion processes can be considered as a sum of three independent component processes. A deterministic drift, a Brownian motion, and lastly a finite activity jump process. The jump process is completely described by a Levy measure ν on \mathbb{R} . The average jump rate, or intensity, is given by the total weight $\lambda := \nu(\mathbb{R})$, and the distribution of jump values follows the Levy probability measure $\nu_1 := \nu/\lambda$.

By definition, the Brownian motion has independent increments, and this is also the case for subsequent jumps in the jump process. In computer simulations, violation of these properties will give incorrect results. Simulation of the Brownian motion is easy, since well known algorithms exist for producing uncorrelated draws from a normal distribution, using a good pseudo-random number generator (PRNG).

On the other hand, the simulation of the jump process requires more care. In principle we can obtain uncorrelated draws from ν_1 by inverting its cumulative distribution function. However, this might not be feasible to do for a given ν_1 . In this case, the MH algorithm might come to the rescue. It can easily produce a Markov chain with values distributed according to ν_1 . However, subsequent values will be correlated, as described below.

In this paper we propose an algorithm to produce uncorrelated jumps along each path, without generating such a multitude of Markov chains. This method is based on SF, which will be defined below. As opposed to MH, it is not an accept/reject algorithm. Therefore it is able to generate an uncorrelated chain of values distributed according to any given probability measure. We will test this algorithm in jump-diffusion computer simulations and compare with MH. We apply the simulation techniques to a Gaussian process, for which the exact distribution at $t = 1$ is known. Convergence towards this exact result is studied.

We also consider adaptive variants of these algorithms. For these, we subdivide \mathbb{R} into appropriate regions, and generate a discrete distribution that allows us to efficiently draw among these regions in an uncorrelated way. As the simulation progresses, this discrete distribution is adaptively improved. The calculation of the jump rate λ is also adaptively improved.

As an interesting application of these simulation techniques, we also look at infinite activity pure jump processes, which are also a Levy process subclass. Here, motion occurs in the form of an infinitude of discontinuous jumps. Some of these processes can be approximated by jump-diffusion by substituting the smallest jumps for an appropriate Brownian motion [6]. The examples we focus on are the NIG and CGMY processes.

In section 3.2 we review the elementary facts about jump-diffusion processes. In section 3.3 we describe the relevant simulation methods, and discuss their pros and cons, as well as provide results from numerical experiments. In section 3.4 we describe a probability space subdivision method and the accompanying adaptive properties of the algorithms, and study the efficiency and correlation strengths of different algorithms by simulation from a Gaussian

distribution. We then proceed in section 3.5 to perform simulations of jump-diffusion processes. We notice how the MH correlations adversely affect the distribution of the simulated process, and compare the SF and MH algorithms for simulation of a process with a Gaussian Levy measure. Lastly, in section 3.6, we review the jump-diffusion approximation of infinite activity pure jump processes, and apply our simulations techniques on the infinite activity NIG and CGMY processes which are relevant in mathematical finance, in order to compare algorithms in these interesting cases.

3.2 Jump-diffusions

First we review some elementary facts about real-valued jump-diffusion Levy processes on a time interval $[0, T]$. Such a process can be expressed as a sum of three simple components

$$L_t = \mu t + B_t + J_t. \quad (3.1)$$

The first part is a deterministic drift with rate μ , B_t is Brownian motion, and J_t is a compound Poisson process. The latter describes completely the discontinuities (jumps) of the paths of L_t , by means of a *Levy measure* ν on \mathbb{R} . We will several times abuse notation by writing ν both for the Levy measure and its density. Firstly, this measure determines the jump intensity (jump rate)

$$\lambda := \nu(\mathbb{R}) < \infty, \quad (3.2)$$

of L_t . Secondly, the corresponding *Levy probability measure*

$$\nu_1 := \nu/\lambda, \quad (3.3)$$

determines the distribution of jump sizes. Note that we do not have $E[L_t] = \mu t$ in general, although this is satisfied in our examples.

Now, J_t can be expressed as

$$J_t = \sum_{j=1}^{N_t} V_j, \quad (3.4)$$

where N_t is a Poisson process of intensity λ , and the jumps $\{V_j\}$ are independent random variables distributed according to the Levy probability measure. We do not simulate the Poisson process N_t directly. Instead, for each path we draw the total number of jumps individually from an exponential distribution with average λT . Then we randomly distribute those jumps on $[0, T]$. We will choose $T = 1$ for our simulations. The lengths of consecutive time intervals between jumps will be independent and exponentially distributed, and give the correct jump intensity. This is described in [24].

Note that for a general Levy process λ is in general not finite, because the Levy measure might diverge too strongly towards the origin. However, the following condition always holds,

$$\int_{\mathbb{R}} \min(1, s^2) \nu(ds) < \infty. \quad (3.5)$$

which restricts the strength on the divergence of ν at the origin.

The difficulty in simulation is to draw independent jump values from ν_1 . We will generate a Markov chain with ν_1 as its invariant measure. However, it is essential for correct simulation that jumps along each path are uncorrelated. Note that existence of correlations between jumps in *different* paths will not ruin the convergence in distribution, but only slow it down.

For a Markov chain $\{X_i\}$, we define the sequential correlation as

$$c := E[(X_{i+1} - \bar{X})(X_i - \bar{X})]/\sigma^2, \quad (3.6)$$

where \bar{X} is the Markov chain average and σ^2 is its standard deviation,

$$\bar{X} := E[X], \quad \sigma^2 := E[X^2 - \bar{X}^2].$$

3.3 Simulation of a probability measure

To simulate the jumps, we must draw independent values from ν_1 . In cases where ν_1 is complicated, it is common to use the Metropolis/Hastings (MH) algorithm [48, 41]. This method generates a Markov chain with ν_1 as its invariant density. Unfortunately, the Markov chain often has large correlations between successive values. Successive values in such a chain cannot be used to represent jumps within a single jump-diffusion path.

It is possible to reduce this problem by several methods. One is to skip a number of terms in the Markov chain to reduce correlation. This results in a loss of efficiency. Another method is to generate multiple independent Markov chains. Each jump-diffusion path then uses values from different Markov chains. This will lead to a correct pathwise simulation, and therefore correct convergence in distribution. The correlations will in this case only slow down the convergence.

We will look at two methods, MH and one based on stochastic step functions (SF). Both rely on a transition probability distribution ρ to determine upcoming values from the previous one. We say that we are dealing with a *local* algorithm if ρ is localized around the current value. An *independent* algorithm results if ρ is independent of the current position.

Use of an independent MH algorithm can reduce correlations dramatically if one is able to find a ρ that approximates ν_1 . The downside is that efficiency will suffer if ρ is not computationally simple. We will focus on generation of low correlation Markov chains in order to get by using only one chain for the Levy process jumps.

3.3.1 Metropolis/Hastings (MH)

The well-known MH algorithm generates a realisation $\{x_i\}$ of a Markov chain distributed according to ν_1 . Its most popular incarnation is local, where the transition probability density

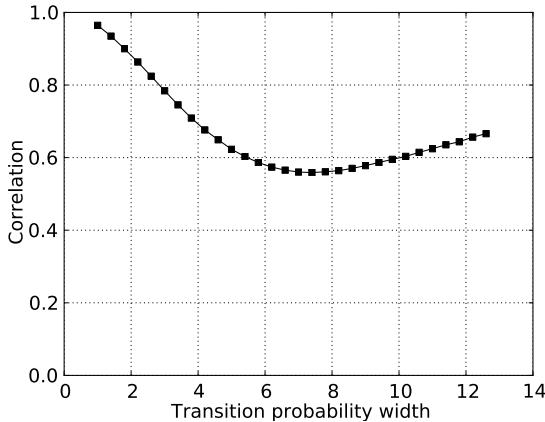


Figure 3.1: Correlation c versus transition probability width w for the local MH algorithm, in a simulation of the unit normal distribution using a uniform centred transition probability distribution. Correlation is minimal around $w \approx 7$, where $c \approx 0.55$.

$\rho(x_i, x_{i+1})$ involved in each transition $x_i \mapsto x_{i+1}$ is localized around the value x_i , and its width is adjusted to give an average acceptance rate somewhere around $1/2$. The resulting correlation between successive values can be classified into two types:

- *Rejection correlation*: Since both the local and independent algorithms are based on acceptance/rejectance, correlations are introduced by repeated values due to rejections. With an acceptance rate around $1/2$, repeated values will often occur.
- *Transition correlation*: In order for the local algorithm to be efficient, its transition probability measure $\rho(x_i, x_{i+1})$ must in many cases be quite strongly localized; otherwise too small an acceptance rate will result. Thus the difference between subsequent variates of the Markov chain will tend to be small.

To reduce transition correlation, the width of the transition distribution can be increased. But this typically leads to a reduced acceptance ratio, which gives an increased rejection correlation, and vice versa.

As a numerical example, consider a unit variance normal distribution, simulated with a simple localized uniform transition probability measure. The correlation defined in (3.6) as a function of transition probability width is displayed in Figure 3.1. As expected, the correlation has a minimum. Towards smaller widths the transition correlation increases, and towards larger widths the rejection correlation increases. Thus there is a lower bound on the amount of correlation for this algorithm, and it seems that the local MH algorithm is therefore not suited for our purposes. We will from now on focus on the independent MH algorithm.

3.3.2 Stochastic step function (SF)

Next, we propose a simulation algorithm based on stochastic step functions, that can be adjusted to completely eliminate correlations. This possibility of vanishing correlations is an attractive property that allows it to be used as a reference algorithm. It can also be adjusted to run more efficiently, with nonzero correlation.

Let us give a pathwise definition of the SF process on $[0, \infty)$ for some measure density ν on $\Omega \subset \mathbb{R}$, not necessarily normalized. Consider a sequence $\{\tau_i\}_{i=0}^\infty \subset (0, \infty)$ of *resting times* and corresponding *jump times* $\{t_i\}_{i=0}^\infty$, defined recursively by

$$t_0 = 0, \quad t_{i+1} := t_i + \tau_i.$$

In addition, let $\{s_i\}_{i=0}^\infty \subset \Omega$ be a Markov chain with initial value $s_0 = 0$ and transition probability density $\rho(s_i, s_{i+1})$. Assume that $\rho(s_i, \cdot)$ is absolutely continuous with respect to the Lebesgue measure on \mathbb{R} , and homogeneous, i.e. it can be expressed as $\rho(s_{i+1} - s_i)$.

From this data, we are now ready to define our stochastic step function process X_t , by defining its paths as the piecewise constant *cadlag* functions (right-continuous with left limits) of the form

$$X_t = \sum_{i=1}^{\infty} s_i \chi_{I_i}(t), \quad I_i := [t_i, t_{i+1}),$$

where χ_I is the indicator function for the interval I , and where the resting times are given by $\tau_i := \nu(s_i)$.

We see that the paths of X_t linger for a some time at each of its attained positions, with a resting time equal to the local value of the density ν . Consider the graph of such a path on $[0, T]$ for large T compared to $\sup \nu$. As T increases, the relative path length within a given subset $A \subset \Omega$ on the vertical axis converges towards $\nu(A)/\nu(\Omega) =: \nu_1(A)$. When sampling this path on a uniform time-grid, the obtained values will therefore be distributed according to the probability density ν_1 .

As in the case of local MH, we get transition probability correlations if ρ is localized. In this algorithm, however, there is no accept/reject step, and therefore no rejection correlation. As an example, let us choose ρ to be the uniform probability distribution on Ω . As noted above, if we sample the step function path generated above on a uniform time grid, we get a chain of values distributed according to ν_1 . If we make sure that the time discretisation interval size Δt is larger than $\sup \nu$, repeated values will never occur and there are no correlations.

Note that as in MH algorithms, we do not need to know scale factor relating ν_1 and ν . If $\sup \nu$ is initially unknown, we can simply start with any estimate, and improve it adaptively as the algorithm runs.

It is easy to see that the amount of computer time spent between each discrete time value is proportional to $\sup \nu / \nu(\Omega)$ in its correlation-free mode. If this ratio is large, the algorithm will be inefficient.

One can make the time discretisation finer to increase the rate of variate generation, but this introduces correlations since values for which ν is large will be repeated with complete certainty. This is similar to the Metropolis algorithm, where values of large ν are repeated (although not with complete certainty). The SF algorithm can be made less deterministic in this case by letting the resting times be random variables distributed according to an exponential distribution with mean $\nu(x)$, where x is the current position.

Since we are concerned with minimizing correlations in the context of jump-diffusion simulations, we will use independent algorithms, where $\rho(x_i, x_{i+1})$ is independent of x_i .

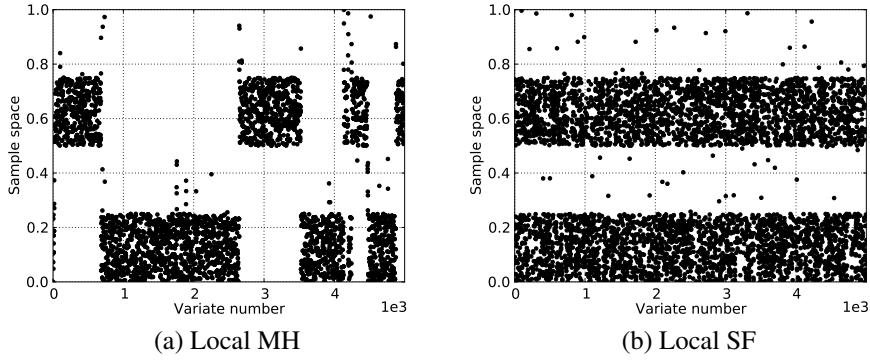


Figure 3.2: Markov chain realisations using local MH and local SF. The SF algorithm mixes much better than MH.

3.3.3 Simulation comparison

Let us illustrate the advantages of the local SF algorithm over local MH by considering an example with a probability distribution with several modes on a sample space $\Omega = [0, 1]$. Our goal is to simulate values from a probability distribution proportional to the following unnormalised density with two strong modes,

$$\nu(x) = \begin{cases} 1 & , x \in [0, 0.25) \cup [0.5, 0.75) \\ 0.01 & , \text{otherwise.} \end{cases} \quad (3.7)$$

For both MH and SF, we used a localized uniform transition probability density ρ of width 1/2. For MH, this gave an acceptance rate of approximatively 0.55. The SF Markov chain realisation was obtained by sampling the stochastic step function using a lattice spacing slightly larger than $\sup \nu$ in order to avoid any repeating values.

The results are shown in Figure 3.2. One sees immediately that the MH algorithm has the potential of getting stuck inside a mode for long periods. This is caused by rejection correlation as found in accept/reject algorithms such as MH. On the other hand, the SF algorithm has no such problems since it lacks such correlations. This serves to illustrate a problem that often can occur with MH. The SF algorithm is a clear favourite in this case if mixing is important for the application of these Markov chains.

3.4 Probability space subdivision and adaptability

The Levy probability measure from which we will simulate in the context of jump-diffusions will often be defined on the unbounded probability space $\Omega = \mathbb{R}$. This presents no difficulty for the MH algorithm. SF algorithms on the other hand need to impose upper and lower cutoffs on Ω , in order to avoid step functions that diverge towards infinity. For simplicity, we impose such cutoffs on both algorithms in our examples. Alternatively, one could perform a topology-changing coordinate transformation on Ω to obtain a compact space, however we will not do this in our examples. From now on we therefore assume $\Omega \subset \mathbb{R}$ to be a bounded interval which we choose symmetrically about the origin. We will only deal with symmetrical Levy measures

in our examples.

In order to reduce correlations in the MH algorithms, and increase efficiency in the SF algorithms, we define a finite disjoint subdivision $\{U_i\}$ of Ω , where $\cup_i U_i = \Omega$. We construct a discrete distribution $\tilde{\nu}$ on the finite set $\{U_i\}$. For each $U_i \subset \Omega$, the value $\tilde{\nu}(U_i)$ must approximate $\nu(U_i)$, i.e. the Levy measure weight of U_i . The simulation algorithm starts with a preliminary estimate of $\tilde{\nu}$ which is adaptively improved throughout the simulation, as explained below in the two cases of MH and SF. This is reminiscent of variance reduction schemes used in Monte Carlo integration, such as stratification and the VEGAS algorithm [46]. We now describe in more details how this is done in each case.

3.4.1 Adaptive Independent MH (AIMH)

As explained above, we will use an independent MH algorithm. The independent transition probability ρ is defined as follows. First, use the discrete unnormalised probability measure $\tilde{\nu}$ to draw a subset U_i . Then a random position within this subdomain is proposed and the value of ν at this position is calculated. Thereafter follows the usual MH accept/reject step.

The initial draw of U_i is done without introducing correlations, using e.g. an efficient algorithm which is independent of the normalization of $\tilde{\nu}$ [54]. The registered values of ν are accumulated, and used periodically in the simulation to improve $\tilde{\nu}$. Essentially, a separate Monte Carlo simulation is being performed within each subdomain U_i to improve the discrete distribution $\tilde{\nu}$, while the algorithm proceeds to generate new draws from ν .

Since subdomains U of Ω are drawn by means of $\tilde{\nu}$ without introducing correlations, the amount of correlation generated by the algorithm as a whole is reduced. Transition correlation is completely eliminated since the algorithm is independent. It is impossible to completely eliminate rejection correlation in an MH algorithm unless ρ is identical to ν_1 . However, since ν is well approximated on each subdomain (as long as the subdivision is sufficiently fine), these will be small, and the acceptance rate will be high. As $\tilde{\nu}$ adapts, the acceptance rate rises and correlation decreases. The subdivision discretisation implies a lower bound for the amount of correlation. But this is still a big improvement on a basic local MH algorithm for the cases we consider. Note that the subdivision cannot be made arbitrarily fine, since the discrete algorithm will then become inefficient.

3.4.2 Adaptive Independent SF (AISF)

It is possible to improve the efficiency of the SF algorithm by a similar construction, without introducing correlations. In this case, the SF algorithm proceeds as follows. First draw a subdomain U_i using $\tilde{\nu}$. Draw a random position x within this subdomain, and record the value $\nu(x)$ for later use to improve $\tilde{\nu}$. For each subset U_i we keep an estimate of $\sup_i \nu$ which is updated for each sample. Each subset is also accompanied by a local time variable t_i . For each position x within U_i , we increase t_i by the resting time $\nu(x)$. We choose new positions independently within U_i until t_i exceeds the current estimate of $\sup_i \nu$.

When the above loop exists, we have determined our new position and we subtract $\sup_i \nu$ from t_i in anticipation of the next time we enter U_i . In a sense we are using a different time discretisation in each subset U_i of Ω .

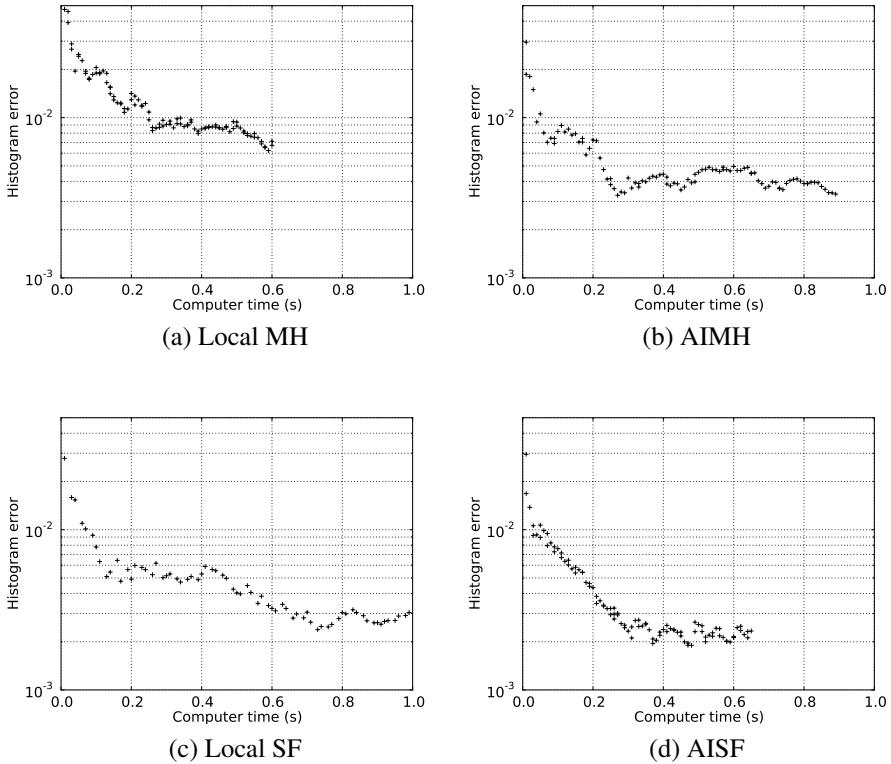


Figure 3.3: Histogram error versus computer time, for simulation from normal distribution.

This modified algorithm avoids spending a lot of time in areas of low probability for two reasons. First, the low probability subdomains will rarely be entered when drawing from the discrete distribution $\tilde{\nu}$. Secondly, the amount of time necessary to exit the loop in those regions is much smaller, since the local sup ν is small. By exploiting the subdivision of Ω , and using an efficient algorithm for drawing from discrete distributions, we improve the efficiency of the SF algorithm drastically in nontrivial cases.

Both $\tilde{\nu}$ and the estimates sup ν are adaptively improved throughout the simulation.

3.4.3 Simulation test

To check the rate of convergence of the different methods, we simulated from a Gaussian distribution with unit variance. The rate of variate generation, correlation, and the convergence towards the known exact result were analysed.

A histogram of the simulated values is gathered, and compared with a histogram calculated from the Gaussian distribution. We define the histogram error using an L^∞ measure, i.e. as the absolute value of the maximum deviation from the exact result. The plots show the histogram error as a function of simulation running time. The simulations were run on one core of an Intel T4400 laptop processor, but only the relative efficiencies matter here. Results are shown in Figure 3.3 and are discussed below.

3.4.3.1 Local MH

Since a low correlation will be important for our later use on Levy processes, we selected the transition probability width 0.75, giving the minimal correlation of approximatively 0.55, according to Figure 3.1. The variate generation rate was approximately $3.4 \cdot 10^6 / s$ in this case.

3.4.3.2 AIMH

As expected, as great improvement of the correlation was obtained compared to local MH. Since the algorithm is more complicated, the variate generation rate decreased to $2.3 \cdot 10^6 / s$, or roughly 60% of local MH. However, the correlation was reduced to around 0.03. This dramatic decrease results in a faster histogram convergence in terms of computer time. The lower correlation nature of this algorithm will be noticeably beneficial when simulating Levy processes.

3.4.3.3 Local SF

We adjusted the basic SF algorithm parameters to give zero correlation, and used a uniform transition probability density on $[-5, 5]$. Thus we are running it quite inefficiently as regards variate generation speed, which turned out to be around $1.4 \cdot 10^6 / s$. Despite this, the histogram converges faster than local MH, due to the lack of correlation.

3.4.3.4 AISF

As expected, when including the domain subdivision and adaptive algorithm, the SF algorithm efficiency increased (in fact doubled) with a variate generation rate of $3.1 \cdot 10^6 / s$. The amount of correlation here is still zero, which leads to the fastest histogram convergence. So this is a great improvement at no cost other than increased algorithm complexity. It is perfectly suited for simulating consecutive jumps in jump-diffusion process paths.

3.5 Jump-diffusion simulation

As previously mentioned, existence of correlations among jumps within the same jump-diffusion path is disastrous. Let us check this by using an local MH algorithm to simulate an exactly solvable stochastic process [47] and compare distributions at $t = 1$. It is defined as

$$X_t := \sigma W_t + J_t,$$

$$J_t = \sum_{i=1}^{N_t} Y_i, \quad Y_i \sim N(\mu, \delta^2), \quad \mu, \delta \in \mathbb{R},$$

where N_t is a Poisson process of intensity λ . We choose $\sigma = 1$, $\mu = 0$, $\delta = 1$, and $\lambda = 10$ in order to get an appreciable number of jumps within the time domain $[0, 1]$.

The exact PDF of this process for $t > 0$ is [24, Eq.(4.12)]

$$p_t(x) = \exp(-\lambda t) \sum_{k=0}^{\infty} \frac{(\lambda t)^k \exp(-\frac{x^2}{2(t+k)})}{k! \sqrt{2\pi(t+k)}}.$$

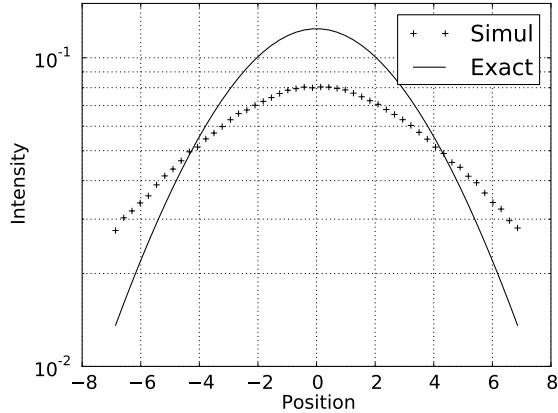


Figure 3.4: Local MH simulation of distribution at $t = 1$ of Gaussian process, and exact result. Due to jump correlations, the simulated histogram has heavy tails.

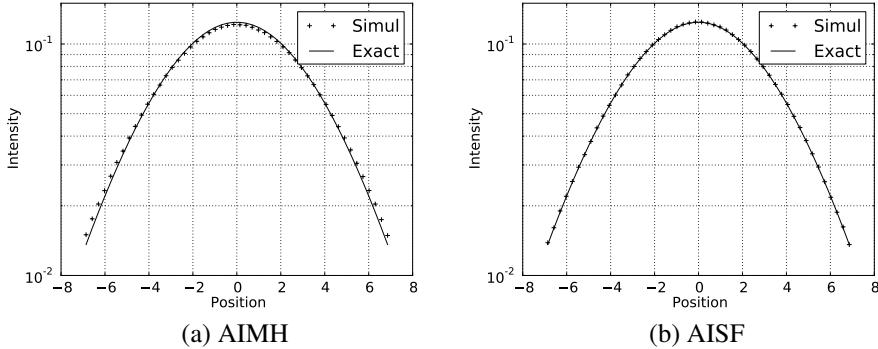


Figure 3.5: Distributions of Gaussian process at $t = 1$ for AIMH and AISF simulations. The continuous curve represents the exact result. Notice the slight heavy tails in the AIMH case, due to nonzero jump correlations.

For local MH simulations, the positive correlations between subsequent jumps leads to heavy distribution tails. Results confirming this are shown in Figure 3.4. The simulation of the jumps used a localized uniform transition probability of width 4, which resulted in a MH acceptance rate of around 0.56.

We now turn to our serious attempts at jump-diffusion simulation. We have simulated the same process using AIMH and AISF algorithms. In this case we use the already known value of $\lambda = 10$ in the simulations, so the adaptability consists solely of the calculation of the discrete subdivision distribution, and in the AISF case also of the local subdivision supremum calculations. As opposed local MH, in this case the distribution at $t = 1$ approaches the exact curve, as seen in Figure 3.5. Note that despite our use of upper/lower cutoffs in the implementation of the discrete subdivision, the tails of the local MH simulation are still somewhat heavy.

We use the same L^∞ measure of convergence detailed above. Convergence results are shown in Figure 3.6. The AISF outperforms the AIMH algorithm due to its lower correlation which leads to faster and more accurate convergence.

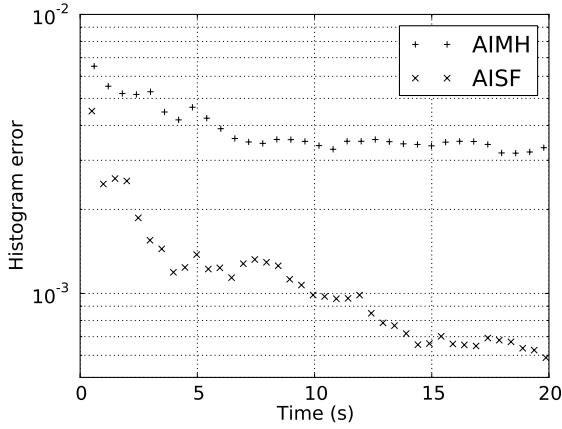


Figure 3.6: Error in distribution at $t = 1$ for the Gaussian process, versus computer simulation time. AISF gives the fastest convergence. The AIMH heavy tails cause a lower bound on the error.

3.6 Application to NIG and CGMY infinite activity pure jump processes

We present two applications within the realm of infinite activity pure jump Levy processes, namely NIG and CGMY. To this end, we employ the jump diffusion approximation of these processes [6].

3.6.1 Jump-diffusion approximation

For an infinite activity Levy process, the intensity λ is undefined, since its defining integral (3.2) diverges. However, by virtue of (3.5), it is possible to approximate the infinitude of small jumps by a Brownian motion process. Consider an infinite activity pure jump Levy process L_t . For $\epsilon > 0$, define the following subsets of Ω ,

$$\begin{aligned} A^{\epsilon,-} &:= \{|x| < \epsilon : x \in \mathbb{R}\} \\ A^{\epsilon,+} &:= \{|x| > \epsilon : x \in \mathbb{R}\}. \end{aligned}$$

These represent small and large jumps, respectively.

We now define the unique Levy measures $\nu^{\epsilon,-}$ and $\nu^{\epsilon,+}$ on these subdomains as follows. For any ν -measurable $U \subset \mathbb{R} - \{0\}$, define

$$\begin{aligned} \nu^{\epsilon,-}(U) &:= \nu(U \cap A^{\epsilon,-}) \\ \nu^{\epsilon,+}(U) &:= \nu(U \cap A^{\epsilon,+}). \end{aligned}$$

Each of these Levy measures in turn define its own Levy process L^ϵ_- and L^ϵ_+ of infinite and finite activity, respectively. We have the unique process decomposition

$$L_t = L_t^{\epsilon,-} + L_t^{\epsilon,+}. \tag{3.8}$$

It follows from (3.5) that the intensity of the large jump component process,

$$\lambda^{\epsilon,+} := \nu^{\epsilon,+}(\mathbb{R})$$

is well-defined, and so $L_t^{\epsilon,+}$ is a compound Poisson process. For small ϵ , the small jump process $L_t^{\epsilon,-}$ can be often be approximated by the following Brownian motion [6]

$$L_t^{\epsilon,-} \approx \sigma(\epsilon)W_t, \quad \sigma(\epsilon)^2 := \text{Var}[L_1^{\epsilon,-}], \quad (3.9)$$

where W_t is a Wiener process. The variance must exist for the approximation to be well-defined. A sufficient condition is [6]

$$\lim_{\epsilon \rightarrow 0} \frac{\sigma(\epsilon)}{\epsilon} = \infty. \quad (3.10)$$

Thus we have the approximation

$$L_t \approx \sigma(\epsilon)W_t + L_t^{\epsilon,+}. \quad (3.11)$$

Any additional nonvanishing deterministic drift component of L_t would appear trivially on the right hand side of this equation.

In these cases, the intensity λ is not known beforehand, and also depends on our choice of ϵ . Since our algorithms already calculate $\tilde{\nu}$, and the individual volumes of each subset in the subdivision of Ω is known, it is easy to use this to update an estimate for λ concurrently.

3.6.2 Simulation of NIG

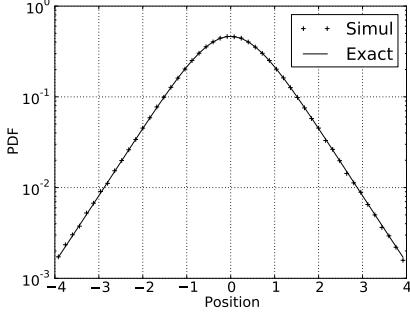
We can now check the quality of the jump-diffusion approximation in conjunction with our AIMH and AISF algorithms. Since the density of NIG is known, and a direct algorithm [24, Alg.6.12] exists for simulating from its distribution at any time, we have ample material to compare our results. The numerical results for the density at $t = 1$ are collected using logarithmic plots in Figure 3.7, in order to emphasise the distributional tail behaviour. In terms of the parametrisation used in [24], we use $\sigma = 1$, $\theta = 0$, $\kappa = 1/2$. Figure 3.7 contains the results from the direct simulation algorithm, as well as the results of the jump-diffusion approximation using AIMH and AISF, where we have used the small jump cutoff value $\epsilon = 0.005$.

The volatility σ of the Brownian component in the jump-diffusion approximation, defined by (3.9), is calculated analytically using an approximate expression for the Levy measure at small $|x|$. The Levy jump density is

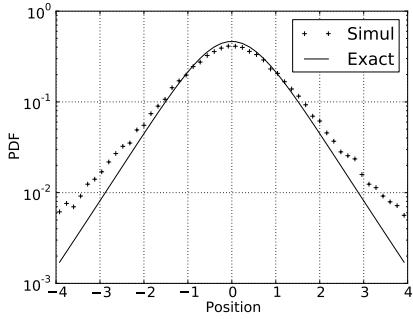
$$\nu(x) = \frac{\alpha\delta}{\pi|x|} e^{\beta x} K_1(\alpha|x|), \quad (3.12)$$

where K_1 is the irregular modified cylindrical Bessel function of first order. As an asymptotic approximation of (3.12) at small $|x|$, we use

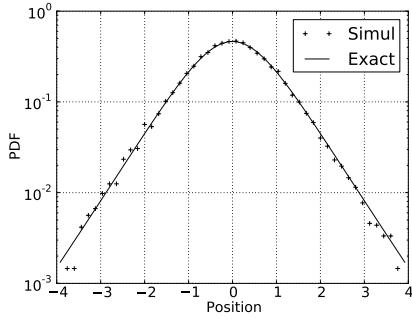
$$e^{\beta x} \approx 1, \quad K_1(\alpha|x|) \approx \frac{1}{\alpha|x|}.$$



(a) Cont/Tankov



(b) AIMH



(c) AISF

Figure 3.7: Simulated and exact NIG PDFs at $t = 1$. The AIMH algorithm has heavy tails.

Inserting the chosen parameter values, and using (3.9) which gives σ^2 as the second moment of the Levy jump density on $[-\epsilon, \epsilon]$, we get

$$\sigma \approx \sqrt{2\epsilon\delta/\pi} \approx 0.067.$$

Notice that this expression satisfies the Asmussen/Rosinski condition (3.10). Owing to this small value, the Brownian part of the Levy process has a negligible influence on the distribution tails at $t = 1$, compared to the contributions from larger jumps.

The direct algorithm from [24] works best, as expected. We are doing this to compare AIMH and AISF, for a known case with an exact solution. The AIMH and AISF algorithms are general and can be used on any Levy process that allows a jump-diffusion approximation. This case gives us an indication of how trustworthy these methods are when applied to more exotic Levy processes for which we do not have an exact result or a simplified simulation methods as in this case.

Most sources of errors are common to both algorithms. These are related to inaccuracy in the calculation of the σ and λ , the latter being related to the cutoff imposed on the jump domain Ω . This cutoff will cause an inaccuracy in λ since the total weight of Ω will not be calculated. Improved subdivision schemes of Ω are possible, e.g. employing coordinate transformations that transform Ω into a compact domain. We have not done this, since it unimportant for the matters we are considering.

We note that AIMH will never be completely correlation-free, and will therefore tend to produce heavy tails as is obvious in Figure 3.7. No such problem exists for AISF. The cutoff on Ω will naturally lead to weak tails, as seen in the plot. This can be remedied by enlarging the

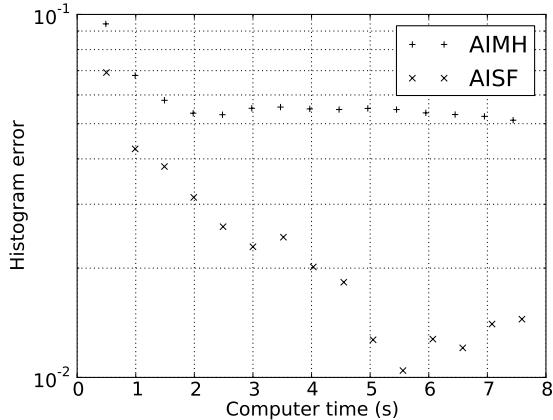


Figure 3.8: NIG histogram error at $t = 1$, versus computer time. The AISF clearly converges faster. AIMH reaches a steady state in the convergence plot due to the error caused by the nonvanishing jump correlations.

cutoff value, and/or treating large jumps differently.

The measurement data for convergence analysis is plotted in Figure 3.8, from which one readily sees that the AISF algorithm converges faster and more exactly.

3.6.3 Simulation of CGMY

As a second example of a pure jump infinite activity process, we turn to CGMY [14]. In this case we have no closed form expression for the distribution. We do however have the characteristic function, from which the distribution can be obtained by means of a numerical inverse Fourier transform. We have performed this calculation, and used the result as a the benchmark against which our path simulation algorithms are compared.

In this case, the Levy density is

$$\nu(x) = \begin{cases} Ce^{-Mx}/x^{1+Y} & , x > 0 \\ Ce^{-G|x|}/|x|^{1+Y} & , x < 0. \end{cases} \quad (3.13)$$

The parameter space for the CGMY model is $C, G, M > 0$ and $Y < 2$. By expanding the Levy density in negative powers of x around the origin and keeping only the lowest order terms, we get

$$\sigma(\epsilon) \sim \epsilon^{1-Y/2},$$

so by (3.10), the jump-diffusion approximation is valid only for $0 < Y \leq 1$. In fact, the volatility can in this case be expressed exactly in terms of the incomplete gamma function, which we will not show here.

We simulated the CGMY process with parameters $C = G = M = 1$, $Y = 1/2$, using $\epsilon = 0.005$, and 100000 paths. In this case, the volatility for our choice of parameter values is

$$\sigma \approx 0.022.$$

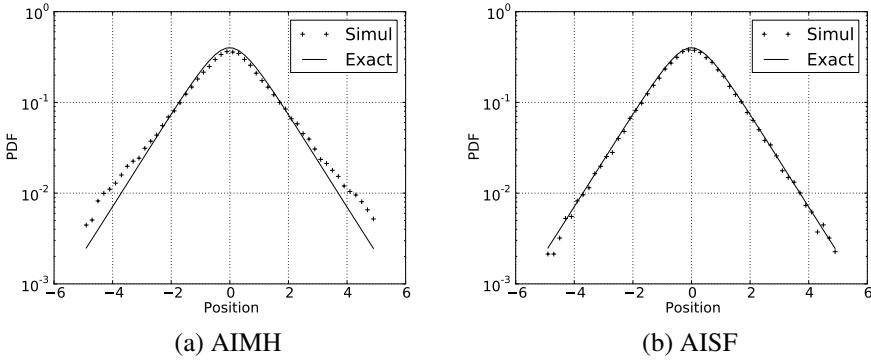


Figure 3.9: CGMY PDF at $t = 1$, simulated and exact results. Note again the heavy tails in the AIMH case.

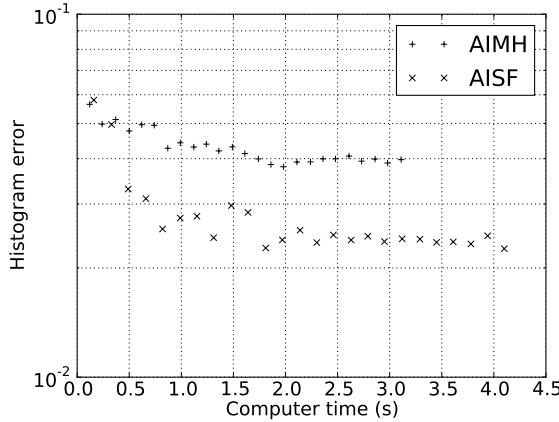


Figure 3.10: CGMY histogram error at $t = 1$, versus computer time.

Also in this case it has negligible influence on the distribution tails at $t = 1$.

The results for the distribution at $t = 1$ of the process is given in Figure 3.9, and the convergence measurements are plotted in Figure 3.10. The conclusions are similar to the NIG case.

3.7 Discussion

We have studied two algorithms for jump-diffusion and infinite activity pure jump process simulation via jump-diffusion approximations. Most importantly, we have studied the proposed SF algorithm based on a stochastic step function. It has some advantages over MH accept/reject algorithms. It is possible to configure it to have an arbitrarily small autocorrelation. Our simulations show that in the case of simulation of Levy processes, this algorithm can represent an improvement over the MH method that we have considered. The numerical results show an improvement in the tails of the distribution of the Levy process at a given time while at the same time converging faster. The MH algorithm will tend to give heavy tails, due to the problem of positive correlations between large jump values.

In our computer simulations, we also implemented a subdomain discretisation with a cor-

responding adaptively improved discrete probability distribution. This method helps to reduce correlations for the MH algorithms, since the subdomains themselves are drawn without using an accept/reject algorithm. In the SF case, it improves the variate generation speed while maintaining the correlation-free nature of the method.

Chapter 4

Worm PIMC for Canonical Ensembles

4.1 Introduction

Path integral Monte Carlo (PIMC) is a well-established simulation method suitable for calculations of statistical properties of a nonrelativistic quantum many-particle systems of scalar bosonic point particles in thermodynamic equilibrium. It is based on a discrete approximation of the path-integral formulation of the density operator (2.24) of such a system. Essentially, it calculates the many-particle low-temperature density matrix, given the high-temperature density matrix as input data, by representing each particle mathematically as a classical path. See [15] for an excellent review, and references therein.

It is not possible to simulate the exact behaviour of each constituent particle in a many-particle system of interacting particles. So we cannot predict the detailed paths of all constituent particles in a complicated system. However, as the number of particles becomes large, other types of descriptive quantities start to make more and more sense. These are the thermodynamical properties such as pressure, temperature, mass density and entropy. For example, for a given mass density and temperature, we can calculate the energy density and specific heat capacity of the system. In addition, we can analyse transitions between different phases of the same system.

Statistical physics works under the assumption that any given non-zero temperature, the constituent particles in a system occupy random energy states, governed by known probability distributions that depend on the type of constituent particles. The system as a whole is in a random energy state, determined by the Gibbs distribution (2.27). This is the case for all thermodynamic systems at a fixed temperature. This has consequences for the distribution of each constituent particle among the available single-particle energy states of the system. For example, in a system of bosonic particles, the constituent particle energy states are distributed according to the Bose-Einstein distribution. Fermion states are distributed according to the Fermi-Dirac distribution. Distinguishable, classical particles are distributed according to the Maxwell-Boltzmann distribution. For a detailed account of statistical physics, refer to [45].

Since the many-particle system is in a random state, and we wish to calculate expectation values of probabilistic quantities, we must perform multi-dimensional integrals of very high dimension, over the space of all possible system states. This can be done using Monte Carlo integration. Since the path integral method represents a d -dimensional quantum system as a

$(d + 1)$ -dimensional classical system of paths, this involves performing sampling across the space of all such possible paths, using the Gibbs measure. The resampling methods involve translation movements of whole particle paths, as well as resampling of the shapes of each particle path. Each configuration of particle paths have a given probability weight according to the Gibbs distribution, and the Metropolis/Hastings (MH) method is used to sample system configurations that are distributed accordingly.

At low temperature, quantum bosonic systems can form a Bose-Einstein condensate (BEC). This happens when a macroscopic fraction of the constituent particles occupy the same quantum state. The experimental consequences are for example the well-known phenomena of superfluidity and superconductivity. These are examples of cooperative or collective behaviour that result from quantum physics. When several particles occupy the same state, their quantum states are *entangled*, or *exchanged*. This is represented in the path integral formalism by the paths of different particles being connected. To sample the entire configuration space of the quantum system, we must include all possible connections between the constituent particles. This is the most difficult part of PIMC simulations, since it is not obvious how one can propose exchange resampling moves that closely follow the Gibbs distribution. In other words, most single exchange resampling moves are energetically unfavourable, the therefore represent low probability configurations, and therefore create barriers between system configuration domains of different exchange configurations.

One way to alleviate this problem is by the introduction of a worm, a non-closed path in the simulation algorithm, in addition to the closed paths that constitute the system [12, 11]. The system configuration is defined as *diagonal* when the worm is not present, and *off-diagonal* when it is present. Measurement of physical observables like energy are made when the system is diagonal. However, while in the off-diagonal state, the worm path is allowed to propagate in several ways to be described below, resulting in changes in the exchange configuration of the system. When the worm closes, the system can be in a completely different exchange configuration. This helps to reduce the energetically unfavourable barrier that separates the different exchange configuration domains.

The Worm PIMC (WPIMC) method was proposed initially for simulations in the *grand canonical ensemble*, in which the system particle number is variable, and a corresponding chemical potential is introduced. However, we would like to formulate the method in the canonical ensemble with a fixed number of particles.

4.2 Physical system

The physical system we wish to simulate consists of N indistinguishable bosonic point particles, each of mass m , inhabiting the space \mathbb{R}^d , characterised by no other properties other than the fact that their pair interaction potential is of a given form. We wish to consider such mutually interacting particles confined by a confining external potential about the origin, in thermal equilibrium at some temperature T low enough so that a nonrelativistic description is adequate. Our goal is to be able to calculate statistical properties as functions of temperature, such as energy and heat capacity, and other properties of the particle configuration having to do with Bose-Einstein condensation, i.e. superfluidity. Since we are working with a fixed number of

particles, at a fixed temperature, we are using the *canonical ensemble*.

The pair potential $v(r)$ falls roughly as r^{-6} at large distances [7, 44]. Since the number of terms in the potential action goes as N^2 , we must introduce an upper distance potential cutoff for large systems. The short distance nature of the potential allows us to do this without introducing too much error. Things would be more difficult for Coulomb interactions.

In a periodic system with no external potential there is a well-known method, based on the assumption of a homogeneous particle distribution, to compensate for the error introduced by the potential cutoff. This is called a standard tail correction (STC). However, we do not employ such a device, although it would be possible to derive a similar STC for particles in a confining external field, although it would be position dependent.

Also, a common device in simulation of very large systems is the partitioning of the particle into a box structure. The box sizes are chosen in correlation with the range of the potential, so particles not in neighbouring boxes do not interact. This allows to reduce the number of potential terms radically for very large systems. This method, along with the methods of keeping track of particles and updating of box particle content, is called Verlet neighbour lists [4]. The limited diffusion speed of individual particles allows to assume that once a particle is within a given box, it is possible to wait for a given amount of time before again checking which particles are in which box. Such methods are also applicable to particles in a confining potential.

4.3 Mathematical description

It is well-known that a mathematical description of such a physical system can be formulated in terms of path integrals. This is done as follows (see [15] for more detail). Any configuration of well-defined positions for the N particles are encoded in a Nd -vector $\mathbf{R} := (\mathbf{r}_0, \dots, \mathbf{r}_{N-1}) \in \mathbb{R}^{Nd}$, each $\mathbf{r}_n \in \mathbb{R}^d$ being a d -vector describing the position of particle n .

The particles are governed by a Hamiltonian operator

$$H = -\frac{\sigma^2}{2}\Delta + V, \quad (4.1)$$

where the *volatility* σ is given by

$$\sigma := \sqrt{\frac{\hbar^2}{m}}, \quad (4.2)$$

where \hbar is the rescaled Planck's constant $h/2\pi$, and m is the particle mass. It has physical dimensions

$$[\sigma] = L\sqrt{T},$$

where L is a unit of length and T is a unit of temperature. We use Ångström (\AA) and Kelvin (K) as units for these quantities. Note that Kelvin will also be used as the unit of energy, resulting in $k_B = 1$ for Boltzmann's constant. We will see that in the case of free particles, σ^2 is the variance per unit of time of the noise processes that generates the fluctuations of the closed particle paths in the path-integral expression for the density matrix. Note that as mentioned in

(2.10), we are misusing the word time here. More correctly, the stochastic process in question is evolving along the inverse temperature dimension, not physical time. In the case of Helium-4 we have

$$\sigma \approx 3.48 \text{\AA} \sqrt{K} .$$

The Nd -dimensional Laplacian Δ can be written as a sum of individual particle Laplacians

$$\Delta := \sum_n \Delta_n ,$$

each Laplacian Δ_n acting only on the position coordinate $\mathbf{r}_n \in \mathbb{R}^d$ of particle n . The position representation is used for the quantum state space of the system. The notation $|\mathbf{R}\rangle$ is used for the state vector representing the system being in the state of well-defined positions \mathbf{R} . The Hamiltonian H acts on the state space consisting of all such possible states. The potential energy operator $V = V_{\text{int}} + V_{\text{ext}}$ consists in our case of an internal V_{int} and external potential V_{ext} . The internal potential models the interaction between particles and depends on distances between each pair of particles. The external potential acts on each particle independently, only depending on its position. They are both diagonal operators in the position basis of the state space, since they don't depend on derivatives with respect to particle positions. The internal potential is given by a sum over pair potential contributions v . For a many-particle state with position configuration \mathbf{R} , we have

$$V_{\text{int}}(\mathbf{R}) = \sum_{i < j} v(r_{ij}) , \quad (4.3)$$

where the sum is over all pairs (i, j) of particles, and $r_{ij} := |\mathbf{r}_{ij}|$, $\mathbf{r}_{ij} := \mathbf{r}_j - \mathbf{r}_i$. The pair potential v we use for Helium-4 is that of [7].

The density operator $\rho(\beta)$ mentioned in section 2.10 acts as the basic definition of the quantum statistical mechanics. In the position basis $\{|\mathbf{R}\rangle\}$ for the quantum state space on which H acts, we have

$$\rho(\mathbf{R}(0), \mathbf{R}(\beta); \beta) = \langle \mathbf{R}(\beta) | \rho(\beta) | \mathbf{R}(0) \rangle ,$$

which is not diagonal. This is caused by the fact that a position basis state $|\mathbf{R}\rangle$ is not an energy eigenstate since H contains position derivatives. In anticipation of the path-integral formulation, we have alluded to the fact that \mathbf{R} will attain a dependency on a “time” coordinate $t \in [0, \beta]$ along the inverse temperature dimension.

By almost the exact same derivation as used in real-time quantum mechanics, we can derive a path-integral expression for the density matrix element by considering the limiting case towards an infinitely fine division along the time dimension,

$$\rho(\mathbf{R}(0), \mathbf{R}(\beta); \beta) = \int_{\mathcal{C}(\mathbf{R}(0), 0; \mathbf{R}(\beta), \beta)} \mathcal{D}_W \mathbf{R} \exp(- \int_0^\beta V(\mathbf{R}(t)) dt) , \quad (4.4)$$

where the integration goes over all continuous paths $\mathbf{R} : [0, \beta] \rightarrow \mathbb{R}^{Nd}$ with the prescribed

endpoints at $t \in \{0, \beta\}$, and $\mathcal{D}_W(\mathbf{R})$ is the *conditional Wiener measure* generated by Nd -dimensional Brownian motion with volatility σ^1 [16]. Intuitively written,

$$\mathcal{D}_W \mathbf{R} = \exp \left(- \int_0^\beta \frac{\dot{\mathbf{R}}^2}{2\sigma^2} dt \right) \prod_{\tau \in [0, \beta]} \frac{d\mathbf{R}(\tau)}{(2\pi\sigma^2 d\tau)^{Nd/2}}.$$

Convergence of the integral (4.7) is assured when V is bounded from below, as is the case for the Aziz potential, and for external potential that increase quickly enough with increasing $|\mathbf{R}|$. As an example, the Coulomb potential would require further analysis, since it has negative singularities in \mathbf{R} -space at which the integrand blows up.

Thus the n -th particle of this system is represented mathematically as a continuous curve $\mathbf{r}_n : [0, \beta] \rightarrow \mathbb{R}^d$. These paths are collected into $\mathbf{R} : [0, \beta] \rightarrow \mathbb{R}^{Nd}$, defined component-wise as

$$\mathbf{R}(t) = (\mathbf{r}_0(t), \dots, \mathbf{r}_{N-1}(t)).$$

In Equation (4.4), each particle has a well-defined path. Therefore, it describes the density operator for a system of distinguishable particles. However, our example of Helium-4 consists of a system of indistinguishable particles. The partition function $Z(\beta)$ in Equation (2.28) can be expressed by noting that the trace forces $\mathbf{R}(0)$ and $\mathbf{R}(\beta)$ to coincide in Equation 4.4. However, this assumes that the density operator describes indistinguishable particles. Since ours don't, we can fix this by allowing the path $\mathbf{R}(\beta)$ to coincide with $\mathbf{R}(0)$ only up to some permutation P of the N particles, while at the same time summing over all possible such permutations in order to obtain the required complete symmetry under exchange of any pair of constituent particles. Therefore, the correct expression for the partition function is

$$Z(\beta) = \sum_P \int d\mathbf{R} \rho(\mathbf{R}, P\mathbf{R}; \beta), \quad (4.5)$$

Here, P acts on an Nd -vector \mathbf{R} by rearranging the particle labels,

$$\mathbf{P}\mathbf{R} = (\mathbf{r}_{P(0)}, \mathbf{r}_{P(1)}, \dots, \mathbf{r}_{P(N-1)}). \quad (4.6)$$

Therefore, $Z(\beta)$ can be expressed as a path integral over all closed paths

$$Z(\beta) = \int_{\mathcal{C}(0; \beta)} \mathcal{D}_W \mathbf{R} \exp \left(- \int_0^\beta V(\mathbf{R}(\tau)) d\tau \right), \quad (4.7)$$

where the integration now goes of all paths $\mathbf{R} : [0, \beta] \rightarrow \mathbb{R}^{Nd}$ with $\mathbf{R}(\beta) = \mathbf{P}\mathbf{R}(0)$ for any permutation P . We have absorbed the permutation into the integration measure by expanding the function space to include paths with all possible permutation connections at $t = \beta$. We denote such paths as closed, even though strictly speaking we only have that $\mathbf{R}(0)$ and $\mathbf{R}(\beta)$ are equal as sets of d -vectors.

¹Brownian motion B_t of volatility σ satisfies $\text{Var}[B_t] = \sigma^2 t$. Its distribution $\rho(x, t)$ satisfies a diffusion equation $\dot{\rho} = D\Delta\rho$ with diffusion constant $D = \sigma^2/2$.

The partition function (2.28) is a sum over the space of energy eigenvalues $|i\rangle$, of the relative probabilities $\exp(-\beta E_i)$ for the system to exist in each energy eigenstate. The same interpretation is valid for the position state representation used in equation (4.5). It is a sum over the space of position value \mathbf{R} , and the integrand $\rho(\mathbf{R}, P\mathbf{R}; \beta)$ is the relative probability of the system to be in state $|\mathbf{R}\rangle$ with permutation P .

Since $Z(\beta)$ and $\rho(\beta)$ are given as integrals over path spaces, it is also possible to interpret the corresponding integrands as densities on the corresponding infinite dimensional path space. These densities are non-negative, since $V(\mathbf{R}) \in \mathbb{R}$.

The expectation value of any physical observable \mathcal{O} is given by

$$\text{Avg}[\mathcal{O}] = Z(\beta)^{-1} \mathbf{tr}[\mathcal{O}\rho] = Z(\beta)^{-1} \sum_P \int d\mathbf{R}' d\mathbf{R} \mathcal{O}(\mathbf{R}', \mathbf{R}) \rho(\mathbf{R}, P\mathbf{R}', \beta). \quad (4.8)$$

If $\mathcal{O}(\mathbf{R}, \mathbf{R}')$ has nonzero off-diagonal elements, we get contributions from off-diagonal elements of the density matrix ρ . However, the measurements done by the PIMC simulation is only on closed paths, and does not allow modifications of this by such off-diagonal operators. However, expectation values of thermodynamic quantities are available through the partition function. The integral in (4.7) goes over a function space already endowed with the Wiener measure. Now, the product of the Wiener measure density and the integrand defines a measure which governs the relative probabilities of the different paths that contribute to expectation values in the statistical system. This is however not a probability measure, since it is unnormalised. However, as mentioned, the Metropolis/Hastings algorithm does not depend on a normalised density. Therefore, if we can sample paths from a probability density proportional to this complicated density, we have obtained an algorithm which correctly simulates the physical system. In addition, we would be able to determine expectation values of arbitrary operators that act on the state space, using the formula (4.8). We will accomplish this for a discretised system, as described below.

As an example, consider noninteracting particles in a finite periodic box of volume V . If we assume that the length scale of the path fluctuations dictated by σ and β are small compared to the side length of the box, we can neglect corrections to the Wiener density. In that case, since $V = 0$, we get

$$\rho(\mathbf{R}(0), \mathbf{R}(\beta); \beta) = \int_{C(\mathbf{R}(0), 0; \mathbf{R}(\beta), \beta)} \mathcal{D}_W \mathbf{R} = (2\pi\sigma^2\beta)^{-Nd/2} \exp\left(-\frac{(\mathbf{R}(\beta) - \mathbf{R}(0))^2}{2\sigma^2\beta}\right). \quad (4.9)$$

It is not a coincidence that this coincides with the expression for the transition probabilities for the Wiener process of volatility σ , since the noninteracting systems is completely described by such a stochastic process. Therefore, for distinguishable particles, the partition function is

$$Z(\beta) = V^N (2\pi\sigma^2\beta)^{-Nd/2}.$$

According to the usual expression for the energy in terms of the partition function given in

Equation 2.29, the average energy per particle is therefore

$$\frac{E}{N} = \frac{1}{N} \left[-\frac{d}{d\beta} \log Z(\beta) \right] = \frac{d}{2\beta} = \frac{d}{2} T,$$

which is the known correct result for noninteracting distinguishable particles at temperature T .

At each value of the inverse temperature t , $\mathbf{R}(t)$ describes the instantaneous system state at time t by specifying the N positions of the particles. The physical interpretation of this is that the spatial fluctuations in the position of each particle due to quantum effects is modelled by the spatial extension of its path. This has no classical analogy. A classical system can be approached by letting each closed path shrink to a point. The fluctuations due to thermal effects is encoded by the movements due to the acceptance/rejectance of the proposed moves by the Metropolis/Hastings algorithm, to be described later. This is the case for both classical and quantum systems. For a classical particle, the thermal fluctuations simply move it around, but for a quantum particle the thermal fluctuations move it around as well as acting on its quantum state. There are more degrees of freedom in a quantum system, and this is why there is an isomorphism between d -dimensional quantum systems and $(d+1)$ -dimensional classical systems. In this case, the $(d+1)$ -dimensional space $[0, \beta] \times \mathbb{R}^d$ is used to formulate a classical description in terms of the positions $\mathbf{R}(t)$ of the d -dimensional quantum system. The temperature dependence in the simulation comes in both in the length of the added dimension, as well as governing the size of the allowed fluctuations as dictated by the form of the Boltzmann distribution. We postpone until later the meaning of nontrivially permuted connections between different particles.

The relative probability that a randomly chosen system in the statistical ensemble is in a state $|\mathbf{R}\rangle$ is given by $\rho(\mathbf{R}, \mathbf{R}; \beta)$. On the other hand, the off-diagonal $\rho(\mathbf{R}, \mathbf{R}'; \beta)$ represent relative probabilities for transition processes $\mathbf{R}' \rightarrow \mathbf{R}$ during the time interval $[0, \beta]$. We can associate an effective action with such a transition,

$$S(\mathbf{R}, \mathbf{R}'; \beta) := -\log(\rho(\mathbf{R}, \mathbf{R}'; \beta)). \quad (4.10)$$

It is natural to decompose S into kinetic (K) and potential (U) actions. In the noninteracting case, using (4.9), the potential action vanishes and we have

$$K(\mathbf{R}, \mathbf{R}'; \beta) = \frac{Nd}{2} \ln(2\pi\sigma^2\beta) + \frac{(\mathbf{R}' - \mathbf{R})^2}{2\sigma^2\beta}. \quad (4.11)$$

It is convenient to define this to be the exact kinetic action even in the interacting case. In other words, we consider all additions to this as parts of the effective potential action U . Transitions for which the effective action S is large are improbable, and will not contribute much to the partition function Z .

4.4 Discretisation

It is not possible to express the density operator and partition function in closed form for most interacting systems. However, we are able to approximate the density matrix for high temper-

atures. We can use a semiclassical approximation in those cases. Of course, β is fixed by the physical system we consider, but it is possible to express the density matrix $\rho(\mathbf{R}, \mathbf{R}'; \beta)$ in terms of high temperature density matrices $\rho(\mathbf{R}, \mathbf{R}'; \tau)$, i.e. for small τ . Actually, the path-integral (4.4) expresses it in terms of the limit of an infinite number of infinitesimal inverse temperature density matrices. By the composition property of the Wiener measure, it is possible to split the path integral into a finite number of integrals of factors. For any positive integer M , we have the identity

$$\rho(\mathbf{R}_0, \mathbf{R}_M; \beta) = \int d\mathbf{R}_1 \dots d\mathbf{R}_{M-1} \rho(\mathbf{R}_0, \mathbf{R}_1; \tau) \rho(\mathbf{R}_1, \mathbf{R}_2; \tau) \dots \rho(\mathbf{R}_{M-1}, \mathbf{R}_M; \tau), \quad (4.12)$$

where the discrete inverse temperature step is

$$\tau := \beta/M.$$

This follows also by a repeated application of the composition property (2.26). We have defined $M + 1$ uniformly separated discrete time slices $t_m = m\tau$ for $m \in \{0, 1, \dots, M\}$, on $[0, \beta]$. The collective path position at time t_m is denoted $\mathbf{R}_m := \mathbf{R}(t_m)$. The position of particle n at time slice m is written $\mathbf{r}_{n,m} := \mathbf{r}_n(t_m)$. Thus, we have at time slice m ,

$$\mathbf{R}_m = \{\mathbf{r}_{0,m}, \mathbf{r}_{2,m}, \dots, \mathbf{r}_{N-1,m}\}.$$

Each $\mathbf{r}_{n,m}$ is called a bead.

Finally, we define link m , for $m \in \{0, 1, \dots, M - 1\}$ as the time interval $[t_m, t_{m+1}]$. On the m -th link, a transition from \mathbf{R}_m to \mathbf{R}_{m+1} takes place. It is governed by the link density matrix $\rho^m := \rho(\mathbf{R}_m, \mathbf{R}_{m+1}; \tau)$. This is still given by a path integral, but on the time subset corresponding the the link time interval,

$$\rho^m(\mathbf{R}_m, \mathbf{R}_{m+1}; \tau) = \int_{\mathcal{C}(\mathbf{R}_m, t_m; \mathbf{R}_{m+1}, t_{m+1})} d_W \mathbf{R} \exp\left(-\int_{t_m}^{t_{m+1}} V(\mathbf{R}(t)) dt\right). \quad (4.13)$$

Corresponding to this we have an effective link action

$$S^m(\mathbf{R}_m, \mathbf{R}_{m+1}; \tau) = -\log(\rho^m(\mathbf{R}_m, \mathbf{R}_{m+1}; \tau)) \quad (4.14)$$

Similarly as above, the exact kinetic link action for the m 'th link is

$$K^m(\mathbf{R}_m, \mathbf{R}_{m+1}; \tau) = \frac{Nd}{2} \log(4\pi\lambda\tau) + \frac{(\mathbf{R}_{m+1} - \mathbf{R}_m)^2}{2\sigma^2\tau}. \quad (4.15)$$

The problem is now reduced to finding the effective link potential action $U^m(\mathbf{R}_m, \mathbf{R}_{m+1}; \tau)$. The primitive approximation

$$U_{\text{prim}}^m(\mathbf{R}_m, \mathbf{R}_{m+1}; \tau) := \frac{\tau}{2} (V(\mathbf{R}_m) + V(\mathbf{R}_{m+1})) \quad (4.16)$$

is a good approximation for $\tau \ll \beta$ (i.e. for large M). In the limiting case, this is exact.

For small τ , the important paths on the m -th link in the integral (4.13) don't stray much from \mathbf{R}_m or \mathbf{R}_{m+1} because of the damping that occurs in the Wiener measure, since the volatility is limited. Therefore, it makes sense to approximate the potential action by using the average of its values at the path endpoints. Better approximations can be obtained by considering the positions reached by an intermediate Brownian bridge between these endpoints.

With these definitions, and using equation (4.12), it is possible to express the complete density operator as

$$\rho(\mathbf{R}_0, \mathbf{R}_M, \beta) := \int d\mathbf{R}_1 \dots d\mathbf{R}_{M-1} \exp(-\sum_m S^m).$$

Similarly, the partition function is

$$Z(\beta) = \sum_P \int d\mathbf{R}_0 \dots d\mathbf{R}_{M-1} \exp(-\sum_m S^m(\mathbf{R}_m, \mathbf{R}_{m+1}; \tau)), \quad (4.17)$$

where $\mathbf{R}_M = P\mathbf{R}_0$.

4.5 Pair action approximation

In the limit $\tau \rightarrow 0$, the high temperature N -particle single link interaction density matrix splits into a product of contributions for each pair of particles. This is a consequence of assumption (4.3) about the potential. In other words, the total single link potential action U^m can be expressed as a sum over pair contributions,

$$U^m(\mathbf{R}, \mathbf{R}'; \tau) \approx \sum_{i < j} u(\mathbf{r}_{ij}, \mathbf{r}'_{ij}; \tau), \quad (4.18)$$

where the pair link interaction u depends only on the relative separations \mathbf{r}_{ij} and \mathbf{r}'_{ij} for the pair of particles at the beginning and end of the link.

In order to determine u , we consider a pair of particles 1 and 2 on a link $[t, t + \tau]$. We have defined the centre-of-mass and relative coordinates at time t as

$$\begin{aligned} \mathbf{x} &:= (\mathbf{r}_1 + \mathbf{r}_2)/2 \\ \mathbf{d} &:= \mathbf{r}_2 - \mathbf{r}_1, \end{aligned} \quad (4.19)$$

and primed version are used for the same quantities at time $t + \tau$. As before, the kinetic action for the pair is defined exactly, and is given by

$$k := \frac{2 \cdot d}{2} \ln(2\pi\sigma^2\tau) + \frac{(\mathbf{r}'_1 - \mathbf{r}_1)^2}{2\sigma^2\tau} + \frac{(\mathbf{r}'_2 - \mathbf{r}_2)^2}{2\sigma^2\tau}.$$

By the definitions (4.19), we have

$$(\mathbf{r}'_1 - \mathbf{r}_1)^2 + (\mathbf{r}'_2 - \mathbf{r}_2)^2 = 2(\mathbf{x}' - \mathbf{x})^2 + \frac{1}{2}(\mathbf{d}' - \mathbf{d})^2.$$

We now define the centre-of-mass and relative kinetic actions as follows,

$$\begin{aligned} k^{\text{com}} &= \frac{d}{2} \ln(2\pi\sigma_{\text{com}}^2\tau) + \frac{(\mathbf{x}' - \mathbf{x})^2}{2\sigma_{\text{com}}^2\tau} \\ k^{\text{rel}} &= \frac{d}{2} \ln(2\pi\sigma_{\text{com}}^2\tau) + \frac{(\mathbf{d}' - \mathbf{d})^2}{2\sigma_{\text{rel}}^2\tau}, \end{aligned} \quad (4.20)$$

where the volatilities $\sigma_{\text{com}/\text{rel}}$ for centre-of-mass and relative motion are related to the nominal Brownian motion volatility by

$$\sigma_{\text{com}} = \sigma/\sqrt{2} \quad , \quad \sigma_{\text{rel}} = \sigma\sqrt{2}. \quad (4.21)$$

This is intuitively understandable from the fact that for a given erratic motion of \mathbf{r}_1 and \mathbf{r}_2 , the centre of mass motion \mathbf{x} is less erratic due to the averaging, and the opposite holds for the relative coordinate \mathbf{d} . Mathematically, it follows from the procedure of factorizing the kinetic density matrix which is a normalized probability distribution, into two normalized probability distributions depending on centre-of-mass and relative coordinates, respectively. These are the two kinetic density matrices for centre-of-mass and relative movement.

The total slice action s for this pair can now be written

$$s(\mathbf{x}, \mathbf{x}', \mathbf{d}, \mathbf{d}') = k_{\text{com}}(\mathbf{x}, \mathbf{x}') + s_{\text{rel}}(\mathbf{d}, \mathbf{d}'), \quad (4.22)$$

where the relative action s^{rel} is

$$s_{\text{rel}}(\mathbf{d}, \mathbf{d}') := k_{\text{rel}}(\mathbf{d}, \mathbf{d}') + u(\mathbf{d}, \mathbf{d}').$$

Thus the complete pair link density matrix factorizes,

$$\rho(\mathbf{x}, \mathbf{x}', \mathbf{d}, \mathbf{d}') = \rho_{\text{com}}(\mathbf{x}, \mathbf{x}')\rho_{\text{rel}}(\mathbf{d}, \mathbf{d}'). \quad (4.23)$$

Our goal is to determine u . It can be determined through the Feynman-Kac formula

$$\rho_{\text{rel}}(\mathbf{d}, \mathbf{d}'; \tau) = \text{Avg} \left[\exp \left(- \int_t^{t+\tau} V(\mathbf{r}(t')) dt' \right) \right],$$

where the average is taken over the space of all Brownian bridge paths of volatility σ_{rel} from \mathbf{d} to \mathbf{d}' on the time interval $[t, t + \tau]$. This expression lends itself well to a Monte Carlo simulation of the relative coordinate. However, this is inefficient compared to the method we will describe next, called matrix squaring.

Due to the radial symmetry, the relative pair density matrix ρ_{rel} can be expressed as a partial wave expansion using Legendre polynomials [53, 43],

$$\rho_{\text{rel}}(\mathbf{d}, \mathbf{d}'; \tau) = \frac{1}{4\pi dd'} \sum_{l=0}^{\infty} (2l+1) \rho_l(d, d'; \tau) P(\cos \theta).$$

Here, d, d' are the lengths of \mathbf{d}, \mathbf{d}' , and θ is the angle between them. The partial wave relative density matrices $\rho_l(d, d'; \tau)$ are single-particle densities for a particle moving on \mathbb{R}^+ in a mod-

ified potential $\tilde{v}(d) := v(d) + l(l+1)/d^2$, where v is the Helium two-particle potential. The modified potential term pushes the particle away from the origin more and more for higher l , so for a given domain $d \in [0, d_{\max}]$ only a finite number of partial wave terms are needed for a given accuracy.

As any density matrix, ρ_l satisfies the squaring property

$$\rho_l(d, d'; 2\tau) = \int \rho_l(d, r; \tau) \rho_l(r, d'; \tau) dr , \quad (4.24)$$

relating density matrices at temperatures differing by a factor of 2. By performing this integral iteratively m times, it is possible to obtain the density matrix at a temperature T from the density matrix at temperature $T \cdot 2^m$. As a starting point, we use the high temperature semiclassical approximation given in [15, eq. 4.42]

$$\rho_l(d, d'; \tau) = \rho_l^0(d, d'; \tau) \exp \left(-\frac{\tau}{|d - d'|} \int_d^{d'} v(x) dx \right) , \quad (4.25)$$

where the free relative partial wave density matrix is

$$\rho_l^0(d, d'; \tau) = \frac{4\pi dd'}{\left(2\pi\sigma_{\text{rel}}^2\tau\right)^{3/2}} \exp \left(-\frac{d^2 + d'^2}{2\sigma_{\text{rel}}^2\tau} \right) i_l \left(\frac{dd'}{\sigma_{\text{rel}}^2\tau} \right) . \quad (4.26)$$

In this expression, i_l is the regular modified spherical Bessel function of order l , which are related to the regular modified cylindrical Bessel function of half-integer order by $i_l(x) = \sqrt{\pi/2x} I_{l+1/2}(x)$.

The iterative procedure given in equation (4.24) corresponds to simple matrix squaring in the discrete case with a uniform grid. However, errors will arise in $\rho(d, d'; 2\tau)$ at large d, d' since the cutoff neglects important contributions in that case. If d_{\max} is large enough, the semiclassical approximation (4.25) can be applied to correct for this, at each step in the iterative procedure, as detailed in [43].

By subdividing the small interval $[0, \tau]$ into B even smaller equal pieces, we can express the high temperature density matrix as a multi-dimensional integral as before,

$$\exp(-s) = \int dR_1 \dots dR_{B-1} \exp \left(- \sum_{m=0}^{B-1} k^m + u^m \right) ,$$

where $R_i = (\mathbf{x}_i, \mathbf{d}_i)$ are the coordinates of the pair at the i 'th time slice.

where we have moved the kinetic action for the whole time slice over on the right hand side. In this case, the time slices are of length $\tilde{\tau} = \tau/B$. When B is large enough, we can use the primitive approximation for u^m , and perform a Monte Carlo simulation to obtain u , which is the only unknown in this equation. Since u doesn't depend on the centre-of-mass coordinates,

the centre-of-mass kinetic action cancels from this equation, and we obtain

$$\exp(-u) = \exp(k_{\text{rel}}) \int d\mathbf{r}_1 \dots d\mathbf{r}_{B-1} \exp\left(-\sum_{m=0}^{B-1} k_{\text{rel}}^m + u^m\right). \quad (4.27)$$

The probability density for a polygonal Brownian bridge with B segments starting at \mathbf{r}_r and ending at \mathbf{r}'_r is given by

$$P(\mathbf{r}_1 \dots \mathbf{r}_{B-1}) = \exp(k_r) \exp\left(-\sum_{m=0}^{B-1} k_r^m\right).$$

Therefore, we have

$$\exp(-u) = \mathbf{Avg}_{\text{BB}} \left[\exp\left(-\sum_{m=0}^{B-1} u^m\right) \right], \quad (4.28)$$

where the subscript refers to an average over Brownian bridge paths generated by Brownian motion of volatility σ_r , since this is the volatility parameter that occurs in the relative kinetic action k_r . In order to measure energies, we need the τ -derivative of U , and therefore of u . This is accomplished by differentiating (4.27). This gives

$$\dot{u} = \exp(u) \mathbf{Avg}_{\text{BB}} \left[(\sum \dot{k}^m + \dot{u}^m) \exp(-\sum u^m) \right] - \dot{k},$$

where the dot refers to differentiation with respect to τ , as usual. However, we have had better success with smaller errors by simply calculating at two τ -values $\tau - d\tau$ and $\tau + d\tau$, and using

$$\dot{u}(\tau) \approx \frac{u(\tau + d\tau) - u(\tau - d\tau)}{2d\tau}.$$

So-called matrix squaring is a common method used to calculate the pair density matrix. There one uses the fact that the density matrices at temperatures T and $T/2$ are related by

$$\rho(\mathbf{R}, \mathbf{R}''; \tau) = \int d\mathbf{R}' \rho(\mathbf{R}, \mathbf{R}'; \tau/2) \rho(\mathbf{R}', \mathbf{R}''; \tau/2).$$

This equation is used in a recursive manner n times to obtain $\rho(\tau)$ from a high-temperature semi-classical approximation for $\rho(\tau \cdot 2^{-n})$.

4.6 Sampling methods

In order to calculate expectation values of observables, we must generate bead configurations that are distributed according to the Boltzmann weights that appear in the partition function. This is done using a Metropolis/Hastings algorithm.

4.6.1 Metropolis/Hastings

We have seen that we can calculate observables for the statistical system by generating a Markov chain of particle paths that are distributed according to the discretised Boltzmann distribution

$$\pi(\mathbf{R}) = \exp(-S(\mathbf{R})) = \exp(-K(\mathbf{R})) \exp(-U(\mathbf{R})),$$

where we have left the particle permutation out for now, so this is strictly only correct in the case of distinguishable particles. Here, \mathbf{R} denotes the discretised path of all the particles. The kinetic factor is simply the multidimensional distribution of independent Brownian bridge paths for each particle.

4.6.2 Brownian bridge sampling

Thankfully, we are able to generate bead configurations that exactly sample the Boltzmann distribution in the case of $V = 0$. In that case, the paths are independent Brownian motions, as dictated by the Wiener measure, which we can sample exactly by using the algorithm for a Brownian bridge, as described in section 2.6. This resampling move consists of a resampling of a contiguous subset of bead positions along one or more particle paths. New intermediate bead positions are sampled from the distribution of a Brownian bridge. All the kinetic action terms relating to unchanged links are invariant under this operation.

We now refer to section (2.11), for analysis of the Metropolis/Hastings algorithm. Consider a current particle configuration \mathbf{R}_i . We then proposed a resampled configuration \mathbf{R}_{i+1} using a Brownian bridge between two fixed endpoints. The proposed \mathbf{R}_{i+1} only depends on the position of the endpoints, which are fixed. So it does not depend on the previous configuration \mathbf{R}_i between those endpoints. Therefore, $t(\mathbf{R}_i, \mathbf{R}_{i+1})$ is independent of \mathbf{R}_i , and is equal to

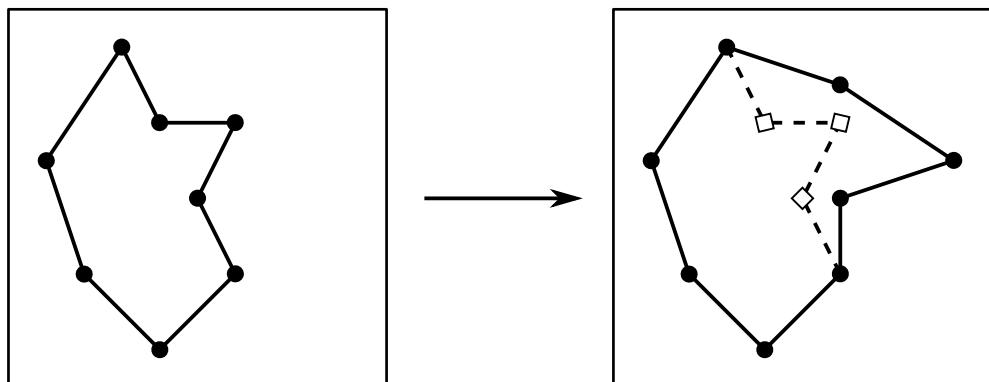
$$t(\mathbf{R}_i, \mathbf{R}_{i+1}) = \exp(-K(\mathbf{R}_{i+1})).$$

Therefore, the probability p to be calculated in the Metropolis/Hastings algorithm becomes

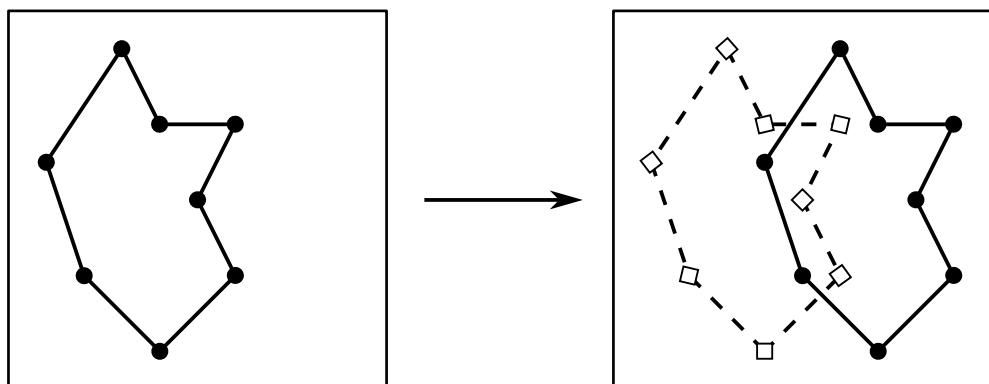
$$p = \min(1, \exp(-U(\mathbf{R}_{i+1}) + U(\mathbf{R}_i))).$$

Since the Brownian bridge resampling enables us to sample the kinetic action exactly, only the potential action will ever partake in the probability calculation for the Metropolis/Hastings algorithm.

As mentioned in 2.6, a bisection algorithm may be employed. This can be used to define a multilevel resampling algorithm. Its purpose is to facilitate an accept/reject Metropolis/Hastings decision before doing all the work that is needed to resample a whole path. A midpoint bead position is resampled, and an approximate action is used to make an accept/reject decision. The idea is that if the initially resampled midpoint bead position is “far from ideal”, it is best to reject the move immediately instead of going on to resample all bead positions along the path, and then reject. For this to work, a family of approximate potential actions $\{U_l\}$ must be defined, and the probability formula for the acceptance/rejection of the fully resampled path must be adjusted so as to give a correct overall probability of accepting/rejecting the path. See [15] for more details.



(a) Brownian motion resampling



(b) Translation resampling

Figure 4.1: Brownian and translation resampling moves. Solid black lines and beads represent the current configuration. Non-filled squares and perforated lines represent the new configuration.

4.6.3 Translation

In this case, whole nonexchanged particle paths are translated, without chaning the relative positions of the continuent beads. The translation distance is adaptively adjusted so as to lead to an acceptance ratio not too far from 1/2. The complete kinetic action is unchanged under this move, since no relative bead positions within any particle are changed.

We translate only non-exchanged particles, since translation of whole systems of connected particles would most likely be rejected anyway. It is also possible to translate several particles independently, and adaptively adjust their number to optimise the algorithm efficiency.

The transition probability density is in this case symmetric and uniform across space. All kinetic parts will cancel from the Metropolis/Hastings probability formula, and we are left with the same simplified expression as for the Brownian bridge resampling case above.

4.7 Exchange effects

A system of indistinguishable quantum particles may find itself in a state in which there is a non-zero probability of finding several particles inhabiting the same quantum state. This effect has no classical analogy. In this mathematical description, this phenomenon occurs for particles that participate in nontrivial permutation connections at $t = \beta$. Due to the image of two or more particles “changing places” when following their paths, this phenomenon is called *exchange*. As can be understood by the fact that the spatial extension of each particle path shrinks towards high temperatures, exchange phenomena are suppressed in that limit even for a quantum system. Of course this is compatible with the fact that the system approaches a classical one at high temperature.

We have seen that the spatial extension of the particle paths is large at low temperatures. This increases the rate of exchange effects in the statistical system. Most importantly, this phenomenon of increased exchange at low temperature is responsible for the λ phase transition for Helium-4 at $T \approx 2.17K$, at which a macroscopic fraction of the constituent particles join into a Bose-Einstein condensate with superfluid properties. At this critical point, the heat capacity of the system diverges, while the energy density remains continuous. Therefore, it is classified as a continuous phase transition, or a second order phase transition.

Chapter 5

Simplicial gauge theory action

5.1 Overview

In this chapter, we describe an alternative finite-element inspired discretisation of gauge field theory. In our terminology, a gauge field theory only describes the dynamics of a single gauge field, and has a continuous gauge symmetry group. Essentially, gauge field theory defines and governs the behaviour a vector field on spacetime, and governing equations enjoy a very large symmetry group that can be likened to a Lie group for every spacetime point. This can be made well-defined through the use of fibre bundles. Usually, in the particular case of non-commutative symmetry groups, we speak of a Yang-Mills field theory [60]. However, we will not make any distinction, and simply write gauge field theory or Yang-Mills field theory, since the fundamental description is identical.

We discuss the numerical consistency of the action in the energy norm in section 5.6. Finally, in section 5.7 we describe some numerical convergence tests, performed on several gauge fields for which the continuum action was exactly calculable. Conclusions are given in section 5.8. We use the same notation as in [21].

5.2 Motivation

5.2.1 General

Gauge field theory lies at the heart of modern physics [55, 56, 51]. We can speak of a classical gauge field theory or a quantum gauge field theory. As noted in chapter 2, both can be defined using a path integral formalism by means of an action. In the classical gauge, we are interested in minimising the action to obtain the equations of motion, which we then attempt to solve. In the quantum case, we are e.g. interested in calculating probabilities for scattering events, or masses of bound states. These depend not only on the classical solution of the field equations, but also on non-solution field configurations, since they also contribute to the path integral. In both cases, the governing quantity for everything is the action.

Quantum gauge field theory has been extremely successful in describing high energy particle physics. This is done using the standard model of particle physics, which is based on the gauge symmetry group $\mathcal{G} = U(1) \times SU(2) \times SU(3)$. In addition to the gauge fields, this model also

describes the behaviour of the fundamental matter particles, such as the electron and the quarks.

In particular, quarks and their mutual interactions are described using a quantum field theory (QFT) with $SU(3)$ gauge symmetry, called by the name Quantum ChromoDynamics (QCD), inspired by a certain analogy involving red, green and blue colours, in addition to the older quantum gauge theory of photons and electrons called Quantum ElectroDynamics (QED).

Despite the massive successes of such models, there are still large calculation difficulties in calculating low energy properties of quarks and gluons. High energy phenomena can be described accurately through the use of perturbation theory, but this is not the case for low-energy phenomena. The problem is that through the effect of renormalisation, the QCD coupling constant increases as interaction energies decrease, in such a way that ordinary perturbation theory is made impossible due to non-convergence. This phenomenon is related to the property of confinement in QCD, whereby interaction strength between quarks increase with distance. Direct paper-and-pen calculation of masses and interactions among low energy bound states of quarks is therefore very problematic.

5.2.2 Discretisation

Therefore, numerical investigations of this theory are important, as an alternative to perturbation theory. At the outset, these are continuum theories, so appropriate numerical descriptions can be obtained by discretising spacetime by some choice of lattice. In so doing, a lot of these difficulties are removed. Usually, this is done on a hypercubic lattice, and such numerical studies have contributed greatly to the advancement of the field of fundamental particle physics, in particular within QCD [59, 26]. Such methods have shown that lattice gauge theory (LGT) is a powerful method of doing nonperturbative quantum gauge theory calculations. They will therefore continue for a long time to be useful in testing QCD against experimental results at low energy.

A hypercubic mesh preserves some discrete subgroups of the translational, mirror and 4d rotational symmetries. Note that a clever way of retaining continuous symmetries while working on a lattice is to use random lattices [20, 18, 19]. The simplicity of a hypercubic lattice can be a great advantage.

But the discrete Wilson action 5.12 on a hypercubic lattice is only one possibility for discretisation of the Yang-Mills action 5.9. Other lattice geometries have also been considered, e.g. random lattices in order to retain rotational symmetry [20, 18, 19].

On a related matter, a notable improvement in computing efficiency for lattice QCD was obtained by using so-called tadpole-improved actions [3]. These allow the use of larger lattice spacing and reduce the required computed resources drastically. We use simple actions not improved by perturbative QCD results. Accuracy is not our primary concern, since this is primarily a proof of concept for non-commutative quantum gauge theory on a simplicial lattice.

5.2.3 Simplicial mesh

In this work we have looked into the possibility of using simplicial 3d lattices (tetrahedral) together with a uniform time discretisation, and expressing the action in a form familiar from finite element methods (FEM), while preserving gauge invariance. This leads to the simplicial

gauge theory (SGT) action. The underlying motivation is to exploit the advantages of such lattices as well as to tap into the wealth of knowledge that have resulted from the development of the finite element method for partial differential equations (PDE). This might help us to formulate a discretised quantum field theory in a more geometry-independent way in order to implement grid refinement, or curved geometries.

The construction of the gauge invariant SGT action functional is inspired by FEM, which is most commonly used for numerical solutions of PDEs, particularly on complicated domains [23, 49, 50, 57, 42]. The formalism therefore includes the use of finite element function spaces on simplicial meshes, and the concept of mass matrices. The term “mass matrix” in this context has nothing to do with physical particle masses, and is therefore not to be confused with the usual mass matrix of quantum states within quantum field theory.

This work is the natural continuation of an earlier work [21]. There, a gauge invariant discrete action to approximate the Yang-Mills-Higgs action was constructed¹ on a simplicial mesh, and consistency of the discrete action was proved for smooth fields. Here, we expand the discussion to 4d spacetime simplexes by including a time dimension, and we expand the consistency proof to fields in the natural energy norm.

More precisely, we define the spatial part of the discrete action as in [21], using Whitney forms [57], Wilson lines and loops [59, 26], and concepts from finite element methods (FEM) [23, 42, 49]. We then expand the spatial mesh to spacetime by introducing a uniform time discretisation. The spacetime mesh is defined by repeating the spatial mesh at every time step. Furthermore, we extend the Whitney forms to spacetime, yielding a natural expansion of the discrete spatial action to a spacetime action.

The purpose of this work is to prove consistency of the resulting approximation scheme for the classical Yang-Mills equations, and to analyse the convergence properties numerically. In chapter 6, we will use this action to perform Monte Carlo quantum field theory simulations. The latter is described in [40].

Simplicial meshes have been used for quantum field theory simulations before [31, 32, 28, 30, 13, 29, 5], with promising numerical results. However, this work goes in a different direction than those earlier works, with a different, more general and more mathematically stringent approach. Our approach results in a formalism that allows different simplicial meshes, and it therefore allows the use of refinement methods.

Through the use of the FEM formulation, and the massive resources of methods available within that subject area, we hope to gain advantages for QCD simulations in future implementations, in particular with regards to the possibilities of grid refinement. This could be useful in modelling some QCD phenomena, e.g. for highly concentrated gluon flux tubes between quarks where an increased lattice resolution might be desired. An earlier work used FEM inspired methods within QFT, although along a different direction involving solutions of operator equations instead of Monte Carlo simulations [9, 10].

The mathematical proof of consistency between the SGT and continuous Yang-Mills gauge theory action is described in the paper [39], along with a description of the more comprehensive Yang-Mills-Higgs model, while quantum simulation is described in [40].

¹Albeit without a Higgs potential.

5.3 The continuous Yang-Mills action

Let $\mathbb{M} = \mathbb{R} \times S$ be a Riemannian spacetime manifold, where \mathbb{R} represents time and S is a bounded domain in three dimensional Euclidean space. We let \mathbb{M} be equipped with a Lorentzian or Euclidean signature and coordinates $x = (t, \mathbf{x})$. Furthermore, let \mathcal{G} be a compact Lie group with associated Lie algebra $\mathfrak{g} = T_{\text{Id}}\mathcal{G}$, and assume that \mathcal{G} can be represented by a subgroup of the complex unitary $n \times n$ matrices, for some n .

It is well-known that \mathcal{G} (the connected component containing the identity) and \mathfrak{g} are related through the exponential map

$$\exp : \mathfrak{g} \rightarrow \mathcal{G}, \quad g \mapsto \exp(g). \quad (5.1)$$

The Hermitian conjugate of a matrix g is denoted g^H and the real valued scalar product is

$$g' \cdot g := \Re \text{tr}(g' g^H), \quad (5.2)$$

The space of smooth k -forms on \mathbb{M} is denoted $\Omega^k(\mathbb{M})$, and the space $\Omega^k(\mathbb{M}) \otimes \mathfrak{g}$ can be identified with the space of smooth \mathfrak{g} -valued k -forms on \mathbb{M} . The bracket of Lie algebra valued forms is defined as

$$[u \otimes g, u' \otimes g'] := (u \wedge u') \otimes [g, g'], \quad (5.3)$$

where u, u' are real valued differential forms and $g, g' \in \mathfrak{g}$.

A connection one-form on \mathbb{M} is an element $\mathbb{A} = (A_0, A) \in \Omega^1(\mathbb{M}) \otimes \mathfrak{g}$, where A_0 represents the time component and A the spatial component. The temporal curvature $\mathcal{F}^t(\mathbb{A})$ and spatial curvature $\mathcal{F}^s(A)$ of such a one-form are given by

$$\mathcal{F}^t(\mathbb{A}) = dA_0 + d_t A + [A_0, A], \quad \mathcal{F}^s(A) = dA + \frac{1}{2}[A, A], \quad (5.4)$$

where d_t and d denote exterior derivatives in the temporal and spatial direction respectively. The independent variable in Yang-Mills theory is such a connection one-form, not necessarily smooth, and the action functional that describes it is given by

$$S[\mathbb{A}] = S[\mathbb{A}]_t + S[A]_s, \quad (5.5)$$

where

$$S[\mathbb{A}]_t := \frac{1}{4g^2} \int_{\mathbb{M}} |\mathcal{F}^t(\mathbb{A})|^2, \quad S[A]_s := \frac{1}{4g^2} \int_{\mathbb{M}} |\mathcal{F}^s(A)|^2.$$

Here, $g > 0$ is the coupling constant, not to be confused with an element of the Lie algebra. This should be clear from the context.

A gauge transformation of a connection one-form \mathbb{A} is associated with a mapping $G : \mathbb{M} \rightarrow \mathcal{G}$, such that

$$A_0(x) \mapsto G(x) (A_0(x) + d_t) G(x)^{-1}, \quad A(x) \mapsto G(x) (A(x) + d) G(x)^{-1}. \quad (5.6)$$

The action $S[\mathbb{A}]$ is invariant under such gauge transformations.

A formulation more familiar to physicists is obtained by expressing the one-form and curvature by means of coordinates. In other words, one decomposes the one-form component A^a in the basis $\{dx^\mu\}$, i.e. $A^a = \sum_\mu A_\mu^a dx^\mu$. The exterior derivative of such a one-form is given by

$$dA^a = \sum_{\mu\nu} \partial_\nu A_\mu^a dx^\nu \wedge dx^\mu = \sum_{\mu\nu} \frac{1}{2} (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a) dx^\mu \wedge dx^\nu. \quad (5.7)$$

Furthermore, the curvature is given by $F^a = \sum_{\mu\nu} \frac{1}{2} F_{\mu\nu}^a dx^\mu \wedge dx^\nu$, where

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - \varepsilon^{abc} A_\mu^b A_\nu^c. \quad (5.8)$$

Finally, the action can be expressed as

$$S = \frac{1}{4g^2} \int_{\mathbb{M}} \sum_{\mu\nu a} F_{\mu\nu}^a F^{a\mu\nu} dx, \quad (5.9)$$

the usual coordinate dependent expression for the Yang-Mills action. Note that the value of e is insignificant for classical field theory, since it doesn't affect the minimum of the action. It does matter for quantum theory, since it affects value of the path integral.

5.4 Wilson action

Consider the special case $\mathbb{M} = [0, 1] \times S$, for $S = [0, L]^3$, and a homogeneous cubic lattice discretisation \mathcal{S} of \mathbb{M} using N divisions in each dimension. In addition, we assume that we are dealing with the Euclidean case. The lattice constant is $h = 1/N$, and L is adjusted to satisfy $L = hN$.

For a gauge field configuration A , define for each edge $e \in \mathcal{S}$ the link gauge variables

$$U_e := \exp(g \int_e A \cdot dl) \in \mathcal{G}, \quad (5.10)$$

where the integration is along edge e . For each face $f \in \mathcal{S}$, encircled cyclically by the edges e_1, e_2, e_3, e_4 , define the face curvature variables

$$\mathcal{W}_f := U_{e_1} U_{e_2} U_{e_3} U_{e_4} \in \mathcal{G}. \quad (5.11)$$

Define the Wilson LGT action for a given configuration of gauge fields as

$$S_W := \beta \sum_{f \in \mathcal{S}} \left(1 - \frac{1}{N_C} \text{TrRe} \mathcal{W}_f \right), \quad (5.12)$$

where N_C is the dimension of the gauge algebra adjoint representation. The subscript alludes to the number of colours, which is three for the case of QCD.

In the limit $N \rightarrow \infty$ with β and L unchanged, S_W converges towards the Euclidean contin-

uum action (5.9) if and only if [26]

$$\beta = \frac{2N_c}{g^2}. \quad (5.13)$$

5.5 Simplicial gauge theory action

In the FEM formulation, we assume \mathcal{T} to be a simplicial complex spanning the spatial domain S . The simplices are referred to as vertices, edges, faces and tetrahedra according to dimension, and are denoted i, e, f and T respectively. The symbol T will also be used for simplices of any dimension. We also suppose an orientation has been chosen for each simplex in \mathcal{T} . In addition, we assume that time is discretised with a time step Δt , and that the simplicial complex \mathcal{T} is repeated at every time step, resulting in a spacetime simplicial complex \mathbb{T} . For a more detailed definition of simplicial complexes, consult [22, section 5].

Thus, as the basic building block in classical FEM theory is a tetrahedron T , the basic building block in this extended FEM version is $T \times I_\tau$, where $I_\tau = [\tau, \tau + \Delta t]$ and τ denotes the temporal nodes.

Furthermore, let $W^k(\mathcal{T})$ ($W^k(T)$) be the space of Whitney k -forms [57, 2] on \mathcal{T} (T), with canonical basis (λ_T) , T ranging over the set of k -dimensional simplexes in \mathcal{T} (T). The 0-forms λ_i are the barycentric coordinate maps taking the value 1 at vertex i and 0 at all others. For an edge $e = \{i, j\}$, the associated Whitney 1-form is defined by

$$\lambda_e = \lambda_i d\lambda_j - \lambda_j d\lambda_i, \quad (5.14)$$

and for a face $f = \{i, j, k\}$ the associated Whitney 2-form is defined by

$$\lambda_f = 2(\lambda_i d\lambda_j \wedge d\lambda_k + \lambda_j d\lambda_k \wedge d\lambda_i + \lambda_k d\lambda_i \wedge d\lambda_j). \quad (5.15)$$

In order to formulate the Yang-Mills theory in a spacetime FEM setting we need to extend these k -forms to \mathbb{T} . In addition, we need to define the temporal edge and temporal face basis functions, which are constructed as in [22].

The spatial Whitney k -forms are extended to be piecewise affine in time and are denoted $\Lambda_{T(\tau)}$, i.e.

$$\lambda_T \rightarrow \Lambda_{T(\tau)} = \lambda_T \otimes P_1^t, \quad (5.16)$$

where P_1^t denote polynomials in the time variable of degree at most one, and $T(\tau) := (\tau, T)$ denotes the spatial simplex T at temporal node τ . More precisely, $\Lambda_{T(\tau)}$ is the piecewise affine function in time, taking the value λ_T at τ and 0 at all other temporal nodes.

The temporal edge basis functions are constructed as follows. To every vertex i in the spatial mesh, there are temporal edges $e_t(\tau) = \{i_\tau, i_{\tau+\Delta t}\}$, where $i_\tau := i(\tau)$. The temporal basis edge function attached to $e_t(\tau)$ is then the piecewise constant function in time defined by

$$\Lambda_{e_t(\tau)}(t) = \begin{cases} \lambda_i \circ \pi \frac{1}{\Delta t} dt, & t \in I_\tau \\ 0, & \text{otherwise.} \end{cases} \quad (5.17)$$

Here, π is the canonical projection onto S , i.e.

$$\pi : \mathbb{M} = \mathbb{R} \times S \rightarrow S , \quad (5.18)$$

and dt is the standard basis one-form in the temporal direction.

Finally, the temporal face elements are constructed as follows. To every spatial edge e there are corresponding temporal faces $f_t(\tau) = e \times I_\tau$. Consider the spatial Whitney edge element λ_e . Then apply the pull-back of π to construct a one-form on spacetime, and then wedge it with dt . More precisely, denote by

$$\pi^* : \Omega(S) \rightarrow \Omega(\mathbb{M}) , \quad (5.19)$$

the pull-back induced by π . Then the temporal basis face function is

$$\Lambda_{f_t(\tau)}(t) = \begin{cases} \pi^*(\lambda_e) \wedge \frac{1}{\Delta t} dt = \lambda_e \circ \pi \wedge \frac{1}{\Delta t} dt , & t \in I_\tau \\ 0 , & \text{otherwise.} \end{cases} \quad (5.20)$$

This construction ensures that the temporal face basis is orthogonal to the spatial face basis. If no confusion can arise, the time dependence index τ that identifies a temporal node is omitted to compactify notation.

Thus, let $\mathbb{A} = (A_0, A) \in \mathbb{W}^1 \otimes \mathfrak{g}$, $\mathbb{A} = \sum_{e_t} A_{0,e_t} \Lambda_{e_t} + \sum_e A_e \Lambda_e$, where the summations are over oriented edges, and we remark that

$$A_{0,e_t} = \int_{e_t} A_0 , \quad A_e = \int_e A . \quad (5.21)$$

The temporal and spatial curvatures of \mathbb{A} are given by

$$\begin{aligned} \mathcal{F}^t(\mathbb{A}) &= \sum_e A_e d_t \Lambda_e + \sum_{e_t} A_{0,e_t} d \Lambda_{e_t} + \sum_{e_t,e} [A_{0,e_t}, A_e] \Lambda_{e_t} \wedge \Lambda_e , \\ \mathcal{F}^s(A) &= \sum_e A_e d \Lambda_e + \frac{1}{2} \sum_{e,e'} [A_e, A_{e'}] \Lambda_e \wedge \Lambda_{e'} . \end{aligned} \quad (5.22)$$

The mass matrices are defined as

$$M_{f_t f'_t} := \int_{\mathbb{M}} \Lambda_{f_t} \cdot \Lambda_{f'_t} , \quad M_{f f'} := \int_{\mathbb{M}} \Lambda_f \cdot \Lambda_{f'} , \quad (5.23)$$

By inspiration from lattice gauge theory [59, 26], we will now in several steps construct an approximation to this action that is gauge invariant, i.e. invariant under the transformation (5.6).

The goal is of course to formulate a gauge invariant action. The intermediate action is not gauge invariant due to terms that contain products of curvatures of different faces with non-coincident distinguished points. However, this can be resolved by parallel transport. We treat the spatial and temporal part separately.

We define the SGT action as

$$S_{\text{SGT}}[\mathbb{A}] := S_{\text{SGT}}[\mathbb{A}]_t + S_{\text{SGT}}[A]_s , \quad (5.24)$$

where

$$S_{\text{SGT}}[\mathbb{A}]_t := \frac{\beta}{2} \Re \sum_{f_t(\tau), f'_t(\tau')} M_{f_t(\tau), f'_t(\tau')} \text{Tr} \left(U_{\dot{f}'_t(\tau) \dot{f}_t(\tau)} [F_{f_t(\tau)}^t - \text{Id}] U_{\dot{f}_t(\tau) \dot{f}'_t(\tau)} \times \right. \\ \left. \times U_{0, \dot{f}'_t(\tau) \dot{f}'_t(\tau')} [F_{f'_t(\tau')}^t - \text{Id}]^H U_{0, \dot{f}'_t(\tau') \dot{f}'_t(\tau)} \right), \quad (5.25)$$

and

$$S_{\text{SGT}}[\mathbb{A}]_s := \frac{\beta}{2} \Re \sum_{f(\tau), f'(\tau')} M_{f(\tau), f'(\tau')} \text{Tr} \left(U_{\dot{f}'(\tau) \dot{f}(\tau)} [F_{f(\tau)}^s - \text{Id}] U_{\dot{f}(\tau) \dot{f}'(\tau)} \times \right. \\ \left. \times U_{0, \dot{f}'(\tau) \dot{f}'(\tau')} [F_{f'(\tau')}^s - \text{Id}]^H U_{0, \dot{f}'(\tau') \dot{f}'(\tau)} \right). \quad (5.26)$$

We can conclude that the simplicial gauge theory action S_{SGT} is discretely gauge invariant.

5.6 Consistency

Since we have formulated the theory in a spacetime FEM setting, we choose to define consistency for the entire action, not only for the spatial part as is usual. This definition encompasses the usual one. Note that we only prove consistency for the gauge field action, without the scalar field. Inclusion of the scalar field is a simple extension of this proof.

We suppose that we have a regular sequence of simplicial meshes \mathcal{T}_n of the spatial domain S . The diameter of a simplex T is denoted h_T , and the largest h_T when $T \in \mathcal{T}_n$ is denoted h_n . In addition, we suppose that time is discretised by a time step Δt_n , and that \mathcal{T}_n is repeated at every time step, resulting in a simplicial mesh \mathbb{T}_n of the spacetime domain \mathbb{M} . We suppose that

$$\max_{\mathbb{T}} \{(h_n), (\Delta t_n)\} \xrightarrow{n \rightarrow \infty} 0. \quad (5.27)$$

The interpolation operators onto the Whitney elements introduced earlier are attached with a subscript n to associate them with the mesh \mathbb{T}_n . Finally, let $X_n = \mathbb{W}^1(\mathbb{T}_n) \otimes \mathfrak{g}$.

Definition 1. We say that two actions S_n and S'_n defined on X_n are consistent with each other, with respect to a norm $\|\cdot\|$, if for any $\mathbb{A} \in X_n$ we have

$$\sup_{\mathbb{A}' \in X_n} \frac{|DS_n[\mathbb{A}]\mathbb{A}' - DS'_n[\mathbb{A}]\mathbb{A}'|}{\|\mathbb{A}'\|} \xrightarrow{n \rightarrow \infty} 0. \quad (5.28)$$

If there is a constant C not depending on n such that quantities a_n and b_n satisfy $a_n \leq C b_n$ for all n , we write $a_n = \mathcal{O}(b_n)$. To compactify notation the subscript n will be suppressed.

We have introduced three different actions $S^J[\mathbb{A}]$, $S^I[\mathbb{A}]$ and $S^L[\mathbb{A}]$, and the plan is to show

1. $S^J[\mathbb{A}]$ consistent with $S[\mathbb{A}]$,
2. $S^I[\mathbb{A}]$ consistent with $S^J[\mathbb{A}]$,
3. $S^L[\mathbb{A}]$ consistent with $S^I[\mathbb{A}]$,

which implies the consistency between $S^L[\mathbb{A}]$ and $S[\mathbb{A}]$.

We will prove consistency in the energy norm, i.e.

$$A_0, A \in L^\infty(\mathbb{R}; H^1(S)), \quad \partial_t A_0, \partial_t A \in L^\infty(\mathbb{R}; L^2(S)), \quad (5.29)$$

where ∂_t is a shorthand for $\partial/\partial t$. To compactify notation we define

$$\begin{aligned} \|\cdot\|_{L^p(L^q)} &:= \|\cdot\|_{L^p(\mathbb{R}; L^q(S))}, \quad \forall 0 < p, q \leq \infty, \\ \|\cdot\|_{L^\infty(H^1)} &:= \|\cdot\|_{L^\infty(\mathbb{R}; H^1(S))}, \\ \|\mathbb{A}\| &:= \|A_0\| + \|A\|. \end{aligned} \quad (5.30)$$

The H^1 spacetime Euclidean seminorm is denoted $|\cdot|_{H^1(\mathbb{M})}$.

Theorem 2. *Assume \mathbb{M} is a bounded domain in \mathbb{R}^{1+3} . Then the SGT action (5.25) is consistent with the continuous Yang-Mills action (5.5), with respect to the norm*

$$\|\mathbb{A}\| := \|\mathbb{A}\|_{L^\infty(H^1)} + \|\partial_t \mathbb{A}\|_{L^\infty(L^2)}, \quad (5.31)$$

under the assumption that the above-mentioned CFL condition holds.

As a consequence of the above estimates, we get the following estimate for the deviation of the SGT action S^L from the continuous action S ,

$$|S(\mathbb{A}) - S^L(\mathbb{A})| \leq \begin{cases} Ch & , \mathbb{A} \in L^\infty(H^1), \partial_t \mathbb{A} \in L^\infty(L^2) \\ Ch^2 & , \mathbb{A} \text{ smooth.} \end{cases} \quad (5.32)$$

5.7 Numerical convergence tests

The preceding sections have defined and proven consistency of the SGT action. However, due to the complexity involved in these quantities, we would like to include some numerical convergence tests as well.

In our computer calculations, we focus on pure gauge theory with $\mathcal{G} = SU(2)$, and used a four-dimensional cubic euclidean domain $[0, 1]^4 \subset \mathbb{R}^4$ with periodic boundary conditions. The lattice structure consisted of a three-dimensional simplicial mesh replicated at each discrete time value. The three-dimensional simplicial mesh consisted of a homogeneous arrangement of N^3 identical cubic building blocks, each building block containing six tetrahedra as shown in figure 5.1. Each such spatial mesh was replicated N times in the time direction to fill the four-dimensional domain, in accordance with the construction detailed in the previous sections. The lattice constant h is defined as the side length of each cubic building block, and also coincides with the time discretisation interval. In the interest of simplicity we enforced temporal gauge, in which the temporal link matrices reduce to identity matrices.

By defining the distinguished points of all spatial and temporal faces to coincide for as many pairs of faces as possible, we only need the parallel transport matrices for terms in the action involving pairs of temporal faces with no common nodes.

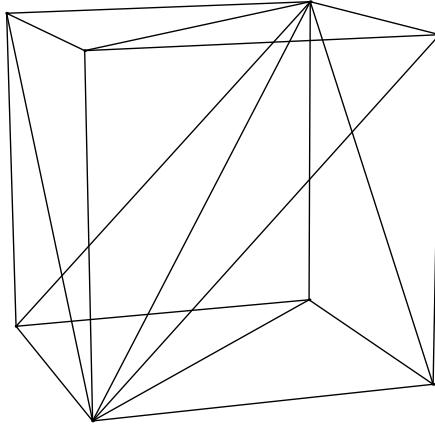


Figure 5.1: Elementary 3d mesh building block containing six tetrahedra, all of which share the single interior diagonal. This particular choice implies an anisotropy in the discretisation. This anisotropy disappears in the continuum limit.

In order to test convergence of the Euclidean SGT, we compared the discrete and continuum action for several different choices of gauge fields for which the continuum action S 5.5 can be calculated exactly. We chose the following cases.

1. Gauge field oriented towards the x -direction in space and towards the generator $t^3 = i\sigma^3/2$ within $\mathfrak{su}(2)$, with a sinusoidal time dependence, where σ^3 is a Pauli matrix. The only nonzero component of the gauge field A is

$$A_x^3(t, x, y, z) := \frac{e}{2\pi} \sin(2\pi t), \quad S = 1.$$

2. Gauge field oriented towards the y -direction in space and t^3 within $\mathfrak{su}(2)$, with a sinusoidal x -dependence. The nonzero component of the gauge field in this case was

$$A_y^3(t, x, y, z) := \frac{e}{2\pi} \sin(2\pi x), \quad S = 1.$$

3. A case with two nonzero components,

$$A_x^1 := \frac{e}{2\pi} \sin(2\pi y), \quad A_y^2 := \frac{e}{2\pi} \sin(2\pi x), \quad S = \frac{1}{2} + \frac{e^2}{8(2\pi)^4}.$$

4. A constant field that only contributes to the nonlinear term in the field strength,

$$A_x^1 := \sqrt{e}, \quad A_y^2 := \sqrt{e}, \quad S = \frac{1}{2}.$$

The link matrices needed to evaluate the SGT action are calculated from these gauge fields by means of the exponential map (5.1)

The first case is insensitive to the spatial face mass matrix elements, while the second is insensitive to the spatial edge mass matrix elements which are used in the definition of the temporal mass matrix elements. In the first two cases, the nonlinear contribution to the continuum

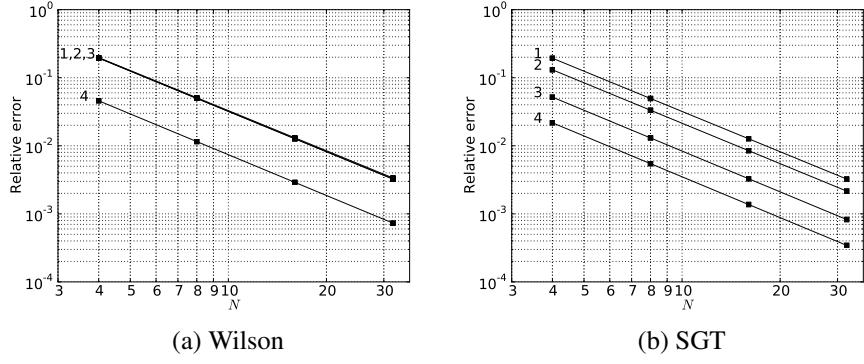


Figure 5.2: The relative error of the Wilson and SGT action versus the number of lattice sites per side N , for the actions 1, 2, 3, 4 described in section 5.7. The squares are the simulation data points and the solid lines are the second order polynomial fits. Errors are proportional to h^2 in all cases.

field strength F vanishes, and the action can be calculated analytically to be unity. In the third case, the nonlinear term survives, and the exact value of the action is

$$S = \frac{1}{2} + \frac{1}{8(2\pi)^4}.$$

In order to provoke a sizeable nonlinear contribution in this case, we chose a small $\beta = 2/e^2 = 1/5$. In all cases, we measured the relative error of the discretised action, versus the lattice size, for lattices sizes N^4 from $N = 4$ to $N = 32$. The results are displayed graphically using double logarithmic plots in figure 5.2 the traditional Wilson action LGT as well as SGT.

In all cases, the errors approach zero as the lattice resolution grows, so the action converges towards the correct continuum result. Using least squares second order polynomial fits, we determined that the relative error depends in the following way on the lattice constant h ,

$$\text{Relative error} \approx Ch^2,$$

where C is some constant depending on the choice of gauge field. This is in accordance with the estimate 5.32.

Note that while the convergence exponent of h is the same in all cases, the prefactor C is smaller in the SGT cases involving time-independent fields, due to its improved spatial discretisation for the same N . Where time-dependence is involved, the errors coincide since the time-discretisation we have chosen for this SGT simulation is of the same quality as for the LGT simulation.

5.8 Discussion

We have proposed a general formulation of lattice gauge theory on simplicial spacetime meshes. For any such constructed spacetime mesh based an arbitrarily shaped 3d simplicial mesh, this action can be used for lattice gauge theory simulations or to study the classical equations of

motion. Traditionally, lattice QCD simulations within physics have used a homogeneous mesh. Mesh refinement is a well established concept within the subject of FEM. We feel that it is well worth the effort to investigate the possibility of mesh refinement within classical and quantum gauge theory.

Quantum lattice gauge theory simulations are very computer intensive. Therefore mesh refinement could be beneficial in cases where it makes sense to focus more computational effort on some subset of the simulation domain.

We have shown the consistency of this SGT numerical approximation to the continuous action, in the sense of approximation theory. The lattice gauge theory formalism is of such a complexity, that it makes sense to complement this theoretical proof with numerical “evidence”. This also serves as an interesting test of whether or not the calculated estimates are saturated in practice. We have provided this for a few different cases of $SU(2)$ gauge fields, for which the action was shown to converge towards the continuum result as the grid fineness increased. Generalization to the gauge group $SU(3)$ of QCD or other gauge groups is in principle a simple matter, although the computational complexity will depend strongly on this.

Chapter 6

Simplicial gauge theory and quantum simulation

6.1 Overview

In this chapter, we employ the discrete SGT action (5.24) in a quantum simulation, measuring the average action density and different Wilson loops. We perform Monte Carlo quantum field theory simulations in order to observe that SGT correctly reproduces the basic aspects of the $SU(2)$ quantum field theory. In this work we shall restrict to a Euclidean spacetime, and the gauge group $SU(2)$. We are assuming some familiarity with concepts from renormalisation of quantum field theories.

After some background in section (6.2), we provide a discussion on the gauge theory coupling constant in the context of lattice field theory and renormalisation, as well as the continuum limit. Then we provide a discussion of some details of the numerical Monte Carlo method as well as numerical results in section (6.4). We end with a short discussion of the SGT discretisation in section (6.5). Lastly, 6.6 contains a short discussion of some aspects of the numerical computer implementation.

6.2 Background

From the gauge theory action, it is possible to deduce the equations of motion by the calculus of variations. These equations allow us to determine the classical solution. One viewpoint is that for any initial condition, the equations of motion allow us to perform a time evolution to obtain the classical time-dependent solution. We say that we are “simulating” the classical system as time evolves, i.e. the mathematical model is evolving in time just like the real system. Of course, solutions procedures for the equations of motion need not do any “evolving in time”.

A quantum system behaves differently. From a given deterministic initial condition, it evolves into a state that may lead to multiple different measurement outcomes. Here we must ask for the probabilities for different outcomes of some measurement on the final state. Each such probability can be calculated, and can be expressed as a path-dependent quantity that is averaged over all possible paths connecting the initial and final state. This is a path integral. Therefore, the classical solution is only one of the many relevant paths between the initial and

final states. These paths do not satisfy any particular equation of motion. Especially not a PDE like the classical equation of motion, since most of the paths are nowhere differentiable in the continuum theory.

Therefore, in the path integral formulation of quantum mechanics, it makes little sense to claim that one is “simulating the system”, if we are to understand the word “simulating” to mean that the mathematical model is following the system as it evolves in time. In the path integral formalism, we are simply calculating average values of different observable quantities, by performing a path integral over a space of paths.

The models are almost always defined so as to also preserve a discrete gauge symmetry. This has the beneficial effect of enforcing a vanishing gluon mass in the discrete model. The ability of a numerical scheme to preserve a symmetry is usually beneficial, since the existence of conserved quantities will usually constrict the numerical solution to not veer too far away from the correct one. However, we must remember that it is possible for a symmetry-preserving scheme to be less accurate than a symmetry-violating scheme.

We perform Monte Carlo quantum pure gauge field theory simulations for the gauge group $SU(2)$ in temporal gauge, as a proof of concept for SGT. Observable measurements include expectation values of the action density as well as a series of different Wilson loops.

- Discuss Monte Carlo simulation.
- Observables.
- Behaviour vs. beta, phase transitions.
- Gauge group $SU(2)$.
- Linear potential.
- Creutz ratios.
- Approx must be good away from the classical solutions.
- It is advantageous to use the language of statistical physics, since we are discussing Euclidean lattice gauge theory which has a statistical interpretation.

6.3 Parameters and continuum limit

The lattice field theory has only the bare coupling constant g as an input parameter. The lattice constant h is a regularisation parameter in the lattice theory, and is initially unknown. It has been absorbed into the independent variables of the lattice theory, the link variables. The drawing of link variables from the distribution defined by the discrete action S_{SGT} happens independently of the value of h . All dimensional observable quantities are automatically calculated in units of powers of the lattice constant h .

However, correlation functions that are calculated by the simulation will depend on spatial lengths, and can be compared with experimental results. This is how h is determined in a given lattice simulation. E.g. the heavy quark-antiquark static potential can be determined for strong coupling by analysing the area-law behaviour of differently sized Wilson loops. At low energy,

this becomes a linear potential, i.e. a nonperturbative demonstration of quark confinement. In reality, other observables from the lattice simulation can be more easily used to determine h , such as quark masses. We indicate the dependency of g on h by writing the bare coupling constant as $g(h)$.

In the lattice theory, $g(h)$ is the bare coupling constant of the regularised theory. This corresponds roughly to the bare coupling constant in a continuum theory with an upper momentum cutoff Λ . When integrating out high-momentum modes and thereby introducing such a momentum cutoff, additional terms are introduced in the Lagrangian. Thus we arrive at an equivalent theory that can be used for calculations of scattering processes involving in- and out-states of momenta used Λ . The result will be independent of Λ .

Our discrete action is simply a discretised version of the non-cutoff continuum action, and the corresponding additional terms that would have been caused by the introduction of such a cutoff by integrating out high-momentum modes, has been neglected. Therefore, the regularised theory is not equivalent to the unregularised one, and therefore the continuum limit must be taken to obtain physically relevant results.

A renormalised coupling constant g_R can be defined, by considering a scattering vertex at some chosen momentum scale r . It will depend on r , h and the bare coupling constant $g(h)$, and therefore we write $g_R(r, h, g(h))$. In the continuum limit $h \rightarrow 0$, it approaches the renormalised coupling constant of the continuum theory, $g_R(r, h, g(h)) \rightarrow g_R(r)$, also defined at the same momentum scale r .

The phenomenon of asymptotic freedom is well-known in $SU(N_c)$ gauge theory, and follows from the noncommutative nature of the gauge group.. At large r , the renormalised coupling constant g_R approaches zero. Therefore, perturbation theory functions well at large energies in such theories. At low energy, under 1GeV in the case of QCD, the renormalised coupling constant approaches unity, rendering perturbation theory useless. This is an indication, but not a proof, of QCD confinement.

Instead of $g(h)$, it is common to instead use $\beta = 2N_c/g(h)^2$ as the free simulation parameter, for $SU(N_c)$ lattice gauge theory. For any choice of β , a corresponding value of h can be determined in the manner described above. The limit $\beta \rightarrow \infty$ (i.e. $g \rightarrow 0$) causes $h \rightarrow 0$, i.e. an approach to the continuum theory.

If, while increasing β in our simulations, it approaches a statistical phase transition $\beta^* < \infty$ at which the correlation length diverges, we would have $h \rightarrow 0$ for which $g \neq 0$. The phase diagram of the Euclidean gauge theory must be determined by nonperturbative calculations, i.e. numerically. It is suspected that such a phase transition does not exist at finite β , and that the limit $\beta \rightarrow \infty$ must be taken to reach the physical continuum limit. Essentially, this means that we must calculate at large β values and extrapolate to $\beta \rightarrow \infty$. Calculations at large β have a strong coupling between plaquettes which causes a reduced rate of convergence. So the choice of β will be a trade-off between computing time and accuracy.

In order for the lattice simulation at strong coupling to correspond to physical reality, there must not be a phase transition that separates it from the asymptotic free continuum limit. This seems to be the case in $SU(N_c)$ lattice gauge theory.

6.4 Computer simulation

For our SGT computer simulations, we chose the Euclidean cubic domain $\mathbb{M} = [0, 1]^4$ with periodic boundary conditions. We simulated the pure gauge SGT action (5.24) in temporal gauge on a simplicial lattice with the gauge group $SU(2)$. Choice of gauge is not necessary, but it does simplify the algorithm slightly, since all temporal edge matrices then reduce to the identity.

The spatial lattice was constructed using a cubic arrangement of N^3 identical building block cubes of size h^3 , each consisting of six tetrahedra as shown in figure 5.1. The resulting spatial mesh was repeated at N consecutive time steps to form a cubic domain of physical volume $(hN)^4$. As described above, each spatial edge is part of two temporal square faces, going forward and backward in time.

The SGT action employs parallel transport matrices in order for gauge invariance to be respected. By defining the distinguished points of all spatial and temporal faces to coincide for as many pairs of faces as possible, we only need the parallel transport matrices for terms in the action involving pairs of temporal faces with no common nodes. More details regarding the exact computer implementation are given in 6.6.

Analogous to the traditional lattice QCD simulations, we performed parallel $SU(2)$ quantum field theory Monte Carlo simulations, for $N = 10$. The Monte Carlo simulation used the Metropolis algorithm to generate a Markov chain of edge $SU(2)$ matrix configurations $\{U_e\}$ that are distributed according to the Boltzmann weights $\exp(-S_{\text{SGT}})$ that occurs in the path integral formulation of the field theory.

Each Monte Carlo step involves randomisation of some edge $SU(2)$ matrices. This is done by multiplication of a small $\mathfrak{su}(2)$ algebra matrix, together with a Metropolis step for acceptance/rejection of the update. The algorithm adapted itself to drive the Monte Carlo acceptance rate towards 1/2. The use of temporal gauge may slow the convergence of this type of numerical simulation, due to introduction of long-range effects. However, Monte Carlo convergence was ascertained and high quality statistical error estimates were made by the use of data blocking [35] independently for each observable. In addition, convergence was verified subjectively by inspection of the time series for observable values with their accompanying distributions, as well as time series for cumulative averages. The data blocking error estimates were found to be smaller than the displayed data points in all the plots.

We simulated at different values of β , at each of which we measured the average action density S/N^4 , and a list of different Wilson loops shown in figure 6.1, all of which are gauge-invariant quantities. For each Wilson loop shape, we average over all possible loop positions, as well as loop orientations in the xy , yz and zx planes. For a given closed path \mathcal{C} , the corresponding Wilson loop variable for gauge group $SU(n)$ is defined as

$$W_{\mathcal{C}} := \frac{1}{n} \Re \text{tr} \prod_{e \in \mathcal{C}} U_e, \quad (6.1)$$

which involves an ordered product of the edge matrices $\{U_e\}$ along the path \mathcal{C} .

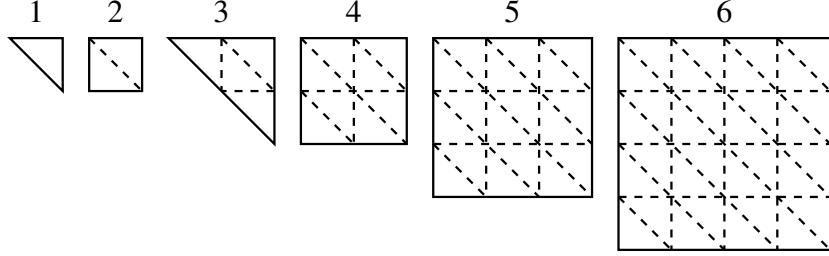


Figure 6.1: The simulated Wilson loop shapes correspond to the outer edges of these figures. They lie in the xy , yz and zx planes.

The lattice partition function for the SGT action is

$$Z(\beta) = \sum_{\{U_e\}_{e \in \mathcal{S}}} \exp(-S_{\text{SGT}}(\{U_e\})), \quad (6.2)$$

where the dependency on β is included in S_{SGT} . The relative contributions of each configuration of link variables $\{U_e\}$ to quantum expectation values is determined by the Boltzmann relative probability weights

$$w(\{U_e\}) = \exp(-S_{\text{SGT}}(\{U_e\})). \quad (6.3)$$

Expectation values for any observable quantity \mathcal{O} , e.g. the action density S/N^4 or a Wilson loop W_C , is given by

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \int \left(\prod_e dU_e \right) \mathcal{O} \exp(-S_{\text{SGT}}), \quad (6.4)$$

where the partition function Z is defined by

$$Z := \int \left(\prod_e dU_e \right) \exp(-S_{\text{SGT}}). \quad (6.5)$$

The integration measure involved in these expressions is a product of the normalised Haar integration measure for each edge group element in the mesh. Note that the normalised Haar measure satisfies

$$\int_{\mathcal{G}} dU = 1. \quad (6.6)$$

To accompany these measurements, we have calculated the strong (small β) and weak (large β) coupling asymptotic behaviour in [40, Appendix B], using methods described in [26]. At strong coupling, this involves various group integrals, while at weak coupling it suffices to use a simple thermodynamic analogy involving DOF counting to determine the limiting behaviour.

The simulated results for the action density and Wilson loops are displayed in figure 6.2. We can see the characteristic and nontrivial behaviour in the medium coupling range $\beta \in (1, 3)$. This coincides qualitatively with LGT simulations [25]. Only qualitative, not exact, agreement

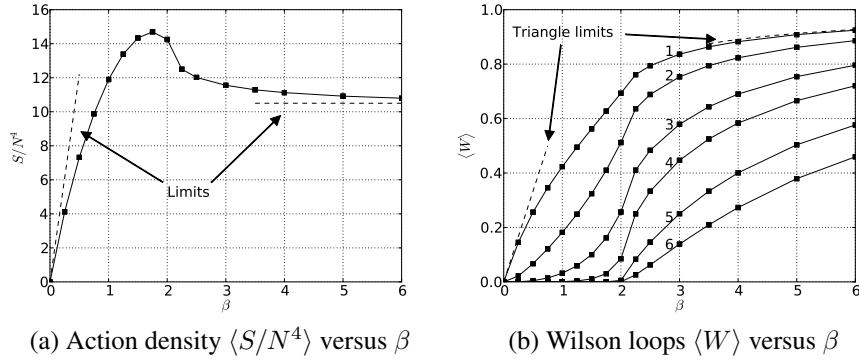


Figure 6.2: Plots showing the β -dependency of the average action density $\langle S/N^4 \rangle$ and the various Wilson loops $\langle W \rangle$ from figure 6.1. Solid squares are data points and solid lines are linear interpolations. The strong and weak coupling asymptotes are included for the action density and the elementary triangular loop. Monte Carlo errors are smaller than the data points.

is expected, since the physical lattice constant will differ in each type of simulation, at β away from the continuum limit. Compared to LGT simulations, the behaviour at small β deviates more from linearity due to the nonlinear aspects of the SGT action. In this region, the actions do not approximate the continuum action, and differences between discrete actions are nonphysical.

The Wilson loops show the same qualitative behaviour as do LGT simulation results, and approaches the calculated asymptotes nicely. Also here, the behaviour is less linear at small β for the same reason as stated above. The typical strong suppression of the Wilson loops as functions of loop area is reproduced, as expected from the area law behaviour that indicates confinement.

The smooth behaviour of these observables is an indication that no phase transition occur. However, for a definitive analysis we would need to define a proper order parameter.

6.5 Discussion

We have implemented the general SGT action on a particular simplicial mesh, and performed Monte Carlo quantum field theory simulations that show sensible results that are qualitatively consistent with standard LGT simulations, as must be the case for this initial proof of concept implementation.

We expect that this method will lend itself nicely to the use of mesh refinement within quantum QCD simulations, and that this will lead to opportunities of novel applications using nontrivial mesh structures, e.g. in the vicinity of gluon flux tubes as mentioned in the introduction.

The nondiagonal nature of the action increases the amount of computer work in the Metropolis step after each proposed update. However, since the number of interactions for each elementary face is finite, the scaling at large meshes for this model will be the same as for traditional QCD. There might be possibilities of real-time adaptive diagonalisation, thereby increasing the algorithm efficiency throughout the initial part of the simulation.

6.6 Program details

Our computer implementation of the simplicial lattice and accompanying SGT action consists of object-oriented C++ code, using MPICH2 [1] for parallelisation, running on a quadruple CPU run-of-the-mill modern workstation computer. The data structures involved are reminiscent of what is used in implementations of the finite element method. This involves different types of mass matrix and connectivity information for elements of the simplicial mesh. The parallelisation consisted of running independent simulations on each node, and averaging the results. We used the `yarn2` algorithm from the TINA pseudo-random number generator [8], which is designed for use in parallel algorithms. Although the edge matrix randomisation appeared to perform stably enough for our purposes, we regularly did projections of the edge matrices onto $SU(2)$ as a precautionary measure.

The Monte Carlo sampling consisted of randomly updating $SU(2)$ edge elements and then using the change in the probability weight 6.3 in a Metropolis acceptance/rejectance test. Link matrix updates were done by multiplication with random $SU(2)$ elements within some adjustable maximal distance L from the identity matrix. The parameter L was adaptively adjusted to drive the Metropolis acceptance rate towards 0.5. Periodic sanity tests were performed to ensure that matrices did not deviate from $SU(2)$.

The parallelisation consisted of performing independent Monte Carlo simulations on each node using MPICH2 [1], and averaging over nodes at the end. Computations were done on consumer grade 64-bit dual-core workstation computers.

As this is only meant as a concrete proof of concept for the possibility of lattice gauge theory simulations using the simplicial gauge theory action, there is lots of room left for optimisation of the algorithms.

Chapter 7

Stochastic Gross-Pitaevskii equation

7.1 Introduction

The well-known Gross-Pitaevskii (GP) equation is a nonlinear Schrödinger equation for the single-particle wavefunctions of each single constituent particle of a Bose-Einstein condensate (BEC), as discussed in section ???. It appears in an approximation of a many-particle system of bosons at very low temperature, so dilute as to justify a point interaction between constituent particles. The complete quantum state of the system is obtained as a product of the single-particle wavefunctions.

A BEC arises due to quantum physics and the bosonic nature of the constituent particles. However, those particles may be bosonic bound states of several bosonic or fermionic particles. For example, Helium-4 atoms are bosons since they are composed of six fermions, and can therefore form a BEC leading to superfluidity. In addition, superconductivity in certain metals occurs because pairs of valence electrons form bosonic bound states, which can act as bosons at low temperature. In both these cases, the experimental phenomenon that arises shows signs of collective (or cooperative) behaviour between a macroscopic fraction of the particles, which in turn causes a reduction in dissipative losses (reduced viscosity and resistance in these examples).

By a careful choice of units and normalisation of the wavefunction, the Gross-Pitaevskii equation can be cast into the form

$$i\dot{\psi} = -\Delta\psi + V\psi + |\psi|^2\psi, \quad \psi : \mathbb{R}^+ \times \mathbb{R}^3 \rightarrow \mathbb{C}, \quad (7.1)$$

where ψ represents the wavefunction of a constituent bosonic particle, and V represents an external potential, which might be confining for the case of a BEC in a potential trap. The total wavefunction of the many-particle system is given by a product of such states.

The nonlinearity is defocusing due to the positive sign. This can be seen intuitively by considering the decomposition $\psi = r \exp(i\phi)$ into two real functions r and ϕ , and then study an infinitesimal time evolution from a sharply localised Gaussian initial condition with $\phi = 0$. The defocusing property means that the wavefunction will tend to be spread out from such configurations.

However, it is more realistic to consider a BEC that interacts with a sea of non-BEC particles, since not all constituent particles in a many-particle system will occupy the lowest energy state, and the temperature in real experiments are nonzero. A possibility is to model such addi-

tional particles as a thermal reservoir. The influence of a thermal reservoir upon the rest of the system is by definition stochastic. This can be modelled by additional stochastic noise terms in the wave equation. We end up with a stochastic PDE (SPDE) called the stochastic GP equation, which involves both additive and multiplicative noise [37, 38]. This equation is quite complicated, containing spacetime white noise fields. Unfortunately, spacetime white noise does not currently lend itself to a rigorous mathematical treatment. Therefore it is necessary to consider spatially correlated noise as an approximation. Also, we expect that the white noise is a mathematical idealisation that arises in the model only because other physical phenomena have been neglected.

In our study of existence proofs for stochastic Gross-Pitaevskii equations, we have come across a proposed proof [27, Theorem 4.1] of the stochastic GP equation with multiplicative noise

$$i\dot{\psi} = \Delta\psi + |\psi|^{2\sigma}\psi + \dot{\eta}\psi, \quad (7.2)$$

where the noise term $\dot{\eta}$ is real-valued. This is necessary for a unitary time evolution. The value $\sigma = 1$ corresponds to the stochastic GP equation. Note that the sign on the Laplacian is different, but this is simply a matter of considering the complex conjugate of the wavefunction. The proof is performed by considering the weak formulation of the SPDE, which means to consider the time-integral of the equation, and searching for solutions in a space of functions with weak derivatives 2.9. An attempt is made to prove that this weak formulation gives rise to a contraction mapping on the solution function space under consideration. Then, Banach's fixed point theorem is invoked to prove the existence and uniqueness of the solution. Needless to say, a multitude of estimates are needed, using methods from stochastic and functional analysis, for example Hölder's inequality and Sobolev embeddings.

Specifically, the proof depends on the Sobolev embedding $H^1(\mathbb{R}^n, \mathbb{C}) \subset L^q(\mathbb{R}^n, \mathbb{C})$ [33], where n is the number of space dimensions and q is a real parameter that arises in the proof. For consistency with the notation in the published proof, we use n instead of d for the number of spatial dimensions. For the Sobolev embedding to be applicable, there are restrictions on the parameter q , depending on the space dimension d . Due to an unfortunate mistake in the proof, the Sobolev embedding is used without justification in two important places. There is no need for us to describe more details from the proof. It suffices to demonstrate the problems in question.

For definitions of the function spaces involved, please read section 2.9. The Sobolev embedding theorem says that [33]

$$H^1(\mathbb{R}^n, \mathbb{C}) \subset L^q(\mathbb{R}^n, \mathbb{C}), \quad \text{for } q = \frac{2n}{n-2}. \quad (7.3)$$

7.2 Regarding the proof of Theorem 4.1

Consider [27, Theorem 4.1]. The following n -dependent parameter ranges for σ are assumed,

$$\begin{cases} 0 < \sigma & , n = 1, 2 \\ 0 < \sigma < 2 & , n = 3 \\ \frac{1}{2} \leq \sigma < \frac{2}{n-2} \quad \text{or} \quad \sigma < \frac{1}{n-1} & , n \geq 4 . \end{cases} \quad (7.4)$$

The theorem essentially states that an *admissible pair* of Lebesgue space exponents (r, p) exists such that a stochastic nonlinear Schrödinger equation has a unique solution in a certain function space characterised by (r, p) . Admissibility is defined as

$$r \geq 2, \quad \frac{2}{r} = n \left(\frac{1}{2} - \frac{1}{p} \right) . \quad (7.5)$$

Dual Lebesgue space exponents are denoted by primed quantities, and are defined by the equation

$$\frac{1}{p} + \frac{1}{p'} = 1 . \quad (7.6)$$

For completeness, the exact function solution space in question is

$$C([0, \tau]; H^1(\mathbb{R}^n)) \cap L^r([0, \tau], W^{1,p}(\mathbb{R}^n)) , \quad (7.7)$$

for some time $\tau > 0$ that is given elsewhere. The proof given is for the special case $\sigma \geq 1/2$.

In the proof, a second admissible pair (γ, s) is introduced after [27, Eq.(4.17)], for which s' satisfies the inequality

$$\frac{2\sigma + 1}{p} - \frac{1}{n} < \frac{1}{s'} < \frac{n+2}{2n} . \quad (7.8)$$

The parameters s' and p are related through another parameter q which arises in the proof,

$$\frac{1}{s'} = \frac{2\sigma}{q} + \frac{1}{p} , \quad (7.9)$$

which is described prior to [27, Eq.(4.18)]. These parameters arise due to the use of Hölder's inequality.

The next step is to use the Sobolev embedding $H^1(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$. It is claimed that the embedding holds because " $q < 2n/(n-3) < 2n/(n-2)$ ". The first part is correct

$$q < \frac{2n}{n-3} ,$$

which follows easily from the left hand side of (7.8) and the left hand side of [27, Eq. 4.14]. However, the second part of the claim is incorrect since $1/(n-3) > 1/(n-2)$, and therefore the Sobolev embedding $H^1(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$ is used without proper justification.

7.3 Regarding the proof of Lemma 4.3

In the proof of [27, Lemma 4.3], in the second estimate on p.121, the interpolation inequality for L^p -spaces, followed by the Sobolev embedding $H^1(\mathbb{R}^n), W^{1,p}(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$ were used, where the parameter q was “as above”. Therefore, for the same reason, the Sobolev embedding used here also lack a proper justification.

7.4 Discussion

We have demonstrated a problem in an existence proof for an SPDE that is closely related to the stochastic Gross-Pitaevskii equations that governs BECs under the influence of a thermal reservoir. It would be extremely pleasing to provide a complete correction of the proof, instead of simply demonstrating a mistake. However, this must, if it is tractable, unfortunately be delayed to a later occasion.

Bibliography

- [1] Mpich2.
- [2] *Whitney forms: a class of finite elements for three-dimensional computations in electromagnetism*, volume 135A, 1988.
- [3] Mark G. Alford, W. Dimm, G. P. Lepage, G. Hockney, and P. B. Mackenzie. Lattice QCD on small computers. *Phys. Lett.*, B361:87–94, 1995.
- [4] M. P. Allen and T. J. Tildesley. *Computer Simulation of Liquids*. Clarendon Press, 1991.
- [5] R. W. B. Ardill, J. P. Clarke, J. M. Drouffe, and K. J. M. Moriarty. Quantum chromodynamics on a simplicial lattice. *Phys. Lett.*, B128:203, 1983.
- [6] Søren Asmussen and Jan Rosinski. Approximations of small jumps of Levy processes with a view towards simulation. *J. Appl. Probab.*, 38(2):482–493, 2001.
- [7] R. A. Aziz, V. P. S. Nain, J. S. Carley, W. L. Taylor, and G. T. McConville. An accurate intermolecular potential for Helium. *The Journal of Chemical Physics*, 70(9):4330–4342, 1979.
- [8] Heiko Bauke. TINA pseudo-RNG library.
- [9] Carl M. Bender and Kimball A. Milton. Approximate determination of the mass gap in quantum field theory using the method of finite elements. *Phys. Rev. D*, 34(10):3149–3155, Nov 1986.
- [10] Carl M. Bender, Kimball A. Milton, and David H. Sharp. Gauge invariance and the finite-element solution of the Schwinger model. *Phys. Rev. D*, 31(2):383–388, Jan 1985.
- [11] M. Boninsegni, N. V. Prokof’ev, and B. V. Svistunov. Worm algorithm and diagrammatic Monte Carlo: A new approach to continuous-space path integral Monte Carlo simulations. *Phys. Rev. E*, 74(3):036701, Sep 2006.
- [12] Massimo Boninsegni, Nikolay Prokof’ev, and Boris Svistunov. Worm algorithm for continuous-space path integral Monte Carlo simulations. *Phys. Rev. Lett.*, 96(7):070601, Feb 2006.
- [13] Kevin E. Cahill and Randolph Reeder. Comparison of the simplicial method with Wilson’s lattice gauge theory for U(1) in three-dimensions. *Phys. Lett.*, B168:381, 1986.

- [14] P. Carr, H. Geman, D. B. Madan, and M. Yor. Stochastic volatility for Levy processes. *Mathematical Finance*, 13(3):345–382, JUL 2003.
- [15] D. M. Ceperley. Path integrals in the theory of condensed Helium. *Rev. Mod. Phys.*, 67(2):279–355, Apr 1995.
- [16] Masud Chaichian and Andrei Demichev. *Path Integrals in Physics Vol.1: Stochastic Processes and Quantum Mechanics*. Institute of Physics Publishing, 2001.
- [17] Masud Chaichian and Andrei Demichev. *Path Integrals in Physics Vol.2: Quantum Field Theory, Statistical Physics and other Modern Applications*. Institute of Physics Publishing, 2001.
- [18] N. H. Christ, R. Friedberg, and T. D. Lee. Gauge theory on a random lattice. *Nucl. Phys.*, B210:310, 1982.
- [19] N. H. Christ, R. Friedberg, and T. D. Lee. Random lattice field theory: General formulation. *Nucl. Phys.*, B202:89, 1982.
- [20] N. H. Christ, R. Friedberg, and T. D. Lee. Weights of links and plaquettes in a random lattice. *Nucl. Phys.*, B210:337, 1982.
- [21] Snorre H. Christiansen and Tore G. Halvorsen. A simplicial gauge theory. *ArXiv e-prints*, June 2010.
- [22] Snorre H. Christiansen, Hans Z. Munthe-Kaas, and Brynjulf Owren. Topics in structure-preserving discretization. *Acta Numerica*, 20:1–119, 2011.
- [23] Philippe G. Ciarlet. *The Finite Element Method for Elliptic Problems*, volume 4 of *Studies in mathematics and its applications*. North-Holland Publishing Company, 1. edition, 1978.
- [24] Rama Cont and Peter Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall/CRC Financial Mathematics Series, 2003.
- [25] Michael Creutz. Monte Carlo study of quantized SU(2) gauge theory. *Phys. Rev.*, D21:2308–2315, 1980.
- [26] Michael Creutz. *Quarks, Gluons and Lattices*. Cambridge, Uk: Univ. Pr. (Cambridge Monographs On Mathematical Physics), 1986.
- [27] A. de Bouard and A. Debussche. The stochastic nonlinear Schrödinger equation in H^1 . *Stochastic Analysis and Applications*, 21(1):97–126, JAN 2003.
- [28] J. M. Drouffe and K. J. M. Moriarty. Gauge theories on a simplicial lattice. *Nucl. Phys.*, B220:253–268, 1983.
- [29] J. M. Drouffe and K. J. M. Moriarty. High-statistics study of the phase transition in U(2) four-dimensional simplicial lattice gauge theory. *Journal of Physics G: Nuclear Physics*, 10(10):L221, 1984.

- [30] J. M. Drouffe and K. J. M. Moriarty. U(2) four-dimensional simplicial lattice gauge theory. *Z. Phys.*, C24:395, 1984.
- [31] J. M. Drouffe, K. J. M. Moriarty, and C. N. Mouhas. Monte Carlo simulation of pure U(N) and SU(N) gauge theories on a simplicial lattice. *Comput. Phys. Commun.*, 30:249, 1983.
- [32] J. M. Drouffe, K. J. M. Moriarty, and C. N. Mouhas. U(1) four-dimensional gauge theory on a simplicial lattice. *J. Phys.*, G10:115, 1984.
- [33] Lawrence C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. Americal Mathematical Society, 2002.
- [34] Richard P. Feynman. Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.*, 20(2):367–387, Apr 1948.
- [35] H. Flyvbjerg and H.G. Petersen. Error estimates on averages of correlated data. *Journal of Chemical Physics*, 91(1):461–466, 1989.
- [36] Piotr Garbaczewski, John R. Klauder, and Robert Olkiewicz. Schrödinger problem, lévy processes, and noise in relativistic quantum mechanics. *Phys. Rev. E*, 51:4114–4131, May 1995.
- [37] C. W. Gardiner, J. R. Anglin, and T. I. A. Fudge. The stochastic Gross-Pitaevskii equation. *JOURNAL OF PHYSICS B-ATOMIC MOLECULAR AND OPTICAL PHYSICS*, 35(6):1555–1582, MAR 28 2002.
- [38] C. W. Gardiner and M. J. Davis. The stochastic Gross-Pitaevskii equation: II. *Journal of Physics B Atomic Molecular Physics*, 36:4731–4753, December 2003.
- [39] Tore Gunnar Halvorsen and Torquil Macdonald Sørensen. Simplicial gauge theory on spacetime. 2011.
- [40] Tore Gunnar Halvorsen and Torquil Macdonald Sørensen. Simplicial gauge theory and quantum gauge theory simulation. *Nuclear Physics B*, 854(1):166 – 183, 2012.
- [41] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
- [42] Ralf Hiptmair. Finite elements in computational electromagnetism. *Acta Numerica*, 11(-1):237–339, 2002.
- [43] A. D. Klemm and R. G. Storer. Structure of quantum fluids - Helium and Neon. *Australian Journal of Physics*, 26(1):43–59, 1973.
- [44] L. Koci, R. Ahuja, A. B. Belonoshko, and B. Johansson. Study of the high-pressure Helium phase diagram using molecular dynamics. *Journal of Physics: Condensed Matter*, 19(1):016206 (9pp), 2007.
- [45] L. D. Landau and E. M. Lifshitz. *Statistical Physics*, volume 5 of *Course of Theoretical Physics*. Butterworth-Heinemann, 3rd edition, 1997.

- [46] G. P. Lepage. New algorithm for adaptive multidimensional integration. *Journal of Computational Physics*, 27(2):192–203, 1978.
- [47] R. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.
- [48] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953.
- [49] Peter Monk. *Finite Element Methods for Maxwell's Equations*. Oxford Science Publications, reprinted edition, 2006.
- [50] Jean-Claude Nédélec. Mixed finite elements in \mathbb{R}^3 . *Num. Math.*, 35:315–341, 1980.
- [51] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995.
- [52] Gert Roepstorff. *Path Integral Approach to Quantum Physics*. Texts and Monographs in Physics. Springer-Verlag, 1994.
- [53] R. G. Storer. Path-integral calculation of the quantum-statistical density matrix for attractive Coulomb forces. *Journal of Mathematical Physics*, 9(6):964–&, 1968.
- [54] Alastair J. Walker. An efficient method for generating discrete random variables with general distributions. *ACM Trans on Mathematical Software*, 3:253–256, 1977.
- [55] Steven Weinberg. *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge, UK: Univ. Pr., 1995.
- [56] Steven Weinberg. *The quantum theory of fields. Vol. 2: Modern applications*. Cambridge, UK: Univ. Pr., 1996.
- [57] Hassler Whitney. *Geometric integration theory*. Princeton University Press, 1957.
- [58] Norbert Wiener. Differential space. *J. Math. and Phys. Sci.*, 2:132, 1923.
- [59] Kenneth G. Wilson. Confinement of quarks. *Phys. Rev. D*, 10(8):2445–2459, Oct 1974.
- [60] Chen-Ning Yang and Robert L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 96:191–195, 1954.
- [61] Bernt Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Universitext. Springer, 6th edition, 2005.

Part II

Papers

Chapter 8

Levy process simulation by stochastic step functions

Submitted in 2011 to “SIAM Journal on Scientific Computing”. Awaiting referee report.

Levy process simulation by stochastic step functions

Torquil Macdonald Sørensen* Fred Espen Benth†

Centre of Mathematics for Applications
University of Oslo, NO-0316 Oslo, Norway

Abstract

We study a Monte Carlo algorithm for simulation of probability distributions based on stochastic step functions, and compare to the traditional Metropolis/Hastings method. Unlike the latter, the step function algorithm can produce an uncorrelated Markov chain. We apply this method to the simulation of Levy processes, for which simulation of uncorrelated jumps are essential.

We perform numerical tests consisting of simulation from probability distributions, as well as simulation of Levy process paths. The Levy processes include a jump-diffusion with a Gaussian Levy measure, as well as jump-diffusion approximations of the infinite activity NIG and CGMY processes.

To increase efficiency of the step function method, and to decrease correlations in the Metropolis/Hastings method, we introduce adaptive hybrid algorithms which employ uncorrelated draws from an adaptive discrete distribution defined on a space of subdivisions of the Levy measure space.

The nonzero correlations in Metropolis/Hastings simulations result in heavy tails for the Levy process distribution at any fixed time. This problem is eliminated in the step function approach. In each case of the Gaussian, NIG and CGMY processes, we compare the distribution at $t = 1$ with exact results and note the superiority of the step function approach.

Keywords: Levy process, Markov chains, Monte Carlo methods,
Simulation of Probability Distributions.

MSC2010: 60G51, 65C05, 60G17, 62M10, 60J22, 60J75

*Email: t.m.sorensen@matnat.uio.no, torquil@gmail.com

†Email: fredb@math.uio.no

1 Introduction

Levy processes are a type of stochastic process whose paths can behave erratically like a Brownian motion, as well as include discontinuities, i.e. jumps. Many observed phenomena in nature and human society display continuous erratic motion similar to Brownian motion, while also having random jumps at random times, and can therefore be modelled using Levy processes. Areas of application include e.g. microscopic physics, chemistry, biology and financial markets. Therefore the study of computer simulation methods for Levy processes is an important subject.

In our case, we will employ a jump-diffusion approximation when we apply our method to infinite activity Levy processes. Jump-diffusion processes can be considered as a sum of three independent component processes. A deterministic drift, a Brownian motion, and lastly a finite activity jump process. The jump process is completely described by a Levy measure ν on \mathbb{R} . The average jump rate, or intensity, is given by the total weight $\lambda := \nu(\mathbb{R})$, and the distribution of jump values follows the Levy probability measure $\nu_1 := \nu/\lambda$.

By definition, the Brownian motion has independent increments, and this is also the case for subsequent jumps in the jump process. In computer simulations, violation of these properties will give incorrect results. Simulation of the Brownian motion is easy, since well known algorithms exist for producing uncorrelated draws from a normal distribution, using a good pseudo-random number generator (PRNG).

On the other hand, the simulation of the jump process requires more care. In principle we can obtain uncorrelated draws from ν_1 by inverting its cumulative distribution function. However, this might not be feasible to do for a given ν_1 . In this case, the Metropolis/Hastings (MH) algorithm might come to the rescue. It can easily produce a Markov chain with values distributed according to ν_1 . However, subsequent values will be correlated, as described below.

In this paper we propose an algorithm to produce uncorrelated jumps along each path, without generating such a multitude of Markov chains. This method is based on stochastic step functions (SF), which will be defined below. As opposed to MH, it is not an accept/reject algorithm. Therefore it is able to generate an uncorrelated chain of values distributed according to any given probability measure. We will test this algorithm in jump-diffusion computer simulations and compare with MH. We apply the simulation techniques to a Gaussian process, for which the exact distribution at $t = 1$ is known. Convergence towards this exact result is studied.

We also consider adaptive variants of these algorithms. For these, we subdivide \mathbb{R} into appropriate regions, and generate a discrete distribution that allows us to efficiently draw among these regions in an uncorrelated way. As the simulation progresses, this discrete distribution is adaptively improved. The calculation of the jump rate λ is also adaptively improved.

As an interesting application of these simulation techniques, we also look at infinite activity pure jump processes, which are also a Levy process subclass. Here, motion occurs in the form of an infinitude of discontinuous jumps. Some

of these processes can be approximated by jump-diffusion by substituting the smallest jumps for an appropriate Brownian motion [1]. The examples we focus on are the NIG and CGMY processes.

In section 2 we review the elementary facts about jump-diffusion processes. In section 3 we describe the relevant simulation methods, and discuss their pros and cons, as well as provide results from numerical experiments. In section 4 we describe our probability space subdivision method and the accompanying adaptability properties of the algorithms, and study the efficiency and correlation strengths of different algorithms by simulation from a Gaussian distribution. We then proceed in section 5 to perform simulations of jump-diffusion processes. We notice how the MH correlations adversely affect the distribution of the simulated process, and compare the SF and MH algorithms for simulation of a process with a Gaussian Levy measure. Lastly, in section 6, we review the jump-diffusion approximation of infinite activity pure jump processes, and apply our simulations techniques on the infinite activity NIG and CGMY processes, in order to compare algorithms in these interesting cases.

2 Jump-diffusion processes

First we review some elementary facts about real-valued jump-diffusion Levy processes on a time interval $[0, T]$. Such a process can be expressed as a sum of three simple components

$$L_t = \mu t + B_t + J_t. \quad (1)$$

The first part is a deterministic drift with rate μ , B_t is Brownian motion, and J_t is a compound Poisson process. The latter describes completely the discontinuities (jumps) of the paths of L_t , by means of a *Levy measure* ν on \mathbb{R} . We will several times abuse notation by writing ν both for the Levy measure and its density. Firstly, this measure determines the jump intensity (jump rate)

$$\lambda := \nu(\mathbb{R}) < \infty, \quad (2)$$

of L_t . Secondly, the corresponding *Levy probability measure*

$$\nu_1 := \nu/\lambda, \quad (3)$$

determines the distribution of jump sizes. Note that we do not have $E[L_t] = \mu t$ in general, although this is satisfied in our examples.

Now, J_t can be expressed as

$$J_t = \sum_{j=1}^{N_t} V_j, \quad (4)$$

where N_t is a Poisson process of intensity λ , and the jumps $\{V_j\}$ are independent random variables distributed according to the Levy probability measure. We do not simulate the Poisson process N_t directly. Instead, for each path we draw

the total number of jumps individually from an exponential distribution with average λT . Then we randomly distribute those jumps on $[0, T]$. We will choose $T = 1$ for our simulations. The lengths of consecutive time intervals between jumps will be independent and exponentially distributed, and give the correct jump intensity. This is described in [3].

Note that for a general Levy process λ is in general not finite, because the Levy measure might diverge too strongly towards the origin. However, the following condition always holds,

$$\int_{\mathbb{R}} \min(1, s^2) \nu(ds) < \infty. \quad (5)$$

which restricts the strength on the divergence of ν at the origin.

The difficulty in simulation is to draw independent jump values from ν_1 . We will generate a Markov chain with ν_1 as its invariant measure. However, it is essential for correct simulation that jumps along each path are uncorrelated. Note that existence of correlations between jumps in *different* paths will not ruin the convergence in distribution, but only slow it down.

For a Markov chain $\{X_i\}$, we define the sequential correlation as

$$c := E[(X_{i+1} - \bar{X})(X_i - \bar{X})]/\sigma^2, \quad (6)$$

where \bar{X} is the Markov chain average and σ^2 is its standard deviation,

$$\bar{X} := E[X], \quad \sigma^2 := E[X^2 - \bar{X}^2].$$

3 Simulation of a probability measure

To simulate the jumps, we must draw independent values from ν_1 . In cases where ν_1 is complicated, it is common to use the Metropolis/Hastings (MH) algorithm [7, 4]. This method generates a Markov chain with ν_1 as its invariant density. Unfortunately, the Markov chain often has large correlations between successive values. Successive values in such a chain cannot be used to represent jumps within a single jump-diffusion path.

It is possible to reduce this problem by several methods. One is to skip a number of terms in the Markov chain to reduce correlation. This results in a loss of efficiency. Another method is to generate multiple independent Markov chains. Each jump-diffusion path then uses values from different Markov chains. This will lead to a correct pathwise simulation, and therefore correct convergence in distribution. The correlations will in this case only slow down the convergence.

We will look at two methods, MH and one based on stochastic step functions (SF). Both rely on a transition probability distribution ρ to determine upcoming values from the previous one. We say that we are dealing with a *local* algorithm if ρ is localized around the current value. An *independent* algorithm results if ρ is independent of the current position.

Use of an independent MH algorithm can reduce correlations dramatically if one is able to find a ρ that approximates ν_1 . The downside is that efficiency will suffer if ρ is not computationally simple. We will focus on generation of low correlation Markov chains in order to get by using only one chain for the Levy process jumps.

3.1 Metropolis/Hastings (MH)

The well-known MH algorithm generates a realisation $\{x_i\}$ of a Markov chain distributed according to ν_1 . Its most popular incarnation is local, where the transition probability density $\rho(x_i, x_{i+1})$ involved in each transition $x_i \mapsto x_{i+1}$ is localized around the value x_i , and its width is adjusted to give an average acceptance rate somewhere around 1/2. The resulting correlation between successive values can be classified into two types:

- *Rejection correlation*: Since both the local and independent algorithms are based on acceptance/rejectance, correlations are introduced by repeated values due to rejections. With an acceptance rate around 1/2, repeated values will often occur.
- *Transition correlation*: In order for the local algorithm to be efficient, its transition probability measure $\rho(x_i, x_{i+1})$ must in many cases be quite strongly localized; otherwise too small an acceptance rate will result. Thus the difference between subsequent variates of the Markov chain will tend to be small.

To reduce transition correlation, the width of the transition distribution can be increased. But this typically leads to a reduced acceptance ratio, which gives an increased rejection correlation, and vice versa.

As a numerical example, consider a unit variance normal distribution, simulated with a simple localized uniform transition probability measure. The correlation defined in (6) as a function of transition probability width is displayed in Figure 1. As expected, the correlation has a minimum. Towards smaller widths the transition correlation increases, and towards larger widths the rejection correlation increases. Thus there is a lower bound on the amount of correlation for this algorithm, and it seems that the local MH algorithm is therefore not suited for our purposes. We will from now on focus on the independent MH algorithm.

3.2 Stochastic step function (SF)

Next, we propose a simulation algorithm based on stochastic step functions, that can be adjusted to completely eliminate correlations. This possibility of vanishing correlations is an attractive property that allows it to be used as a reference algorithm. It can also be adjusted to run more efficiently, with nonzero correlation.

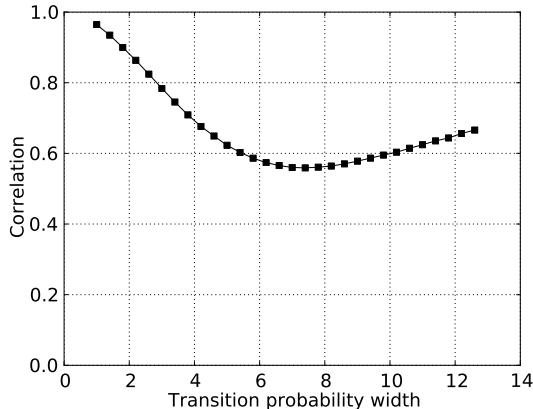


Figure 1: Correlation c versus transition probability width w for the local MH algorithm, in a simulation of the unit normal distribution using a uniform centered transition probability distribution. Correlation is minimal around $w \approx 7$, where $c \approx 0.55$.

Let us give a pathwise definition of the SF process on $[0, \infty)$ for some measure density ν on $\Omega \subset \mathbb{R}$, not necessarily normalized. Consider a sequence $\{\tau_i\}_{i=0}^\infty \subset (0, \infty)$ of *resting times* and corresponding *jump times* $\{t_i\}_{i=0}^\infty$, defined recursively by

$$t_0 = 0, \quad t_{i+1} := t_i + \tau_i.$$

In addition, let $\{s_i\}_{i=0}^\infty \subset \Omega$ be a Markov chain with initial value $s_0 = 0$ and transition probability density $\rho(s_i, s_{i+1})$. Assume that $\rho(s_i, \cdot)$ is absolutely continuous with respect to the Lebesgue measure on \mathbb{R} , and homogeneous, i.e. it can be expressed as $\rho(s_{i+1} - s_i)$.

From this data, we are now ready to define our stochastic step function process X_t , by defining its paths as the piecewise constant *cadlag* functions (right-continuous with left limits) of the form

$$X_t = \sum_{i=1}^{\infty} s_i \chi_{I_i}(t), \quad I_i := [t_i, t_{i+1}),$$

where χ_I is the indicator function for the interval I , and where the resting times are given by $\tau_i := \nu(s_i)$.

We see that the paths of X_t linger for a some time at each of its attained positions, with a resting time equal to the local value of the density ν . Consider the graph of such a path on $[0, T]$ for large T compared to $\sup \nu$. As T increases, the relative path length within a given subset $A \subset \Omega$ on the vertical axis converges towards $\nu(A)/\nu(\Omega) =: \nu_1(A)$. When sampling this path on a uniform time-grid, the obtained values will therefore be distributed according to the probability density ν_1 .

As in the case of local MH, we get transition probability correlations if ρ is localized. In this algorithm, however, there is no accept/reject step, and therefore no rejection correlation. As an example, let us choose ρ to be the uniform probability distribution on Ω . As noted above, if we sample the step function path generated above on a uniform time grid, we get a chain of values distributed according to ν_1 . If we make sure that the time discretization interval size Δt is larger than $\sup \nu$, repeated values will never occur and there are no correlations.

Note that as in MH algorithms, we do not need to know scale factor relating ν_1 and ν . If $\sup \nu$ is initially unknown, we can simply start with any estimate, and improve it adaptively as the algorithm runs.

It is easy to see that the amount of computer time spent between each discrete time value is proportional to $\sup \nu / \nu(\Omega)$ in its correlation-free mode. If this ratio is large, the algorithm will be inefficient.

One can make the time discretization finer to increase the rate of variate generation, but this introduces correlations since values for which ν is large will be repeated with complete certainty. This is similar to the Metropolis algorithm, where values of large ν are repeated (although not with complete certainty). The SF algorithm can be made less deterministic in this case by letting the resting times be random variables distributed according to an exponential distribution with mean $\nu(x)$, where x is the current position.

Since we are concerned with minimizing correlations in the context of jump-diffusion simulations, we will use independent algorithms, where $\rho(x_i, x_{i+1})$ is independent of x_i .

3.3 Simulation comparison

Let us illustrate the advantages of the local SF algorithm over local MH by considering an example with a probability distribution with several modes on a sample space $\Omega = [0, 1]$. Our goal is to simulate values from a probability distribution proportional to the following unnormalized density with two strong modes,

$$\nu(x) = \begin{cases} 1 & , x \in [0, 0.25] \cup [0.5, 0.75] \\ 0.01 & , \text{otherwise.} \end{cases} \quad (7)$$

For both MH and SF, we used a localized uniform transition probability density ρ of width 1/2. For MH, this gave an acceptance rate of approximatively 0.55. The SF Markov chain realisation was obtained by sampling the stochastic step function using a lattice spacing slightly larger than $\sup \nu$ in order to avoid any repeating values.

The results are shown in Figure 2. One sees immediately that the MH algorithm has the potential of getting stuck inside a mode for long periods. This is caused by rejection correlation as found in accept/reject algorithms such as MH. On the other hand, the SF algorithm has no such problems since it lacks such correlations. This serves to illustrate a problem that often can occur with MH. The SF algorithm is a clear favourite in this case if mixing is important for the application of these Markov chains.

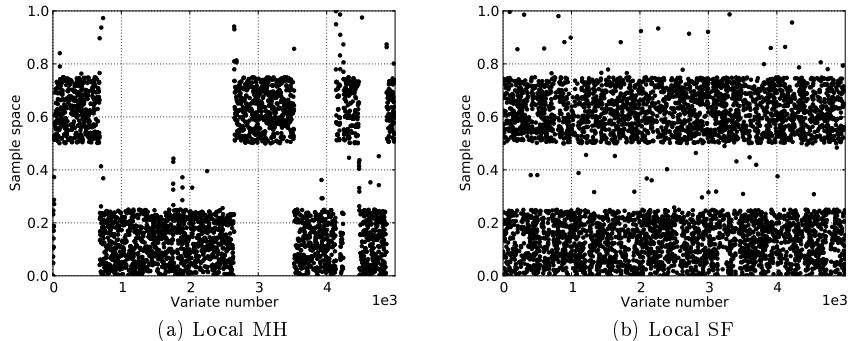


Figure 2: Markov chain realisations using local MH and local SF. The SF algorithm mixes much better than MH.

4 Probability space subdivision and adaptability

The Levy probability measure from which we will simulate in the context of jump-diffusions will often be defined on the unbounded probability space $\Omega = \mathbb{R}$. This presents no difficulty for the MH algorithm. SF algorithms on the other hand need to impose upper and lower cutoffs on Ω , in order to avoid step functions that diverge towards infinity. For simplicity, we impose such cutoffs on both algorithms in our examples. Alternatively, one could perform a topology-changing coordinate transformation on Ω to obtain a compact space, however we will not do this in our examples. From now on we therefore assume $\Omega \subset \mathbb{R}$ to be a bounded interval which we choose symmetrically about the origin. We will only deal with symmetrical Levy measures in our examples.

In order to reduce correlations in the MH algorithms, and increase efficiency in the SF algorithms, we define a finite disjoint subdivision $\{U_i\}$ of Ω , where $\cup_i U_i = \Omega$. We construct a discrete distribution $\tilde{\nu}$ on the finite set $\{U_i\}$. For each $U_i \subset \Omega$, the value $\tilde{\nu}(U_i)$ must approximate $\nu(U_i)$, i.e. the Levy measure weight of U_i . The simulation algorithm starts with a preliminary estimate of $\tilde{\nu}$ which is adaptively improved throughout the simulation, as explained below in the two cases of MH and SF. This is reminiscent of variance reduction schemes used in Monte Carlo integration, such as stratification and the VEGAS algorithm [5]. We now describe in more details how this is done in each case.

4.1 Adaptive Independent MH (AIMH)

As explained above, we will use an independent MH algorithm. The independent transition probability ρ is defined as follows. First, use the discrete unnormalized probability measure $\tilde{\nu}$ to draw a subset U_i . Then a random position within this subdomain is proposed and the value of ν at this position is calculated. Thereafter follows the usual MH accept/reject step.

The initial draw of U_i is done without introducing correlations, using e.g. an efficient algorithm which is independent of the normalization of $\tilde{\nu}$ [8]. The registered values of ν are accumulated, and used periodically in the simulation to improve $\tilde{\nu}$. Essentially, a separate Monte Carlo simulation is being performed within each subdomain U_i to improve the discrete distribution $\tilde{\nu}$, while the algorithm proceeds to generate new draws from ν .

Since subdomains U of Ω are drawn by means of $\tilde{\nu}$ without introducing correlations, the amount of correlation generated by the algorithm as a whole is reduced. Transition correlation is completely eliminated since the algorithm is independent. It is impossible to completely eliminate rejection correlation in an MH algorithm unless ρ is identical to ν_1 . However, since ν is well approximated on each subdomain (as long as the subdivision is sufficiently fine), these will be small, and the acceptance rate will be high. As $\tilde{\nu}$ adapts, the acceptance rate rises and correlation decreases. The subdivision discretization implies a lower bound for the amount of correlation. But this is still a big improvement on a basic local MH algorithm for the cases we consider. Note that the subdivision cannot be made arbitrarily fine, since the discrete algorithm will then become inefficient.

4.2 Adaptive Independent SF (AISF)

It is possible to improve the efficiency of the SF algorithm by a similar construction, without introducing correlations. In this case, the SF algorithm proceeds as follows. First draw a subdomain U_i using $\tilde{\nu}$. Draw a random position x within this subdomain, and record the value $\nu(x)$ for later use to improve $\tilde{\nu}$. For each subset U_i we keep an estimate of $\sup_i \nu$ which is updated for each sample. Each subset is also accompanied by a local time variable t_i . For each position x within U_i , we increase t_i by the resting time $\nu(x)$. We choose new positions independently within U_i until t_i exceeds the current estimate of $\sup_i \nu$.

When the above loop exists, we have determined our new position and we subtract $\sup_i \nu$ from t_i in anticipation of the next time we enter U_i . In a sense we are using a different time discretization in each subset U_i of Ω .

This modified algorithm avoids spending a lot of time in areas of low probability for two reasons. First, the low probability subdomains will rarely be entered when drawing from the discrete distribution $\tilde{\nu}$. Secondly, the amount of time necessary to exit the loop in those regions is much smaller, since the local $\sup \nu$ is small. By exploiting the subdivision of Ω , and using an efficient algorithm for drawing from discrete distributions, we improve the efficiency of the SF algorithm drastically in nontrivial cases.

Both $\tilde{\nu}$ and the estimates $\sup \nu$ are adaptively improved throughout the simulation.

4.3 Simulation test

To check the rate of convergence of the different methods, we simulated from a Gaussian distribution with unit variance. The rate of variate generation,

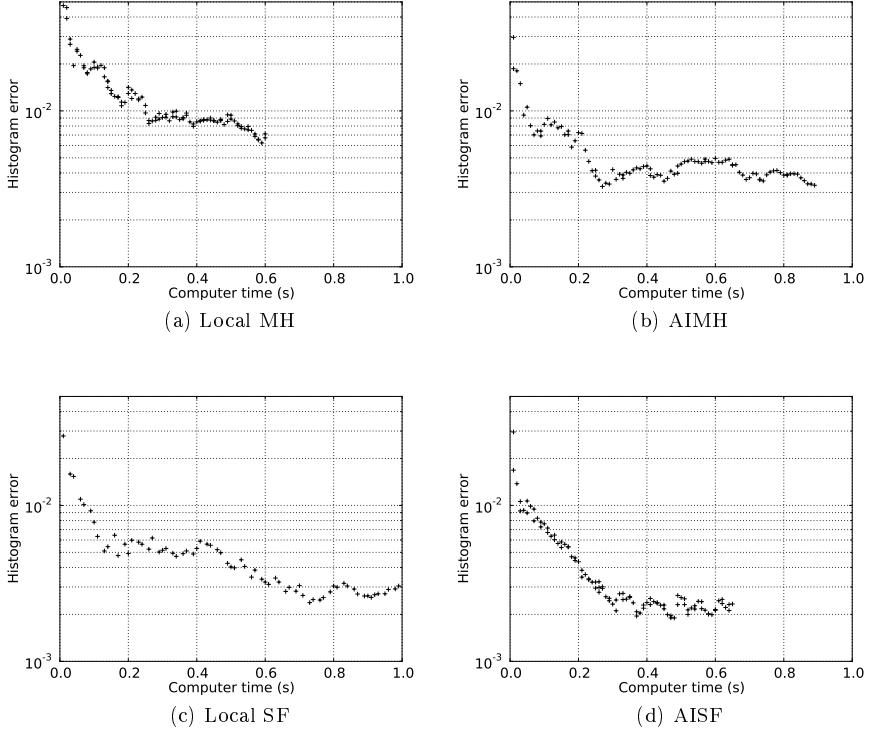


Figure 3: Histogram error versus computer time, for simulation from normal distribution.

correlation, and the convergence towards the known exact result were analyzed.

A histogram of the simulated values is gathered, and compared with a histogram calculated from the Gaussian distribution. We define the histogram error using an L^∞ measure, i.e. as the absolute value of the maximum deviation from the exact result. The plots show the histogram error as a function of simulation running time. The simulations were run on one core of an Intel T4400 laptop processor, but only the relative efficiencies matter here. Results are shown in Figure 3 and are discussed below.

4.3.1 Local MH

Since a low correlation will be important for our later use on Levy processes, we selected the transition probability width 0.75, giving the minimal correlation of approximatively 0.55, according to Figure 1. The variate generation rate was approximately $3.4 \cdot 10^6 / s$ in this case.

4.3.2 AIMH

As expected, as great improvement of the correlation was obtained compared to local MH. Since the algorithm is more complicated, the variate generation rate decreased to $2.3 \cdot 10^6/s$, or roughly 60% of local MH. However, the correlation was reduced to around 0.03. This dramatic decrease results in a faster histogram convergence in terms of computer time. The lower correlation nature of this algorithm will be noticeably beneficial when simulating Levy processes.

4.3.3 Local SF

We adjusted the basic SF algorithm parameters to give zero correlation, and used a uniform transition probability density on $[-5, 5]$. Thus we are running it quite inefficiently as regards variate generation speed, which turned out to be around $1.4 \cdot 10^6/s$. Despite this, the histogram converges faster than local MH, due to the lack of correlation.

4.3.4 AISF

As expected, when including the domain subdivision and adaptive algorithm, the SF algorithm efficiency increased (in fact doubled) with a variate generation rate of $3.1 \cdot 10^6/s$. The amount of correlation here is still zero, which leads to the fastest histogram convergence. So this is a great improvement at no cost other than increased algorithm complexity. It is perfectly suited for simulating consecutive jumps in jump-diffusion process paths.

5 Jump-diffusion simulation

As previously mentioned, existence of correlations among jumps within the same jump-diffusion path is disastrous. Let us check this by using an local MH algorithm to simulate an exactly solvable stochastic process [6] and compare distributions at $t = 1$. It is defined as

$$X_t := \sigma W_t + J_t,$$

$$J_t = \sum_{i=1}^{N_t} Y_i, \quad Y_i \sim N(\mu, \delta^2), \quad \mu, \delta \in \mathbb{R},$$

where N_t is a Poisson process of intensity λ . We choose $\sigma = 1$, $\mu = 0$, $\delta = 1$, and $\lambda = 10$ in order to get an appreciable number of jumps within the time domain $[0, 1]$.

The exact PDF of this process for $t > 0$ is [3, Eq.(4.12)]

$$p_t(x) = \exp(-\lambda t) \sum_{k=0}^{\infty} \frac{(\lambda t)^k \exp(-\frac{x^2}{2(t+k)})}{k! \sqrt{2\pi(t+k)}}.$$

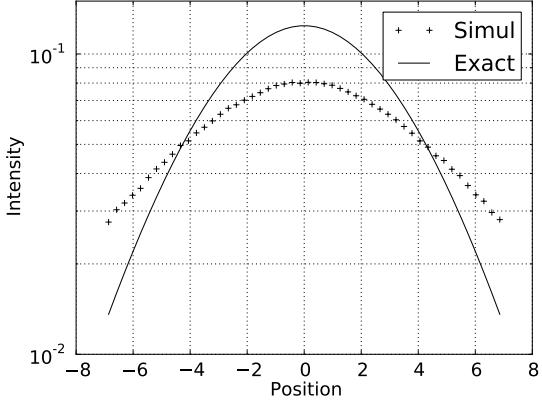


Figure 4: Local MH simulation of distribution at $t = 1$ of Gaussian process, and exact result. Due to jump correlations, the simulated histogram has heavy tails.

For local MH simulations, the positive correlations between subsequent jumps leads to heavy distribution tails. Results confirming this are shown in Figure 4. The simulation of the jumps used a localized uniform transition probability of width 4, which resulted in a MH acceptance rate of around 0.56.

We now turn to our serious attempts at jump-diffusion simulation. We have simulated the same process using AIMH and AISF algorithms. In this case we use the already known value of $\lambda = 10$ in the simulations, so the adaptability consists solely of the calculation of the discrete subdivision distribution, and in the AISF case also of the local subdivision supremum calculations. As opposed local MH, in this case the distribution at $t = 1$ approaches the exact curve, as seen in Figure 5. Note that despite our use of upper/lower cutoffs in the implementation of the discrete subdivision, the tails of the local MH simulation are still somewhat heavy.

We use the same L^∞ measure of convergence detailed above. Convergence results are shown in Figure 6. The AISF outperforms the AIMH algorithm due to its lower correlation which leads to faster and more accurate convergence.

6 Application to NIG and CGMY infinite activity pure jump processes

We present two applications within the realm of infinite activity pure jump Levy processes, namely NIG and CGMY. To this end, we employ the jump diffusion approximation of these processes [1].

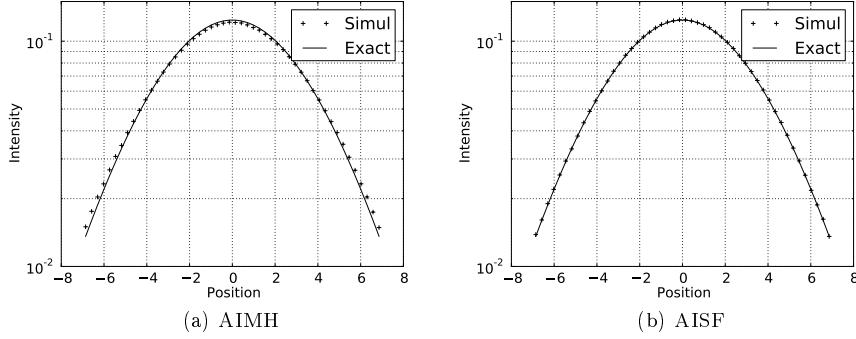


Figure 5: Distributions of Gaussian process at $t = 1$ for AIMH and AISF simulations. The continuous curve represents the exact result. Notice the slight heavy tails in the AIMH case, due to nonzero jump correlations.

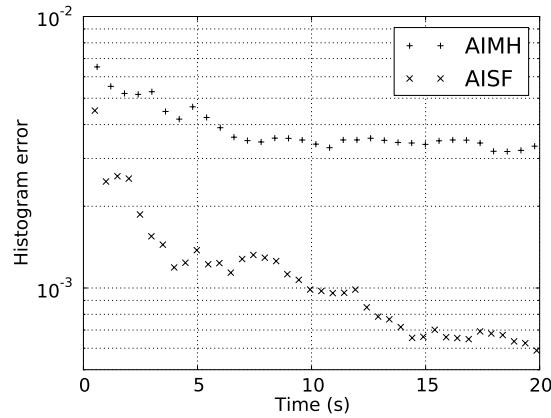


Figure 6: Error in distribution at $t = 1$ for the Gaussian process, versus computer simulation time. AISF gives the fastest convergence. The AIMH heavy tails cause a lower bound on the error.

6.1 Jump-diffusion approximation

For an infinite activity Levy process, the intensity λ is undefined, since its defining integral (2) diverges. However, by virtue of (5), it is possible to approximate the infinitude of small jumps by a Brownian motion process. Consider an infinite activity pure jump Levy process L_t . For $\epsilon > 0$, define the following subsets of Ω ,

$$\begin{aligned} A^{\epsilon,-} &:= \{|x| < \epsilon : x \in \mathbb{R}\} \\ A^{\epsilon,+} &:= \{|x| > \epsilon : x \in \mathbb{R}\}. \end{aligned}$$

These represent small and large jumps, respectively.

We now define the unique Levy measures $\nu^{\epsilon,-}$ and $\nu^{\epsilon,+}$ on these subdomains as follows. For any ν -measurable $U \subset \mathbb{R} - \{0\}$, define

$$\begin{aligned} \nu^{\epsilon,-}(U) &:= \nu(U \cap A^{\epsilon,-}) \\ \nu^{\epsilon,+}(U) &:= \nu(U \cap A^{\epsilon,+}). \end{aligned}$$

Each of these Levy measures in turn define its own Levy process L_-^ϵ and L_+^ϵ of infinite and finite activity, respectively. We have the unique process decomposition

$$L_t = L_t^{\epsilon,-} + L_t^{\epsilon,+}. \quad (8)$$

It follows from (5) that the intensity of the large jump component process,

$$\lambda^{\epsilon,+} := \nu^{\epsilon,+}(\mathbb{R})$$

is well-defined, and so $L_t^{\epsilon,+}$ is a compound Poisson process. For small ϵ , the small jump process $L_t^{\epsilon,-}$ can be often be approximated by the following Brownian motion [1]

$$L_t^{\epsilon,-} \approx \sigma(\epsilon)W_t, \quad \sigma(\epsilon)^2 := \text{Var}[L_1^{\epsilon,-}], \quad (9)$$

where W_t is a Wiener process. The variance must exist for the approximation to be well-defined. A sufficient condition is [1]

$$\lim_{\epsilon \rightarrow 0} \frac{\sigma(\epsilon)}{\epsilon} = \infty. \quad (10)$$

Thus we have the approximation

$$L_t \approx \sigma(\epsilon)W_t + L_t^{\epsilon,+}. \quad (11)$$

Any additional nonvanishing deterministic drift component of L_t would appear trivially on the right hand side of this equation.

In these cases, the intensity λ is not known beforehand, and also depends on our choice of ϵ . Since our algorithms already calculate $\tilde{\nu}$, and the individual volumes of each subset in the subdivision of Ω is known, it is easy to use this to update an estimate for λ concurrently.

6.2 Simulation of NIG

We can now check the quality of the jump-diffusion approximation in conjunction with our AIMH and AISF algorithms. Since the density of NIG is known, and a direct algorithm [3, Alg.6.12] exists for simulating from its distribution at any time, we have ample material to compare our results. The numerical results for the density at $t = 1$ are collected using logarithmic plots in Figure 7, in order to emphasise the distributional tail behaviour. In terms of the parametrisation used in [3], we use $\sigma = 1$, $\theta = 0$, $\kappa = 1/2$. Figure 7 contains the results from the direct simulation algorithm, as well as the results of the jump-diffusion approximation using AIMH and AISF, where we have used the small jump cutoff value $\epsilon = 0.005$.

The volatility σ of the Brownian component in the jump-diffusion approximation, defined by (9), is calculated analytically using an approximate expression for the Levy measure at small $|x|$. The Levy jump density is

$$\nu(x) = \frac{\alpha\delta}{\pi|x|} e^{\beta x} K_1(\alpha|x|), \quad (12)$$

where K_1 is the irregular modified cylindrical Bessel function of first order. As an asymptotic approximation of (12) at small $|x|$, we use

$$e^{\beta x} \approx 1, \quad K_1(\alpha|x|) \approx \frac{1}{\alpha|x|}.$$

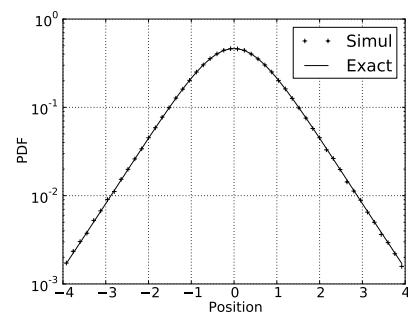
Inserting the chosen parameter values, and using (9) which gives σ^2 as the second moment of the Levy jump density on $[-\epsilon, \epsilon]$, we get

$$\sigma \approx \sqrt{2\epsilon\delta/\pi} \approx 0.067.$$

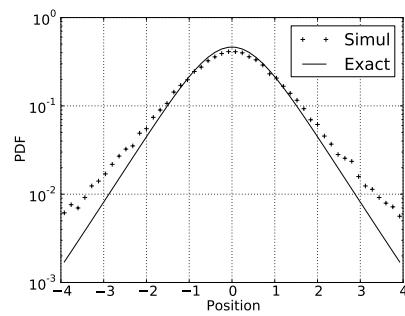
Notice that this expression satisfies the Asmussen/Rosinski condition (10). Owing to this small value, the Brownian part of the Levy process has a negligible influence on the distribution tails at $t = 1$, compared to the contributions from larger jumps.

The direct algorithm from [3] works best, as expected. We are doing this to compare AIMH and AISF, for a known case with an exact solution. The AIMH and AISF algorithms are general and can be used on any Levy process that allows a jump-diffusion approximation. This case gives us an indication of how trustworthy these methods are when applied to more exotic Levy processes for which we do not have an exact result or a simplified simulation methods as in this case.

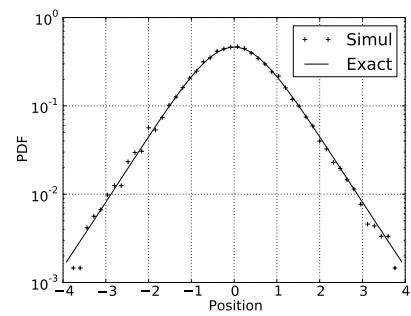
Most sources of errors are common to both algorithms. These are related to inaccuracy in the calculation of the σ and λ , the latter being related to the cutoff imposed on the jump domain Ω . This cutoff will cause an inaccuracy in λ since the total weight of Ω will not be calculated. Improved subdivision schemes of Ω are possible, e.g. employing coordinate transformations that transform Ω into a compact domain. We have not done this, since it unimportant for the matters we are considering.



(a) Cont/Tankov



(b) AIMH



(c) AISF

Figure 7: Simulated and exact NIG PDFs at $t = 1$. The AIMH algorithm has heavy tails.

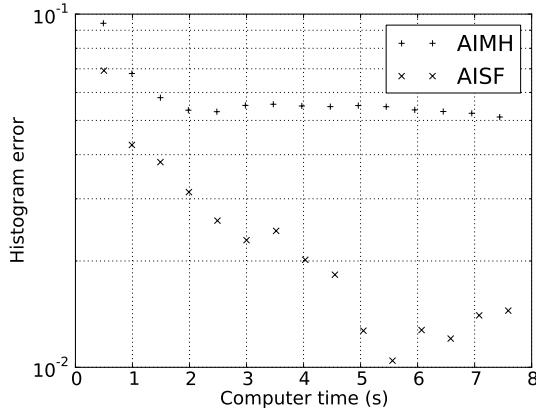


Figure 8: NIG histogram error at $t = 1$, versus computer time. The AISF clearly converges faster. AIMH reaches a steady state in the convergence plot due to the error caused by the nonvanishing jump correlations.

We note that AIMH will never be completely correlation-free, and will therefore tend to produce heavy tails as is obvious in Figure 7. No such problem exists for AISF. The cutoff on Ω will naturally lead to weak tails, as seen in the plot. This can be remedied by enlarging the cutoff value, and/or treating large jumps differently.

The measurement data for convergence analysis is plotted in Figure 8, from which one readily sees that the AISF algorithm converges faster and more exactly.

6.3 Simulation of CGMY

As a second example of a pure jump infinite activity process, we turn to CGMY [2]. In this case we have no closed form expression for the distribution. We do however have the characteristic function, from which the distribution can be obtained by means of a numerical inverse Fourier transform. We have performed this calculation, and used the result as a the benchmark against which our path simulation algorithms are compared.

In this case, the Levy density is

$$\nu(x) = \begin{cases} Ce^{-Mx}/x^{1+Y} & , x > 0 \\ Ce^{-G|x|}/|x|^{1+Y} & , x < 0. \end{cases} \quad (13)$$

The parameter space for the CGMY model is $C, G, M > 0$ and $Y < 2$. By expanding the Levy density in negative powers of x around the origin and keeping only the lowest order terms, we get

$$\sigma(\epsilon) \sim \epsilon^{1-Y/2},$$

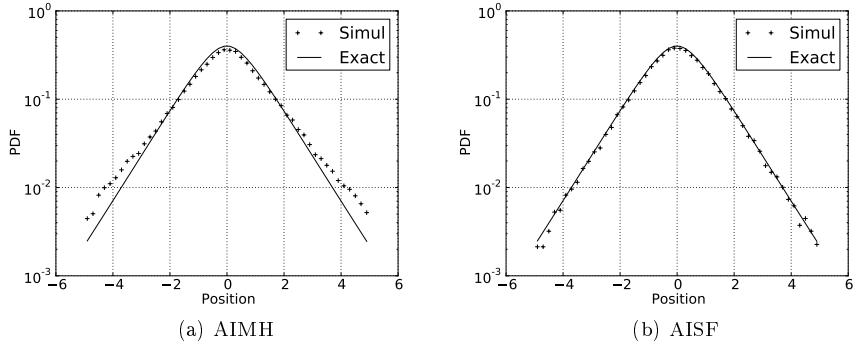


Figure 9: CGMY PDF at $t = 1$, simulated and exact results. Note again the heavy tails in the AIMH case.

so by (10), the jump-diffusion approximation is valid only for $0 < Y \leq 1$. In fact, the volatility can in this case be expressed exactly in terms of the incomplete gamma function, which we will not show here.

We simulated the CGMY process with parameters $C = G = M = 1$, $Y = 1/2$, using $\epsilon = 0.005$, and 100000 paths. In this case, the volatility for our choice of parameter values is

$$\sigma \approx 0.022.$$

Also in this case it has negligible influence on the distribution tails at $t = 1$.

The results for the distribution at $t = 1$ of the process is given in Figure 9, and the convergence measurements are plotted in Figure 10. The conclusions are similar to the NIG case.

7 Conclusions

We have studied two algorithms for jump-diffusion and infinite activity pure jump process simulation via jump-diffusion approximations. Most importantly, we have studied the proposed SF algorithm based on a stochastic step function. It has some advantages over MH accept/reject algorithms. It is possible to configure it to have an arbitrarily small autocorrelation. Our simulations show that in the case of simulation of Levy processes, this algorithm can represent an improvement over the MH method that we have considered. The numerical results show an improvement in the tails of the distribution of the Levy process at a given time while at the same time converging faster. The MH algorithm will tend to give heavy tails, due to the problem of positive correlations between large jump values.

In our computer simulations, we also implemented a subdomain discretization with a corresponding adaptively improved discrete probability distribution.

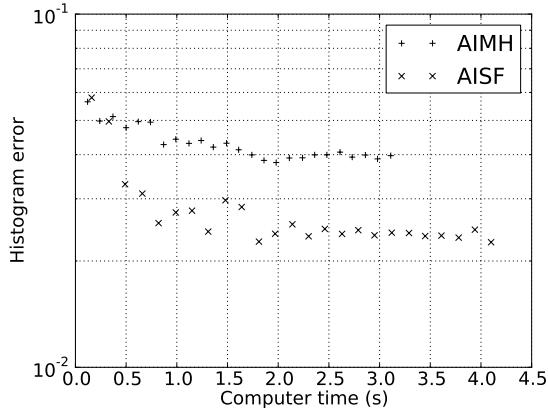


Figure 10: CGMY histogram error at $t = 1$, versus computer time.

This method helps to reduce correlations for the MH algorithms, since the sub-domains themselves are drawn without using an accept/reject algorithm. In the SF case, it improves the variate generation speed while maintaining the correlation-free nature of the method.

References

- [1] S. ASMUSSEN AND J. ROSINSKI, *Approximations of small jumps of levy processes with a view towards simulation*, J. Appl. Probab., 38 (2001), pp. 482–493.
- [2] P. CARR, H. GEMAN, D. MADAN, AND M. YOR, *Stochastic volatility for Levy processes*, Mathematical Finance, 13 (2003), pp. 345–382.
- [3] R. CONT AND P. TANKOV, *Financial Modelling with Jump Processes*, Chapman & Hall/CRC Financial Mathematics Series, 2003.
- [4] W. K. HASTINGS, *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika, 57 (1970), pp. 97–109.
- [5] G. P. LEPAGE, *New algorithm for adaptive multidimensional integration*, Journal of Computational Physics, 27 (1978), pp. 192–203.
- [6] R. MERTON, *Option pricing when underlying stock returns are discontinuous*, Journal of Financial Economics, 3 (1976), pp. 125–144.
- [7] N. METROPOLIS, A. W. ROSENBLUTH, M. N. ROSENBLUTH, A. H. TELLER, AND E. TELLER, *Equation of state calculations by fast computing machines*, The Journal of Chemical Physics, 21 (1953), pp. 1087–1092.

- [8] A. J. WALKER, *An efficient method for generating discrete random variables with general distributions*, ACM Trans on Mathematical Software, 3 (1977), pp. 253–256.

Chapter 9

Simplicial gauge theory on spacetime

Submitted in 2011 to “Numerische Mathematik”. Awaiting referee report.

Simplicial gauge theory on spacetime

Tore Gunnar Halvorsen

Torquil Macdonald Sørensen

the date of receipt and acceptance should be inserted later

Abstract We define a discrete gauge-invariant Yang-Mills-Higgs action on spacetime simplicial meshes. The formulation is a generalization of classical lattice gauge theory, and we prove consistency of the action in the sense of approximation theory. In addition, we perform numerical tests of convergence towards exact continuum results for several choices of gauge fields in pure gauge theory.

Keywords Lattice gauge theory · QCD · Simplicial complex · Yang-Mills theory · Finite element method

Mathematics Subject Classification (2000) 35Q40 · 65M50 · 74S05 · 81T13 · 81T25

1 Introduction

The electromagnetic, weak and strong nuclear interactions are known to be described well by Yang-Mills gauge quantum field theory [25, 21, 22]. Therefore, numerical investigations of this theory are important. At the outset, this is a continuum theory, so a numerical description can be obtained by discretizing spacetime and the field variables. Usually, this is done on a hypercubic lattice, and such numerical studies have contributed greatly to the advancement of the field of fundamental particle physics, in

T. G. Halvorsen
Department of Mathematical Sciences, Norwegian University of Science and Technology, NO-7491
Trondheim, Norway
E-mail: toregha@gmail.com

T. M. Sørensen
Centre of Mathematics for Applications, University of Oslo, P.O.Box 1053 Blindern, NO-0316 Oslo,
Norway
E-mail: t.m.sorensen@matnat.uio.no, torquil@gmail.com

particular within the theory of strong interactions, quantum chromodynamics (QCD) [24].

This article is the natural prolongation of an earlier work [6]. In [6], we defined a gauge invariant discrete action to approximate the Yang-Mills-Higgs (YMH) action¹ on a simplicial mesh approximation of a spatial domain, and we proved consistency of the discrete action for smooth fields. In this article we expand the discussion to 4d spacetime simplexes by including a time dimension, and we expand the consistency proof to fields in the natural energy norm.

More precisely, we define the spatial part of the discrete action as in [6], using Whitney forms [23], Wilson lines and loops [24, 10], and concepts from finite element methods (FEM) [9, 18, 19]. We then expand the spatial mesh to spacetime by introducing a uniform time discretization. The spacetime mesh is defined by repeating the spatial mesh at every time step. Furthermore, we extend the Whitney forms to spacetime, yielding a natural expansion of the discrete spatial action to a spacetime action.

The purpose of this work is to prove consistency of the resulting approximation scheme for the classical YMH equations, and also to develop an analogue of lattice gauge theory (LGT), i.e. performing quantum field theory simulations using this action. The latter is described in a companion article [17].

Simplicial lattices in the context of Yang-Mills theory have been studied before, with promising results [14, 15, 11, 13, 4, 12, 1, 3, 2]. However, this work goes in a different direction than those earlier works, with a different, more general and mathematical stringent approach. A general approach that is mesh-independent is essential if mesh-refinement is to be performed.

We use the same notation as in [6], and the article is organized as follows. In section 2 we introduce the continuous YMH action, and also develop the proposed simplicial gauge theory. In section 3 we review some properties of the differential of the exponential map for matrix groups and the Baker-Campbell-Hausdorff formula. We prove consistency of the action in the energy norm in section 4. Finally, in section 5 we describe some numerical convergence tests, performed on several gauge fields for which the continuum action was exactly calculable. Conclusions are given in section 6.

2 Construction

2.1 The continuous Yang-Mills action

Let $\mathbb{M} = \mathbb{R} \times S$ be a riemannian spacetime manifold, where \mathbb{R} represents time and S is a bounded domain in three dimensional euclidean space. We let \mathbb{M} be equipped with a lorentzian or euclidean signature and coordinates $x = (t, \mathbf{x})$. Furthermore, let \mathcal{G} be a compact Lie group with associated Lie algebra \mathfrak{g} , and assume that \mathcal{G} can be represented by a subgroup of the complex unitary $n \times n$ matrices, for some n . The

¹ Albeit without a scalar potential.

hermitian conjugate of a matrix g is denoted g^H and the real valued scalar product is

$$g' \cdot g := \Re \text{tr}(g' g^H). \quad (2.1)$$

The space of smooth k -forms on \mathbb{M} is denoted $\Omega^k(\mathbb{M})$, and the space $\Omega^k(\mathbb{M}) \otimes \mathfrak{g}$ can be identified with the space of smooth \mathfrak{g} -valued k -forms on \mathbb{M} . The bracket of Lie algebra valued forms is defined as

$$[u \otimes g, u' \otimes g'] := (u \wedge u') \otimes [g, g'], \quad (2.2)$$

where u, u' are real valued differential forms and $g, g' \in \mathfrak{g}$.

A connection one-form on \mathbb{M} is an element $\mathbb{A} = (A_0, A) \in \Omega^1(\mathbb{M}) \otimes \mathfrak{g}$, where A_0 represents the time component and A the spatial component. The temporal curvature $\mathcal{F}^t(\mathbb{A})$ and spatial curvature $\mathcal{F}^s(A)$ of such a one-form are given by

$$\mathcal{F}^t(\mathbb{A}) = dA_0 + d_t A + [A_0, A], \quad \mathcal{F}^s(A) = dA + \frac{1}{2}[A, A], \quad (2.3)$$

where d_t and d denote exterior derivatives in the temporal and spatial direction respectively. We will also use such forms with lower regularity, see e.g. [7]. The independent variable in Yang-Mills theory is such a connection one-form (not necessarily smooth), and the action functional that describes it is given by $S[\mathbb{A}] = S[\mathbb{A}]_t + S[A]_s$, where

$$S[\mathbb{A}]_t := \int_{\mathbb{M}} |\mathcal{F}^t(\mathbb{A})|^2, \quad S[A]_s := \int_{\mathbb{M}} |\mathcal{F}^s(A)|^2. \quad (2.4)$$

A gauge transformation of a connection one-form \mathbb{A} is associated with a choice of $G \in \mathcal{G}$ for each $x \in \mathbb{M}$, such that

$$A_0(x) \mapsto G(x)(A_0(x) + d_t)G^{-1}(x), \quad A(x) \mapsto G(x)(A(x) + d)G^{-1}(x). \quad (2.5)$$

The action $S[\mathbb{A}]$ is invariant under such gauge transformations.

2.2 The interpolated FEM action

In the FEM formulation, we assume \mathcal{T} to be a simplicial complex spanning the spatial domain S . The simplexes are referred to as vertexes, edges, faces and tetrahedra according to dimension, and are labeled i, e, f and T respectively. The symbol T will also be used for simplexes of any dimension. We also suppose an orientation has been chosen for each simplex in \mathcal{T} . In addition, we assume that time is discretized with a time step Δt , and that the simplicial complex \mathcal{T} is repeated at every time step, resulting in a spacetime simplicial complex \mathbb{T} . For a more detailed definition of simplicial complexes, consult [8, section 5].

Thus, as the basic building block in classical FEM theory is a tetrahedron T , the basic building block in this extended FEM version is $T \times I_\tau$, where $I_\tau = [\tau, \tau + \Delta t]$ and τ denotes the temporal nodes.

Furthermore, let $W^k(\mathcal{T})$ ($W^k(T)$) be the space of Whitney k -forms [23] on \mathcal{T} (T), with canonical basis (λ_T) , T ranging over the set of k -dimensional simplexes in \mathcal{T} (T). The 0-forms λ_i are the barycentric coordinate maps taking the value 1 at

vertex i and 0 at all others. For an edge $e = \{i, j\}$, the associated Whitney 1-form is defined by

$$\lambda_e = \lambda_i d\lambda_j - \lambda_j d\lambda_i, \quad (2.6)$$

and for a face $f = \{i, j, k\}$ the associated Whitney 2-form is defined by

$$\lambda_f = 2(\lambda_i d\lambda_j \wedge d\lambda_k + \lambda_j d\lambda_k \wedge d\lambda_i + \lambda_k d\lambda_i \wedge d\lambda_j). \quad (2.7)$$

In order to formulate the Yang-Mills theory in a spacetime FEM setting we need to extend these k-forms to \mathbb{T} . In addition, we need to define the temporal edge and temporal face basis functions, which are constructed as in [8].

The spatial Whitney k-forms are extended to be piecewise affine in time and are denoted $(\Lambda_{T(\tau)})$, i.e.

$$\lambda_T \rightarrow \Lambda_{T(\tau)} = \lambda_T \otimes P_1^t, \quad (2.8)$$

where P_1^t denote polynomials in the time variable of degree at most one, and $T(\tau) := (\tau, T)$ denotes the spatial simplex T at temporal node τ . More precisely, $\Lambda_{T(\tau)}$ is the piecewise affine function in time, taking the value λ_T at τ and 0 at all other temporal nodes. This is consistent with the requirement that the tangential part of the curvature of the gauge potential should be continuous across faces.

The temporal edge basis functions are constructed as follows. To every vertex i in the spatial mesh, there are temporal edges $e_t(\tau) = \{i_\tau, i_{\tau+\Delta t}\}$, where $i_\tau := i(\tau)$. The temporal basis edge function attached to $e_t(\tau)$ is then the piecewise constant function in time defined by

$$\Lambda_{e_t(\tau)}(t) = \begin{cases} \lambda_i \circ \pi \frac{1}{\Delta t} dt, & t \in I_\tau \\ 0, & \text{otherwise.} \end{cases} \quad (2.9)$$

Here, π is the canonical projection onto S , i.e.

$$\pi : \mathbb{M} = \mathbb{R} \times S \rightarrow S, \quad (2.10)$$

and dt is the standard basis one-form in the temporal direction.

Finally, the temporal face elements are constructed as follows. To every spatial edge e there are corresponding temporal faces $f_t(\tau) = e \times I_\tau$. Consider the spatial Whitney edge element λ_e . Then apply the pull-back of π to construct a one-form on spacetime, and then wedge it with dt . More precisely, denote by

$$\pi^* : \Omega(S) \rightarrow \Omega(\mathbb{M}), \quad (2.11)$$

the pull-back induced by π . Then the temporal basis face function is

$$\Lambda_{f_t(\tau)}(t) = \begin{cases} \pi^*(\lambda_e) \wedge \frac{1}{\Delta t} dt = \lambda_e \circ \pi \wedge \frac{1}{\Delta t} dt, & t \in I_\tau \\ 0, & \text{otherwise.} \end{cases} \quad (2.12)$$

This construction ensures that the temporal face basis is orthogonal to the spatial face basis. The space spanned by $(\Lambda_{i(\tau)})$, $(\Lambda_{e(\tau)}) := ((\Lambda_{e_t(\tau)}, \Lambda_{e(\tau)}))$ and $(\Lambda_{f(\tau)}) := ((\Lambda_{f_t(\tau)}, \Lambda_{f(\tau)}))$ are denoted $\mathbb{W}^0(\mathbb{T})$, $\mathbb{W}^1(\mathbb{T})$ and $\mathbb{W}^2(\mathbb{T})$ respectively. If no confusion can arise, the time dependence index τ that identifies a temporal node is usually omitted to compactify notation.

Thus, let $\mathbb{A} = (A_0, A) \in \mathbb{W}^1 \otimes \mathfrak{g}$, $\mathbb{A} = \sum_{e_t} A_{0,e_t} \Lambda_{e_t} + \sum_e A_e \Lambda_e$, where the summations are over oriented edges, and we remark that

$$A_{0,e_t} = \int_{e_t} A_0, \quad A_e = \int_e A. \quad (2.13)$$

The temporal and spatial curvatures of \mathbb{A} are given by

$$\begin{aligned} \mathcal{F}^t(\mathbb{A}) &= \sum_e A_e d_t \Lambda_e + \sum_{e_t} A_{0,e_t} d \Lambda_{e_t} + \sum_{e_t, e} [A_{0,e_t}, A_e] \Lambda_{e_t} \wedge \Lambda_e, \\ \mathcal{F}^s(A) &= \sum_e A_e d \Lambda_e + \frac{1}{2} \sum_{e, e'} [A_e, A_{e'}] \Lambda_e \wedge \Lambda_{e'}. \end{aligned} \quad (2.14)$$

Since $\Lambda_e \wedge \Lambda_{e'} \notin \mathbb{W}^2(\mathbb{T})$ we choose to interpolate $\mathcal{F}^t(\mathbb{A})$ and $\mathcal{F}^s(A)$ onto $\mathbb{W}^2(\mathbb{T})$, instead of working with higher order Whitney elements.

Let (I^t, I^s) and (J^t, J^s) denote interpolation operators onto the temporal and spatial Whitney one- and two-forms, respectively. They are projection operators defined by

$$\begin{aligned} I^t u &= \sum_{e_t} \left(\int_{e_t} u \right) \Lambda_{e_t}, & I^s u &= \sum_e \left(\int_e u \right) \Lambda_e, \\ J^t u &= \sum_{f_t} \left(\int_{f_t} u \right) \Lambda_{f_t}, & J^s u &= \sum_f \left(\int_f u \right) \Lambda_f, \end{aligned} \quad (2.15)$$

and are well defined in particular as maps $\Omega^1(\mathbb{M}) \rightarrow \mathbb{W}^1(\mathbb{T})$ and $\Omega^2(\mathbb{M}) \rightarrow \mathbb{W}^2(\mathbb{T})$, respectively. If we define $I := I^t + I^s$, $J := J^t + J^s$, and $d := (d_t, d)$, we have $d \circ I = J \circ d$ by Stokes' theorem. In particular, $d \circ I^s = J^s \circ d$, consistent with the classical Whitney elements.

The degrees of freedom of the interpolated curvatures are

$$\begin{aligned} J_{f_t}^t(\mathcal{F}^t(\mathbb{A})) &:= \int_{f_t} \mathcal{F}^t(\mathbb{A}) = \sum_{e \in f_t} A_e + \sum_{e_t \in f_t} A_{0,e_t} + \sum_{e, e_t \in f_t} C_{e_t e} [A_{0,e_t}, A_e], \\ J_f^s(\mathcal{F}^s(A)) &:= \int_f \mathcal{F}^s(A) = \sum_{e \in f} A_e + \frac{1}{2} \sum_{e, e' \in f} C_{ee'} [A_e, A_{e'}]. \end{aligned} \quad (2.16)$$

where

$$C_{e_t e} := \int_{f_t} \Lambda_{e_t} \wedge \Lambda_e, \quad C_{ee'} := \int_f \Lambda_e \wedge \Lambda_{e'}.$$

The interpolated FEM action is then $S^J[\mathbb{A}] = S^J[\mathbb{A}]_t + S^J[A]_s$, where

$$\begin{aligned} S^J[\mathbb{A}]_t &:= \int_{\mathbb{M}} |J^t \mathcal{F}^t(\mathbb{A})|^2 = \Re \sum_{f_t, f'_t} M_{f_t f'_t} \text{tr} \left(J_{f_t}^t(\mathcal{F}^t(\mathbb{A})) J_{f'_t}^t(\mathcal{F}^t(\mathbb{A}))^H \right), \\ S^J[A]_s &:= \int_{\mathbb{M}} |J^s \mathcal{F}^s(A)|^2 = \Re \sum_{f, f'} M_{ff'} \text{tr} \left(J_f^s(\mathcal{F}^s(A)) J_{f'}^s(\mathcal{F}^s(A))^H \right). \end{aligned} \quad (2.17)$$

where (\cdot) denotes the scalar product of alternating forms w.r.t. the Lorentzian signature, and the mass matrices are defined as

$$M_{f_t f'_t} := \int_{\mathbb{M}} \Lambda_{f_t} \cdot \Lambda_{f'_t}, \quad M_{ff'} := \int_{\mathbb{M}} \Lambda_f \cdot \Lambda_{f'}, \quad (2.18)$$

By inspiration from lattice gauge theory [24, 10], we will now in several steps construct an approximation to this action that is gauge invariant, i.e. invariant under the transformation (2.5).

2.3 An intermediate action

The spatial part

Let $\{i, j, k, l\}$ be the vertices of T , and pick $A \in W^1(T) \otimes \mathfrak{g}$. Attached to an edge $e = \{i, j\}$ oriented from i to j one has an element $A_e = A_{ij} \in \mathfrak{g}$, where we recall that $A_e = \int_e A$, and parallel transport from i to j is given by $U_{ij} = \exp(A_{ij})$. We suppose U_{ij} to be close enough to 1 so that its logarithm is unambiguous. We use the sign convention $A_{ij} = -A_{ji}$, which corresponds to $U_{ji} = U_{ij}^{-1}$.

The discrete spatial curvature associated with a face $f = \{i, j, k\}$ is defined by

$$F_{ijk}^s := U_{ij} U_{jk} U_{ki}. \quad (2.19)$$

and is in analogy with square Wilson loops in classical lattice gauge theory [24]. By definition, this formula locates the curvature at vertex i . The curvature at vertex j is related to this by

$$F_{jki}^s = U_{ji} F_{ijk}^s U_{ij}, \quad (2.20)$$

which gives a formula for parallel transport of curvature from i to j . Concerning orientation of a face, we notice

$$F_{ijk}^s = (F_{ikj}^s)^{-1}. \quad (2.21)$$

When f is oriented $i \rightarrow j \rightarrow k$, and the curvature is located at i , we write $F_f^s = F_{ijk}^s$. Thus, we have defined the spatial curvature of a pointed oriented face f . The distinguished point of f is denoted \dot{f} , which in this example is $\dot{f} = i$.

A discrete gauge transformation is associated with a choice of $G_i \in \mathcal{G}$ for each vertex i . One then transforms A such that

$$U_{ij} \mapsto G_i U_{ij} G_j^{-1}, \quad (2.22)$$

implying that the spatial curvature transforms as

$$F_f^s \mapsto G_i F_f^s G_i^{-1}, \quad (2.23)$$

similarly to the gauge transformation of the field strength tensor in continuous Yang-Mills gauge theory.

The temporal part

Concerning the temporal part of the curvature, the construction is similar. Let $\{(i_\tau, j_\tau, j_{\tau+\Delta t}, i_{\tau+\Delta t}\}$ be the vertices of the temporal face $f_t(\tau)$, and pick $\mathbb{A} = (A_0, A) \in \mathbb{W}^1 \otimes \mathfrak{g}$. Attached to a spatial edge $e(\tau) = \{i_\tau, j_\tau\}$ one has as before an element $A_{e(\tau)} := A_{i_\tau j_\tau} \in \mathfrak{g}$, and parallel transport from i_τ to j_τ is again given by $U_{ij}(\tau) = \exp(A_{i_\tau j_\tau})$.

Attached to a temporal edge $e_t(\tau) = \{i_\tau, i_{\tau+\Delta t}\}$ one has an element $A_{0,e_t(\tau)} = A_{0,i_\tau i_{\tau+\Delta t}} \in \mathfrak{g}$, where we recall $A_{0,e_t} = \int_{e_t} A_0$, and parallel transport from i_τ to $i_{\tau+\Delta t}$ is given by $U_{0,i_\tau i_{\tau+\Delta t}} = \exp(A_{0,i_\tau i_{\tau+\Delta t}})$. We use the sign convention $A_{0,i_\tau i_{\tau+\Delta t}} = -A_{0,i_{\tau+\Delta t} i_\tau}$ which corresponds to $U_{0,i_\tau i_{\tau+\Delta t}} = U_{0,i_{\tau+\Delta t} i_\tau}^{-1}$. Again, we suppose U_{ij} and $U_{0,i_\tau i_{\tau+\Delta t}}$ to be close enough to the identity, so that their logarithms are unambiguous.

The discrete temporal curvature associated to $f_t(\tau)$ is

$$F_{i_\tau, j_\tau, j_{\tau+\Delta t}, i_{\tau+\Delta t}}^t = U_{ij}(\tau) U_{0,i_\tau j_{\tau+\Delta t}} U_{ji}(\tau + \Delta t) U_{0,i_{\tau+\Delta t} i_\tau}, \quad (2.24)$$

and again we see the similarity with classical lattice gauge theory. This formula locates the temporal curvature at vertex i_τ . The curvature at vertex $i_{\tau+\Delta t}$ is

$$F_{i_{\tau+\Delta t}, i_\tau, j_\tau, j_{\tau+\Delta t}}^t = U_{0,i_{\tau+\Delta t} i_\tau} F_{i_\tau, j_\tau, j_{\tau+\Delta t}, i_{\tau+\Delta t}}^t U_{0,i_\tau i_{\tau+\Delta t}}, \quad (2.25)$$

which gives a formula for parallel transport of curvature from i_τ to $i_{\tau+\Delta t}$. Concerning the orientation of a temporal face, we notice

$$F_{i_\tau, j_\tau, j_{\tau+\Delta t}, i_{\tau+\Delta t}}^t = (F_{i_\tau, i_{\tau+\Delta t}, j_{\tau+\Delta t}, j_\tau}^t)^{-1}. \quad (2.26)$$

When $f_t(\tau)$ is oriented $i_\tau \rightarrow j_\tau \rightarrow j_{\tau+\Delta t} \rightarrow i_{\tau+\Delta t}$, and the curvature is located at i_τ , we write $F_{f_t(\tau)}^t = F_{i_\tau, j_\tau, j_{\tau+\Delta t}, i_{\tau+\Delta t}}^t$. The distinguished point of this pointed oriented face $f_t(\tau)$ is denoted $\dot{f}_t(\tau)$.

Under a discrete gauge transformation, one transforms A_0 such that

$$U_{0,i_\tau i_{\tau+\Delta t}} \mapsto G_{i_\tau} U_{0,i_\tau i_{\tau+\Delta t}} G_{i_{\tau+\Delta t}}^{-1}, \quad (2.27)$$

implying that the temporal curvature transforms as

$$F_{f_t(\tau)}^t \mapsto G_{i_\tau} F_{f_t(\tau)}^t G_{i_\tau}^{-1}. \quad (2.28)$$

The action

We define the intermediate action as $S^I[\mathbb{A}] := S^I[\mathbb{A}]_t + S^I[A]_s$, where

$$S^I[\mathbb{A}]_t = \Re \sum_{f_t, f'_t} M_{f_t f'_t} \text{tr} \left[(F_{f_t}^t(\mathbb{A}) - \mathbb{1})(F_{f'_t}^t(\mathbb{A}) - \mathbb{1})^H \right], \quad (2.29)$$

and

$$S^I[A]_s = \Re \sum_{f, f'} M_{ff'} \text{tr} \left[(F_f^s(A) - \mathbb{1})(F_{f'}^s(A) - \mathbb{1})^H \right]. \quad (2.30)$$

2.4 The simplicial gauge theory action

In this section we use the same notation as in the construction of the intermediate action, equations (2.29) and (2.30).

The goal is of course to formulate a gauge invariant action. The intermediate action is not gauge invariant due to terms that contain products of curvatures of different faces with non-coincident distinguished points. However, this can be resolved by parallel transport. We treat the spatial and temporal part separately.

The spatial part

Let f and f' be two spatial faces of a tetrahedron T . The spatial curvature at f and f' can be connected by the parallel transport operator $U_{\hat{f}\hat{f}'}$. However, the curvature associated to the face f at time τ will interact with the curvature associated to the face f' not only at time τ , but also at times $\tau \pm \Delta t$. This is resolved by the parallel operator in time $U_{0,\hat{f}(\tau)\hat{f}'(\tau')}$.

Thus the troubled term

$$(F_{f(\tau)}^s - \mathbb{1})(F_{f'(\tau')}^s - \mathbb{1})^H \quad (2.31)$$

is replaced by

$$U_{\hat{f}'(\tau)\hat{f}(\tau)}(F_{f(\tau)}^s - \mathbb{1})U_{\hat{f}(\tau)\hat{f}'(\tau)}U_{0,\hat{f}'(\tau)\hat{f}'(\tau')}(F_{f'(\tau')}^s - \mathbb{1})^H U_{0,\hat{f}'(\tau')\hat{f}'(\tau)}, \quad (2.32)$$

and from the transformation properties of F and U , its trace is now gauge invariant.

The temporal part

Let $f_t(\tau)$ and $f'_t(\tau')$ be two temporal faces. By properties of the canonical basis (Λ_{f_t}) we know that interactions between the temporal curvature occur only at equal time intervals. Thus the term

$$(F_{f_t(\tau)}^t - \mathbb{1})(F_{f'_t(\tau')}^t - \mathbb{1})^H \quad (2.33)$$

is replaced by

$$U_{\hat{f}'_t(\tau)\hat{f}_t(\tau)}(F_{f_t(\tau)}^t - \mathbb{1})U_{\hat{f}_t(\tau)\hat{f}'_t(\tau)}U_{0,\hat{f}'_t(\tau)\hat{f}'_t(\tau')}(F_{f'_t(\tau')}^t - \mathbb{1})^H U_{0,\hat{f}'_t(\tau')\hat{f}'_t(\tau)}, \quad (2.34)$$

and we note that $\tau = \tau'$ or $\tau = \tau' + \Delta t$. By the transformation properties of F and U , the trace of this term is also gauge invariant.

The action

We define the simplicial gauge theory (SGT) action as $S^L[\mathbb{A}] := S^L[\mathbb{A}]_t + S^L[A]_s$, where

$$\begin{aligned} S^L[\mathbb{A}]_t := & \Re \sum_{f_t(\tau), f'_t(\tau')} M_{f_t(\tau), f'_t(\tau')} \text{tr} \left(U_{\hat{f}'_t(\tau)\hat{f}_t(\tau)} \left[F_{f_t(\tau)}^t - \mathbb{1} \right] U_{\hat{f}_t(\tau)\hat{f}'_t(\tau)} \times \right. \\ & \left. \times U_{0,\hat{f}'_t(\tau)\hat{f}'_t(\tau')} \left[F_{f'_t(\tau')}^t - \mathbb{1} \right]^H U_{0,\hat{f}'_t(\tau')\hat{f}'_t(\tau)} \right), \end{aligned} \quad (2.35)$$

and

$$\begin{aligned} S^L[\mathbb{A}]_s := \Re \sum_{f(\tau), f'(\tau')} M_{f(\tau), f'(\tau')} \text{tr} & \left(U_{\hat{f}'(\tau) \hat{f}(\tau)} \left[F_{f(\tau)}^s - \mathbb{1} \right] U_{\hat{f}(\tau) \hat{f}'(\tau)} \times \right. \\ & \left. \times U_{0, \hat{f}'(\tau) \hat{f}'(\tau')} \left[F_{f'(\tau')}^s - \mathbb{1} \right]^H U_{0, \hat{f}'(\tau') \hat{f}'(\tau)} \right). \end{aligned} \quad (2.36)$$

Thus we can conclude that

Theorem 1 *The simplicial gauge theory action S^L is discretely gauge invariant.*

2.5 Gauge invariant scalar field action

To complete the Yang-Mills-Higgs-action, we add a discrete gauge invariant action for the scalar Higgs field. No consistence proof for it will be provided, however. The gauge group \mathcal{G} can be represented by a subgroup of the complex unitary $n \times n$ matrices, and in that representation the basic scalar fields will form an n -tuple, i.e.

$$x \mapsto \phi(x) = \begin{pmatrix} \phi_1(x) \\ \vdots \\ \phi_n(x) \end{pmatrix} \in \mathbb{C}^n. \quad (2.37)$$

The action describing this multi-component field is given by $S[\phi, \mathbb{A}] := S[\phi, A_0]_t + S[\phi, A]_s$, where

$$S[\phi, A_0]_t := \int_{\mathbb{M}} |D_{A_0} \phi|^2, \quad S[\phi, A]_s := \int_{\mathbb{M}} |D_A \phi|^2, \quad (2.38)$$

and $D_{A_0} \phi = (d_t + A_0) \phi$ and $D_A \phi = (d + A) \phi$ are the gauge-covariant derivatives on scalar fields. We observe that the action is invariant under the following set of local gauge transformations

$$\begin{aligned} \phi(x) &\mapsto G(x) \phi(x), \\ A_0(x) &\mapsto G(x) (A_0(x) + d_t) G^{-1}(x), \\ A(x) &\mapsto G(x) (A(x) + d) G^{-1}(x), \end{aligned} \quad (2.39)$$

where $G(x) \in \mathcal{G}$. This is so, since the covariant derivatives transform as $D_{A_0} \phi(x) \mapsto G(x) D_{A_0} \phi(x)$ and $D_A \phi(x) \mapsto G(x) D_A \phi(x)$, and $G^H G = \mathbb{1}$.

Our aim is to construct a FEM inspired action for this scalar field that is gauge invariant, and the key ingredient is again the parallel transport operator U introduced in previous sections.

The scalar field is formally a zero-form, and in FEM it has degrees of freedom at the nodes of the mesh. Thus, let $\phi \in \mathbb{W}^0 \otimes \mathbb{C}^n$. Therefore, we can write

$$\phi = \sum_{i_\tau} \phi_{i_\tau} \Lambda_{i_\tau}, \quad \phi_{i_\tau} = \phi(i_\tau). \quad (2.40)$$

The covariant derivatives of ϕ are one-forms, they have degrees of freedom on the edges of the mesh, and are approximated as in LGT [24, 10]. More precisely, let e_t

be a temporal edge and e a spatial edge of the mesh. Their origins are denoted \dot{e}_t and \dot{e} , and their targets \ddot{e}_t and \ddot{e} , respectively. Then the components of the covariant derivative of ϕ along these edges are approximated as

$$\begin{aligned} (D_{A_0}\phi)_{e_t} &\approx (\delta_{A_0}\phi)_{e_t} = \phi_{\ddot{e}_t} - U_{0,\dot{e}_t\dot{e}_t}\phi_{\dot{e}_t}, \\ (D_A\phi)_e &\approx (\delta_A\phi)_e = \phi_{\ddot{e}} - U_{\ddot{e}\dot{e}}\phi_{\dot{e}}. \end{aligned} \quad (2.41)$$

We observe that

$$(\delta_{A_0}\phi)_{e_t} \mapsto G_{\ddot{e}_t}(\delta_{A_0}\phi)_{e_t}, \quad (\delta_A\phi)_e \mapsto G_{\ddot{e}}(\delta_A\phi)_e \quad (2.42)$$

whenever

$$\phi_{i_\tau} \mapsto G_{i_\tau}\phi_{i_\tau}, \quad U_{0,i_\tau+\Delta t,i_\tau} \mapsto G_{i_\tau+\Delta t}U_{0,i_\tau+\Delta t,i_\tau}G_{i_\tau}^{-1}, \quad U_{ij} \mapsto G_iU_{ij}G_j^{-1}. \quad (2.43)$$

Thus the components of the approximated covariant derivatives transform as in the continuous case. However, this is not enough to ensure local gauge invariance of the action, since the inner product of edge basis functions involves interactions between different edges.

The temporal part

By using the approximation of the temporal covariant derivative in equation (2.41), a FEM inspired approximation of $S[\phi, A_0]_t$ reads

$$S^F[\phi, A_0]_t = \Re \sum_{e_t, e'_t} M_{e_t e'_t} (\delta_{A_0}\phi)_{e_t}^H (\delta_{A_0}\phi)_{e'_t}, \quad M_{e_t e'_t} = \int_{\mathbb{M}} \Lambda_{e_t} \cdot \Lambda_{e'_t}. \quad (2.44)$$

We note that this approximation is not gauge invariant since the mass matrix $M_{e_t e'_t}$ is not diagonal. This problem can be overcome by either mass-lumping of $M_{e_t e'_t}$ [5], or by using the parallel transport operator U to localize the terms in the action. The mass-lumping procedure severely restricts the structure of the mesh for time-dependent problems, thus we choose the second alternative.

We make the replacement

$$(\delta_{A_0}\phi)_{e_t}^H (\delta_{A_0}\phi)_{e'_t} \rightarrow (\delta_{A_0}\phi)_{e_t}^H U_{\ddot{e}_t \ddot{e}'_t} (\delta_{A_0}\phi)_{e'_t}, \quad (2.45)$$

in equation (2.44), and approximate the temporal part of the action as

$$S^L[\phi, A_0]_t = \Re \sum_{e_t(\tau), e'_t(\tau)} M_{e_t(\tau) e'_t(\tau)} (\delta_{A_0}\phi)_{e_t(\tau)}^H U_{\ddot{e}_t(\tau) \ddot{e}'_t(\tau)} (\delta_{A_0}\phi)_{e'_t(\tau)}. \quad (2.46)$$

By the transformation properties of $(\delta_{A_0}\phi)$ and U , this is gauge-invariant.

The spatial part

By using the approximation of the spatial covariant derivative in equation (2.41), a FEM inspired approximation of $S[\phi, A]_s$ reads

$$S^F[\phi, A]_s = \Re \sum_{e, e'} M_{ee'} (\delta_A \phi)_e^H (\delta_A \phi)_{e'}, \quad M_{ee'} = \int_{\mathbb{M}} \Lambda_e \cdot \Lambda_{e'}. \quad (2.47)$$

Again, this approximation is not gauge invariant since the mass matrix $M_{ee'}$ is not diagonal, and again we choose to resolve this problem by parallel transport.

We make the replacement

$$(\delta_A \phi)_{e(\tau)}^H (\delta_A \phi)_{e'(\tau')} \rightarrow (\delta_A \phi)_{e(\tau)}^H U_{0, \ddot{e}(\tau) \ddot{e}(\tau')} U_{\ddot{e}(\tau') \ddot{e}'(\tau')} (\delta_A \phi)_{e'(\tau')} \quad (2.48)$$

in equation (2.47), and approximate the spatial part of the action as

$$S^L[\phi, A]_s = \Re \sum_{e(\tau), e'(\tau')} M_{e(\tau)e'(\tau')} (\delta_A \phi)_{e(\tau)}^H U_{0, \ddot{e}(\tau) \ddot{e}(\tau')} U_{\ddot{e}(\tau') \ddot{e}'(\tau')} (\delta_A \phi)_{e'(\tau')}. \quad (2.49)$$

By the transformation properties of $(\delta_A \phi)$ and U , this is gauge invariant. We therefore have

Theorem 2 *The action $S^L[\phi, \mathbb{A}] := S^L[\phi, A_0]_t + S^L[\phi, A]_s$ is discretely gauge invariant.*

3 The differential of the exponential map for matrix groups

Let \mathcal{G} be a compact matrix Lie group with associated Lie algebra $\mathfrak{g} = T_1 \mathcal{G}$. It is well-known that the structures of \mathcal{G} (the connected component containing the identity) and \mathfrak{g} are related through the exponential map

$$\exp : \mathfrak{g} \rightarrow \mathcal{G}, \quad A \mapsto \exp(A), \quad (3.1)$$

which for matrix Lie groups is given by the usual power series expansion

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!} =: e^A. \quad (3.2)$$

In order to apply Hamilton's variational principle to the gauge invariant SGT action introduced above, we need to calculate the differential of the exponential map in an arbitrary direction. In other words, we need to find an expression for

$$D(e^A) \cdot B = \frac{d}{d\tau} e^{A+\tau B} \Big|_{\tau=0}. \quad (3.3)$$

If A and B commute, i.e. $[A, B] := AB - BA = 0$, the differential is straightforward to calculate and equals

$$D(e^A) \cdot B = e^A B. \quad (3.4)$$

This is not the case if $[A, B] \neq 0$.

However, one can prove the following useful formula [16]

Proposition 1 Let X and Y be $n \times n$ ($n \in \mathbb{N}$) complex matrices. Then the following relation holds

$$\frac{d}{d\tau} e^{X+\tau Y} \Big|_{\tau=0} = e^X \left[\frac{\mathbb{1} - e^{-ad_X}}{ad_X} Y \right], \quad (3.5)$$

where $ad_X : \mathfrak{g} \rightarrow \mathfrak{g}$ and $ad_X Y = [X, Y]$. More generally, if $\tau \mapsto X(\tau)$ is a smooth matrix evaluated function, then

$$\frac{d}{d\tau} e^{X(\tau)} = e^{X(\tau)} \left[\frac{\mathbb{1} - e^{-ad_{X(\tau)}}}{ad_{X(\tau)}} \left(\frac{dX(\tau)}{d\tau} \right) \right]. \quad (3.6)$$

By the cyclic invariance of the trace operator, i.e. if A, B, C are $n \times n$ matrices then $tr(ABC) = tr(BCA) = tr(CAB)$, one can also prove

$$\frac{d}{d\tau} tr(e^{X+\tau Y}) \Big|_{\tau=0} = tr(e^X Y). \quad (3.7)$$

3.1 The Baker-Campbell-Hausdorff formula

For complex numbers x and y we know that

$$e^x e^y = e^{x+y}. \quad (3.8)$$

This is not the case when x and y are replaced by matrices X and Y , but the following relation holds [16, 20]:

Proposition 2 Let X and Y be $n \times n$ ($n \in \mathbb{N}$) complex matrices. Then

$$\begin{aligned} e^X e^Y &= e^Z, \quad Z = \sum_{n=1}^{\infty} c_n, \quad c_1 = X + Y, \\ (n+1)c_{n+1} &= \frac{1}{2}[X - Y, c_n] + \sum_{p=1}^{\lfloor n/2 \rfloor} \frac{B_{2p}}{(2p)!} \sum_{\substack{k_i > 0, n \geq 1 \\ k_1 + \dots + k_{2p} = n}} [c_{k_1}, [\dots [c_{k_{2p}}, X + Y] \dots]], \end{aligned} \quad (3.9)$$

where B_j is the j -th Bernoulli number.

The first four terms in the expansion of Z are

$$\begin{aligned} c_1 &= X + Y, \quad c_2 = \frac{1}{2}(XY - YX), \\ c_3 &= \frac{1}{12}(X^2Y + YX^2 - 2XYX + Y^2X + XY^2 - 2YXY), \\ c_4 &= \frac{1}{24}(X^2Y^2 - 2XYXY - Y^2X^2 + 2YXXY). \end{aligned} \quad (3.10)$$

4 Consistency

Since we have formulated the theory in a spacetime FEM setting, we choose to define consistency for the entire action, not only for the spatial part as is usual. This definition encompasses the usual one. Note that we only prove consistency for the gauge field action, without the scalar field. Inclusion of the scalar field is a simple extension of this proof.

We suppose that we have a regular sequence of simplicial meshes \mathcal{T}_n of the spatial domain S . The diameter of a simplex T is denoted h_T , and the largest h_T when $T \in \mathcal{T}_n$ is denoted h_n . In addition, we suppose that time is discretized by a time step Δt_n , and that \mathcal{T}_n is repeated at every time step, resulting in a simplicial mesh \mathbb{T}_n of the spacetime domain \mathbb{M} . We suppose that

$$\max_{\mathbb{T}} \{(h_n), (\Delta t_n)\} \xrightarrow{n \rightarrow \infty} 0. \quad (4.1)$$

The interpolation operators onto the Whitney elements introduced earlier are attached with a subscript n to associate them with the mesh \mathbb{T}_n . Finally, let $X_n = \mathbb{W}^1(\mathbb{T}_n) \otimes \mathfrak{g}$.

Definition 1 We say that two actions S_n and S'_n defined on X_n are consistent with each other, with respect to a norm $\|\cdot\|$, if for any $\mathbb{A} \in X_n$ we have

$$\sup_{\mathbb{A}' \in X_n} \frac{|DS_n[\mathbb{A}]\mathbb{A}' - DS'_n[\mathbb{A}]\mathbb{A}'|}{\|\mathbb{A}'\|} \xrightarrow{n \rightarrow \infty} 0. \quad (4.2)$$

If there is a constant C not depending on n such that quantities a_n and b_n satisfy $a_n \leq Cb_n$ for all n , we write $a_n = \mathcal{O}(b_n)$. To compactify notation the subscript n will be suppressed.

We have introduced three different actions $S^J[\mathbb{A}]$, $S^I[\mathbb{A}]$ and $S^L[\mathbb{A}]$, and the plan is to show

1. $S^J[\mathbb{A}]$ consistent with $S[\mathbb{A}]$,
2. $S^I[\mathbb{A}]$ consistent with $S^J[\mathbb{A}]$,
3. $S^L[\mathbb{A}]$ consistent with $S^I[\mathbb{A}]$,

which implies the consistency between $S^L[\mathbb{A}]$ and $S[\mathbb{A}]$.

We will prove consistency in the energy norm, i.e.

$$A_0, A \in L^\infty(\mathbb{R}; H^1(S)), \quad \partial_t A_0, \partial_t A \in L^\infty(\mathbb{R}; L^2(S)), \quad (4.3)$$

where ∂_t is a shorthand for $\partial/\partial t$. To compactify notation we define

$$\begin{aligned} \|\cdot\|_{L^p(L^q)} &:= \|\cdot\|_{L^p(\mathbb{R}; L^q(S))}, \quad \forall 0 < p, q \leq \infty, \\ \|\cdot\|_{L^\infty(H^1)} &:= \|\cdot\|_{L^\infty(\mathbb{R}; H^1(S))}, \\ \|\mathbb{A}\| &:= \|A_0\| + \|A\|. \end{aligned} \quad (4.4)$$

The H^1 spacetime Euclidean seminorm is denoted $|\cdot|_{H^1(\mathbb{M})}$.

4.1 Consistency between $S^J[\mathbb{A}]$ and $S[\mathbb{A}]$

The two actions are given in equations (2.17) and (2.4) respectively, and

$$\begin{aligned} |DS[\mathbb{A}] \cdot \mathbb{A}' - DS^J[\mathbb{A}] \cdot \mathbb{A}'| &\leq K_t + K_s, \\ K_t &:= |DS[\mathbb{A}]_t \cdot \mathbb{A}' - DS^J[\mathbb{A}]_t \cdot \mathbb{A}'|, \\ K_s &:= |DS[\mathbb{A}]_s \cdot \mathbb{A}' - DS^J[\mathbb{A}]_s \cdot \mathbb{A}'|. \end{aligned} \quad (4.5)$$

We treat K_t and K_s separately. To compactify notation we define

$$\begin{aligned} d_{\mathbb{A}} \mathbb{A}' &:= dA'_0 + d_t A' + [A'_0, A] + [A_0, A'], \\ d_A A' &:= dA' + [A, A'], \\ \partial A &:= \partial_t A + \nabla A, \\ \partial A_0 &:= \partial_t A_0 + \nabla A_0, \end{aligned} \quad (4.6)$$

where ∇ is the spatial gradient on k-forms.

We first estimate the term K_t . From equations (2.17) and (2.4) we get

$$\begin{aligned} K_t &= 2 \left| \int_{\mathbb{M}} \left(\langle \mathcal{F}^t(\mathbb{A}), d_{\mathbb{A}} \mathbb{A}' \rangle - \langle J^t \mathcal{F}^t(\mathbb{A}), J^t d_{\mathbb{A}} \mathbb{A}' \rangle \right) \right| \\ &\leq 2 \|\mathcal{F}^t(\mathbb{A})\|_{L^2(L^2)} \|d_{\mathbb{A}} \mathbb{A}' - J^t d_{\mathbb{A}} \mathbb{A}'\|_{L^2(L^2)} + \\ &\quad + 2 \|\mathcal{F}^t(\mathbb{A}) - J^t \mathcal{F}^t(\mathbb{A})\|_{L^2(L^2)} \|J^t d_{\mathbb{A}} \mathbb{A}'\|_{L^2(L^2)}. \end{aligned} \quad (4.7)$$

The interpolation operator J^t possesses two important properties. First of all it is a projection operator, and second it is stable $L^2 \rightarrow L^2$ by scaling. This implies that

$$\begin{aligned} \|d_{\mathbb{A}} \mathbb{A}' - J^t d_{\mathbb{A}} \mathbb{A}'\|_{L^2(L^2)} &\leq Ch \left(|[A'_0, A]|_{H^1(\mathbb{M})} + |[A_0, A']|_{H^1(\mathbb{M})} \right) \\ &\leq Ch \left(\|A'_0 \partial A\|_{L^2(L^2)} + \|A \partial A'_0\|_{L^2(L^2)} + \|A' \partial A_0\|_{L^2(L^2)} + \|A_0 \partial A'\|_{L^2(L^2)} \right) \\ &\leq Ch \left(\|A'_0\|_{L^\infty(L^6)} \|\partial A\|_{L^\infty(L^3)} + \|A\|_{L^\infty(L^6)} \|\partial A'_0\|_{L^\infty(L^3)} + \right. \\ &\quad \left. + \|A'\|_{L^\infty(L^6)} \|\partial A_0\|_{L^\infty(L^3)} + \|A_0\|_{L^\infty(L^6)} \|\partial A'\|_{L^\infty(L^3)} \right) \\ &\leq Ch^{1/2} \left(\|A'_0\|_{L^\infty(L^6)} \|\partial A\|_{L^\infty(L^2)} + \|A\|_{L^\infty(L^6)} \|\partial A'_0\|_{L^\infty(L^2)} + \right. \\ &\quad \left. + \|A'\|_{L^\infty(L^6)} \|\partial A_0\|_{L^\infty(L^2)} + \|A_0\|_{L^\infty(L^6)} \|\partial A'\|_{L^\infty(L^2)} \right) \\ &\leq Ch^{1/2} \left(\|\mathbb{A}'\|_{L^\infty(H^1)} + \|\partial_t \mathbb{A}'\|_{L^\infty(L^2)} \right) \end{aligned} \quad (4.8)$$

where C is a generic constant, and

$$\begin{aligned} \|\partial \mathbb{A}'\|_{L^\infty(L^2)} &:= \|\partial A'_0\|_{L^\infty(L^2)} + \|\partial A'\|_{L^\infty(L^2)}, \\ \|\partial_t \mathbb{A}'\|_{L^\infty(L^2)} &:= \|\partial_t A'_0\|_{L^\infty(L^2)} + \|\partial_t A'\|_{L^\infty(L^2)}. \end{aligned}$$

By similar arguments, we have

$$\begin{aligned} \|\mathcal{F}^t(\mathbb{A}) - J^t \mathcal{F}^t(\mathbb{A})\|_{L^2(L^2)} \\ \leq Ch^{1/2} \left(\|A_0\|_{L^\infty(L^6)} \|\partial A\|_{L^\infty(L^2)} + \|A\|_{L^\infty(L^6)} \|\partial A_0\|_{L^\infty(L^2)} \right). \end{aligned} \quad (4.9)$$

Combining the above estimates together with the stability of J^t we get

$$K_t \leq Ch^{1/2} \left(\|\mathbb{A}'\|_{L^\infty(H^1)} + \|\partial_t \mathbb{A}'\|_{L^\infty(L^2)} \right). \quad (4.10)$$

Remark 1 If \mathbb{A} is smooth, then

$$K_t \leq Ch \left(\|\mathbb{A}'\|_{L^\infty(H^1)} + \|\partial_t \mathbb{A}'\|_{L^\infty(L^2)} \right). \quad (4.11)$$

From equations (2.17) and (2.4), we estimate the term K_s as

$$\begin{aligned} K_s &= 2 \left| \int_{\mathbb{M}} \left(\langle \mathcal{F}^s(A), d_AA' \rangle - \langle J^s \mathcal{F}^s(A), J^s d_AA' \rangle \right) \right| \\ &\leq 2 \|\mathcal{F}^s(A)\|_{L^2(L^2)} \|d_AA' - J^s d_AA'\|_{L^2(L^2)} \\ &\quad + 2 \|\mathcal{F}^s(A) - J^s \mathcal{F}^s(A)\|_{L^2(L^2)} \|J^s d_AA'\|_{L^2(L^2)}. \end{aligned} \quad (4.12)$$

Remark 2 By similar arguments as we used to bound K_t we get

$$K_s \leq Ch^{1/2} \left(\|A'\|_{L^\infty(H^1)} + \|\partial_t A'\|_{L^\infty(L^2)} \right). \quad (4.13)$$

If A is smooth, then

$$K_s \leq Ch \left(\|A'\|_{L^\infty(H^1)} + \|\partial_t A'\|_{L^\infty(L^2)} \right). \quad (4.14)$$

Summing up, we get

$$\begin{aligned} |DS[\mathbb{A}] \cdot \mathbb{A}' - DS^J[\mathbb{A}] \cdot \mathbb{A}'| &\leq C \left(\|\mathbb{A}'\|_{L^\infty(H^1)} + \|\partial_t \mathbb{A}'\|_{L^\infty(L^2)} \right) \times \\ &\quad \times \begin{cases} h^{1/2} & \mathbb{A} \in L^\infty(H^1), \partial_t \mathbb{A} \in L^\infty(L^2) \\ h & \mathbb{A} \text{ smooth,} \end{cases} \end{aligned} \quad (4.15)$$

implying consistency between $S[\mathbb{A}]$ and $S^J[\mathbb{A}]$.

4.2 Consistency between $S^I[\mathbb{A}]$ and $S^J[\mathbb{A}]$

The interpolated FEM action is given in equation (2.17), and the differential of it is

$$\begin{aligned} DS^J[\mathbb{A}] \cdot \mathbb{A}' &= DS^J[\mathbb{A}]_t \cdot \mathbb{A}' + DS^J[A]_s \cdot A', \\ DS^J[\mathbb{A}]_t \cdot \mathbb{A}' &= \Re \sum_{f_i, f'_i} M_{f_i f'_i} \operatorname{tr}(J_{f_i}^t(\mathcal{F}^t(\mathbb{A})) J_{f'_i}^t (\mathrm{d}_{\mathbb{A}} \mathbb{A}')^H + \text{h.c.}), \\ DS^J[A]_s \cdot A' &= \Re \sum_{f, f'} M_{ff'} \operatorname{tr}(J_f^s(\mathcal{F}^s(A)) J_{f'}^s (d_AA')^H + \text{h.c.}). \end{aligned} \quad (4.16)$$

Let $f_t = \{i_\tau, j_\tau, j_{\tau+\Delta t}, i_{\tau+\Delta t}\}$ be a temporal face, oriented from $i_\tau \rightarrow j_\tau \rightarrow j_{\tau+\Delta t} \rightarrow i_{\tau+\Delta t}$, and write $e_1 = \{(i_\tau, j_\tau)\}$, $e_2 = \{j_\tau, j_{\tau+\Delta t}\}$,

$e_3 = \{j_{\tau+\Delta t}, i_{\tau+\Delta t}\}$, and $e_4 = \{i_{\tau+\Delta t}, i_\tau\}$. For such a face, the constants $C_{e_i e} = -C_{e e_i}$ in the definition of J^t take the value $C_{e_1 e_2} = C_{e_2 e_3} = C_{e_3 e_4} = C_{e_4 e_1} = \frac{1}{4}$. Hence

$$\begin{aligned} J_{f_t}^t(\mathrm{d}_{\mathbb{A}} \mathbb{A}') &= A'_{e_1} + A'_{0,e_2} + A'_{e_3} + A'_{0,e_4} + \frac{1}{4}([A'_{e_1}, A_{0,e_2}] + [A_{e_1}, A'_{0,e_2}]) + \\ &\quad \frac{1}{4}([A'_{0,e_2}, A_{e_3}] + [A_{0,e_2}, A'_{e_3}]) + \frac{1}{4}([A'_{e_3}, A_{0,e_4}] + [A_{e_3}, A'_{0,e_4}]) + \quad (4.17) \\ &\quad + \frac{1}{4}([A'_{0,e_4}, A_{e_1}] + [A_{0,e_4}, A'_{e_1}]). \end{aligned}$$

In addition, we know that $A_{e_1} + A_{e_3} = (dA)_{f_t}$ and $A_{0,e_2} + A_{0,e_4} = (dA_0)_{f_t}$, implying that $A_{e_3} = -A_{e_1} + (d_t A)_{f_t}$ and $A_{0,e_4} = -A_{0,e_2} + (dA_0)_{f_t}$, with similar relations for A'_{e_3} and A'_{0,e_4} . This gives

$$\begin{aligned} J_{f_t}^t(\mathrm{d}_{\mathbb{A}} \mathbb{A}') &= A'_{e_1} + A'_{0,e_2} + A'_{e_3} + A'_{0,e_4} + ([A'_{e_1}, A_{0,e_2}] + [A_{e_1}, A'_{0,e_2}]) + \quad (4.18) \\ &\quad + \mathcal{O}(A_0 d_t A' + A d A'_0 + \mathbb{A} \leftrightarrow \mathbb{A}')_{f_t}, \end{aligned}$$

where e.g. $\mathcal{O}(A_0 d_t A')_{f_t}$ means $\mathcal{O}((A_0)_{e_t} (d_t A')_{f_t})$, $e_t \in f_t$.

Let $f = \{i, j, k\}$ be a spatial face, oriented from $i \rightarrow j \rightarrow k$, and write $e_1 = \{i, j\}$, $e_2 = \{j, k\}$, $e_3 = \{k, i\}$. For such a face, the constants $C_{ee'} = -C_{e'e}$ in the definition of J^s take the value $C_{e_1 e_2} = C_{e_3 e_1} = C_{e_2 e_3} = \frac{1}{6}$. Hence

$$\begin{aligned} J_f^s(d_A A') &= A'_{e_1} + A'_{e_2} + A'_{e_3} + \frac{1}{6}([A'_{e_1}, A_{e_2}] + [A_{e_1}, A'_{e_2}]) + \quad (4.19) \\ &\quad \frac{1}{6}([A'_{e_1}, A_{e_3}] + [A_{e_1}, A'_{e_3}]) + \frac{1}{6}([A'_{e_2}, A_{e_3}] + [A_{e_2}, A'_{e_3}]). \end{aligned}$$

In addition, we know that $A_{e_1} + A_{e_2} + A_{e_3} = (dA)_f$, which means that $A_{e_3} = -A_{e_1} - A_{e_2} + (dA)_f$, with a similar relation for A'_{e_3} , implying that

$$J_f^s(d_A A') = A'_{e_1} + A'_{e_2} + A'_{e_3} + \frac{1}{2}([A'_{e_1}, A_{e_2}] + [A_{e_1}, A'_{e_2}]) + \mathcal{O}(A d A' + A' d A)_f. \quad (4.20)$$

Again, $\mathcal{O}(A d A')_f = \mathcal{O}(A_e (d A')_f)$, $e \in f$.

Furthermore we need to calculate the differential of $S^I[A]$. The intermediate action is given in equations (2.29) and (2.30), and the differential of it is

$$\begin{aligned} DS^I[\mathbb{A}] \cdot \mathbb{A}' &= DS^I[\mathbb{A}]_t \cdot \mathbb{A}' + DS^I[A]_s \cdot A', \\ DS^I[\mathbb{A}]_t \cdot \mathbb{A}' &= \Re \sum_{f_t, f'_t} M_{f_t f'_t} \mathrm{tr} \left(\frac{dF_{f_t}^t(\mathbb{A} + \tau \mathbb{A}')}{d\tau} \Big|_{\tau=0} (F_{f'_t}^t(\mathbb{A}) - \mathbb{1})^H + \mathrm{h.c.} \right), \quad (4.21) \\ DS^I[A]_s \cdot A' &= \Re \sum_{f, f'} M_{ff'} \mathrm{tr} \left(\frac{dF_f^s(A + \tau A')}{d\tau} \Big|_{\tau=0} (F_{f'}^s(A) - \mathbb{1})^H + \mathrm{h.c.} \right). \end{aligned}$$

Thus we need to calculate

$$\frac{dF_{f_t}^t(\mathbb{A} + \tau \mathbb{A}')}{d\tau} \Big|_{\tau=0}, \quad \frac{dF_f^s(A + \tau A')}{d\tau} \Big|_{\tau=0}. \quad (4.22)$$

We choose to calculate this by first writing the exponential functions in F as a single exponential using the BCH formula, and then using the formula for the differential, proposition 1.

Let f_t be as above. Then $F_{f_t}^t(\mathbb{A}) = U_{e_1}U_{e_2}U_{e_3}U_{e_4}$, where $U_{e_i} = \exp(A_{e_i})$ and with similar expressions for U_{e_2}, U_{e_3} and U_{e_4} . By the BCH formula, proposition 2, $F_{f_t}^t(\mathbb{A} + \tau\mathbb{A}')$ can be written as

$$F_{f_t}^t(\mathbb{A} + \tau\mathbb{A}') = e^{A_{e_1}(\tau)}e^{A_{0,e_2}(\tau)}e^{A_{e_3}(\tau)}e^{A_{0,e_4}(\tau)} = e^{W_{f_t}(A_{e_1}, A_{0,e_2}, A_{e_3}, A_{0,e_4})(\tau)}, \quad (4.23)$$

where $\mathbb{A}(\tau) = \mathbb{A} + \tau\mathbb{A}'$, and $W_{f_t}(A_{e_1}, A_{0,e_2}, A_{e_3}, A_{0,e_4})(\tau)$ is given from the recursion formula in proposition 2. If we write $W_{f_t}(\tau) := W_{f_t}(A_{e_1}, A_{0,e_2}, A_{e_3}, A_{0,e_4})(\tau) = \sum_{n=1}^{\infty} d_n$, then the two first terms are

$$\begin{aligned} d_1 &= A_{e_1}(\tau) + A_{0,e_2}(\tau) + A_{e_3}(\tau) + A_{0,e_4}(\tau), \\ d_2 &= \frac{1}{2}[A_{e_1}(\tau), A_{0,e_2}(\tau)] + \frac{1}{2}[A_{e_1}(\tau), A_{e_3}(\tau)] + \frac{1}{2}[A_{e_1}(\tau), A_{0,e_4}(\tau)] + \\ &\quad + \frac{1}{2}[A_{0,e_2}(\tau), A_{e_3}(\tau)] + \frac{1}{2}[A_{0,e_2}(\tau), A_{0,e_4}(\tau)] + \frac{1}{2}[A_{e_3}(\tau), A_{0,e_4}(\tau)]. \end{aligned} \quad (4.24)$$

We note from proposition 2 that $d_n \propto \mathcal{O}(\frac{1}{n}\mathbb{A}(\tau)^n)$. If we again use $A_{e_3} = -A_{e_1} + (d_t A)_{f_t}$ and $A_{0,e_4} = -A_{0,e_2} + (dA_0)_{f_t}$, with similar relations for A'_{e_3} and A'_{0,e_4} , then we get

$$\begin{aligned} W_{f_t}(\tau = 0) &= J_{f_t}^t(\mathcal{F}^t(\mathbb{A})) + \mathcal{O}(\mathbb{A}^3 + \mathbb{A}\mathrm{d}\mathbb{A})_{f_t}, \\ \frac{dW_{f_t}(\tau)}{d\tau} \Big|_{\tau=0} &= J_{f_t}^t(\mathrm{d}_{\mathbb{A}}\mathbb{A}') + \mathcal{O}(\mathbb{A}^2\mathbb{A}' + \mathbb{A}\mathrm{d}\mathbb{A}' + \mathbb{A}'\mathrm{d}\mathbb{A})_{f_t}, \end{aligned} \quad (4.25)$$

which again implies that

$$F_{f_t}^t(\mathbb{A}) = e^{W_{f_t}(\tau=0)} = \mathbb{1} + J_{f_t}^t(\mathcal{F}^t(\mathbb{A})) + \mathcal{O}(\mathbb{A}^3 + \mathbb{A}\mathrm{d}\mathbb{A})_{f_t}. \quad (4.26)$$

Using proposition 1, we get

$$\frac{dF_{f_t}^t(\mathbb{A} + \tau\mathbb{A}')}{d\tau} \Big|_{\tau=0} = F_{f_t}^t(\mathbb{A}) \left[\frac{\mathbb{1} - e^{-ad_{W_{f_t}(0)}}}{ad_{W_{f_t}(0)}} \frac{dW_{f_t}(\tau)}{d\tau} \Big|_{\tau=0} \right], \quad (4.27)$$

where

$$\frac{\mathbb{1} - e^{-ad_{W_f(0)}}}{ad_{W_f(0)}} = \mathbb{1} - \frac{1}{2}ad_{W_f(0)} + \frac{1}{3!}ad_{W_f(0)}^2 + \dots, \quad (4.28)$$

and where we recall that $ad_X Y = [X, Y]$.

Finally, combining the above estimates gives

$$\begin{aligned} (F_{f_t'}(\mathbb{A}) - \mathbb{1})^H F_{f_t}(\mathbb{A}) \left[\frac{\mathbb{1} - e^{-ad_{W_{f_t}(0)}}}{ad_{W_{f_t}(0)}} \frac{dW_{f_t}(\tau)}{d\tau} \Big|_{\tau=0} \right] &= \\ &= J_{f_t'}^t(\mathcal{F}^t(\mathbb{A}))^H \frac{dW_{f_t}(\tau)}{d\tau} \Big|_{\tau=0} + \mathcal{O}(\mathbb{A}^4\mathbb{A}')_{f_t f_t'} \\ &= J_{f_t'}^t(\mathcal{F}^t(\mathbb{A}))^H J_{f_t}^t(\mathrm{d}_{\mathbb{A}}\mathbb{A}') \\ &\quad + \mathcal{O}(\mathbb{A}^4\mathbb{A}' + (\mathbb{A}^3 + \mathbb{A}\mathrm{d}\mathbb{A})\mathrm{d}\mathbb{A}' + (\mathbb{A}^2\mathrm{d}\mathbb{A} + (\mathrm{d}\mathbb{A})^2)\mathbb{A}')_{f_t f_t'}. \end{aligned} \quad (4.29)$$

Let f be as above. Then $F_f^s(A) = U_{e_1}U_{e_2}U_{e_3}$, where $U_{e_1} = \exp(A_{e_1})$ and with similar expressions for U_{e_2} and U_{e_3} . By the BCH formula, proposition 2, $F_f^s(A)$ can be written as

$$F_f^s(A + \tau A') = e^{A_{e_1}(\tau)} e^{A_{e_2}(\tau)} e^{A_{e_3}(\tau)} = e^{W_f(A_{e_1}, A_{e_2}, A_{e_3})(\tau)}, \quad (4.30)$$

where $A(\tau) = A + \tau A'$, and $W_f(A_{e_1}, A_{e_2}, A_{e_3})(\tau)$ is given from the recursion formula in proposition 2. If we write $W_f(\tau) := W_f(A_{e_1}, A_{e_2}, A_{e_3})(\tau) = \sum_{n=1}^{\infty} d_n$, then the first two terms are

$$\begin{aligned} d_1 &= A_{e_1}(\tau) + A_{e_2}(\tau) + A_{e_3}(\tau), \\ d_2 &= \frac{1}{2}[A_{e_1}(\tau), A_{e_2}(\tau)] + \frac{1}{2}[A_{e_1}(\tau), A_{e_3}(\tau)] + \frac{1}{2}[A_{e_2}(\tau), A_{e_3}(\tau)], \end{aligned} \quad (4.31)$$

and we note that $d_n \propto \mathcal{O}(\frac{1}{n}A(\tau)^n)$. If we again use that $A_{e_3} = -A_{e_1} - A_{e_2} + (dA)_f$, with a similar relation for A'_{e_3} , then we get

$$\begin{aligned} W_f(\tau = 0) &= J_f^s(\mathcal{F}^s(A)) + \mathcal{O}(A^3 + AdA)_f, \\ \frac{dW_f(t)}{dt} \Big|_{t=0} &= J_f^s(d_A A') + \mathcal{O}(A^2 A' + AdA' + A'dA)_f, \end{aligned} \quad (4.32)$$

which again implies that

$$F_f^s(A) = e^{W_f(t=0)} = \mathbb{1} + J_f^s(\mathcal{F}^s(A)) + \mathcal{O}(A^3 + AdA)_f. \quad (4.33)$$

Using proposition 1, we get

$$\frac{dF_f^s(A + \tau A')}{d\tau} \Big|_{\tau=0} = F_f^s(A) \left[\frac{\mathbb{1} - e^{-ad_{W_f(0)}}}{ad_{W_f(0)}} \frac{dW_f(\tau)}{d\tau} \Big|_{\tau=0} \right], \quad (4.34)$$

and by combining the above estimates gives

$$\begin{aligned} (F_{f'}^s(A) - \mathbb{1})^H F_f^s(A) \left[\frac{\mathbb{1} - e^{-ad_{W_f(0)}}}{ad_{W_f(0)}} \frac{dW_f(\tau)}{d\tau} \Big|_{\tau=0} \right] &= \\ &= J_{f'}^s(\mathcal{F}^s(A))^H \frac{dW_f(\tau)}{d\tau} \Big|_{\tau=0} + \mathcal{O}(A^4 A')_{ff'} \\ &= J_{f'}^s(\mathcal{F}^s(A))^H J_f^s(d_A A') \\ &\quad + \mathcal{O}(A^4 A' + (A^3 + AdA)dA' + (A^2 dA + (dA)^2)A')_{ff'}. \end{aligned} \quad (4.35)$$

Summing up, we get

$$\begin{aligned}
DS^I[\mathbb{A}]_t \cdot \mathbb{A}' &= \Re \sum_{f_t, f'_t} M_{f_t f'_t} \text{tr} \left((F_{f'_t}^t(\mathbb{A}) - \mathbb{1})^H F_{f_t}^t(\mathbb{A}) \left[\frac{\mathbb{1} - e^{-ad_{W_{f_t}(0)}}}{ad_{W_{f_t}(0)}} \frac{dW_{f_t}(\tau)}{d\tau} \Big|_{\tau=0} \right] + \text{h.c.} \right) \\
&= \Re \sum_{f_t, f'_t} M_{f_t f'_t} \text{tr} \left(J_{f'_t}^t(\mathcal{F}^t(\mathbb{A}))^H J_{f_t}^t(d\mathbb{A} \mathbb{A}') + \text{h.c.} \right) + \\
&\quad + \Re \sum_{f_t, f'_t} M_{f_t f'_t} \text{tr} \left(\mathcal{O} (\mathbb{A}^4 \mathbb{A}' + (\mathbb{A}^3 + \mathbb{A} d\mathbb{A}) d\mathbb{A}' + (\mathbb{A}^2 d\mathbb{A} + (d\mathbb{A})^2) \mathbb{A}')_{f_t f'_t} \right) \\
&= DS^J[\mathbb{A}]_t \cdot \mathbb{A}' + \\
&\quad + \Re \sum_{f_t, f'_t} M_{f_t f'_t} \text{tr} \left(\mathcal{O} (\mathbb{A}^4 \mathbb{A}' + (\mathbb{A}^3 + \mathbb{A} d\mathbb{A}) d\mathbb{A}' + (\mathbb{A}^2 d\mathbb{A} + (d\mathbb{A})^2) \mathbb{A}')_{f_t f'_t} \right). \tag{4.36}
\end{aligned}$$

and

$$\begin{aligned}
DS^I[A]_s \cdot A' &= \Re \sum_{f, f'} M_{ff'} \text{tr} \left((F_{f'}(A) - \mathbb{1})^H F_f(A) \left[\frac{\mathbb{1} - e^{-ad_{W_f(0)}}}{ad_{W_f(0)}} \frac{dW_f(\tau)}{d\tau} \Big|_{\tau=0} \right] + \text{h.c.} \right) \\
&= \Re \sum_{f, f'} M_{ff'} \text{tr} \left(J_{f'}^s(\mathcal{F}^s(A))^H J_f^s(d_A A') + \text{h.c.} \right) + \\
&\quad + \Re \sum_{f, f'} M_{ff'} \text{tr} \left(\mathcal{O} (A^4 A' + (A^3 + A dA) dA' + (A^2 dA + (dA)^2) A')_{ff'} \right) \\
&= DS^J[A]_s \cdot A' + \\
&\quad + \Re \sum_{f, f'} M_{ff'} \text{tr} \left(\mathcal{O} (A^4 A' + (A^3 + A dA) dA' + (A^2 dA + (dA)^2) A')_{ff'} \right). \tag{4.37}
\end{aligned}$$

If we assume that the mesh satisfies a CFL condition, i.e. there exists a constant C such that

$$0 < \frac{1}{C} \leq \frac{\Delta t_n}{h_n} \leq C, \quad \forall n, \tag{4.38}$$

then we can deduce the bounds

$$\begin{aligned}
|A_e| &\leq Ch^{1/2} \|A\|_{L^\infty(L^6)}, \quad |A_{0,e_t}| \leq Ch^{1/2} \|A_0\|_{L^\infty(L^6)}, \\
|(dA)_f| &\leq Ch^{1/2} \|\partial A\|_{L^\infty(L^2)}, \quad |(d_t A)_{f_t}| \leq Ch^{1/2} \|\partial A\|_{L^\infty(L^2)}, \tag{4.39} \\
|(dA_0)_{f_t}| &\leq Ch^{1/2} \|\partial A_0\|_{L^\infty(L^2)},
\end{aligned}$$

and we can conclude

$$\begin{aligned}
|DS^J[\mathbb{A}] \cdot \mathbb{A}' - DS^I[\mathbb{A}] \cdot \mathbb{A}'| &\leq C \left(\|\mathbb{A}'\|_{L^\infty(H^1)} + \|\partial_t \mathbb{A}'\|_{L^\infty(L^2)} \right) \times \\
&\quad \times \begin{cases} h^{1/2} & \mathbb{A} \in L^\infty(H^1), \partial_t \mathbb{A} \in L^\infty(L^2) \\ h & \mathbb{A} \text{ smooth,} \end{cases} \tag{4.40}
\end{aligned}$$

implying consistency between $S^J[\mathbb{A}]$ and $S^I[\mathbb{A}]$.

4.3 Consistency between $S^I[\mathbb{A}]$ and $S^L[\mathbb{A}]$

The only difference between S^I and S^L is the parallel transport operators in S^L introduced to make S^L gauge invariant. They are given as

$$\begin{aligned} U_{\dot{f}\dot{f}'} &= \exp(A_{\dot{f}\dot{f}'}) = \mathbb{1} + A_{\dot{f}\dot{f}'} + \mathcal{O}(A^2), \\ U_{0,\dot{f}(\tau)\dot{f}(\tau')} &= \exp(A_{0,\dot{f}(\tau)\dot{f}(\tau')}) = \mathbb{1} + A_{0,\dot{f}(\tau)\dot{f}(\tau')} + \mathcal{O}(A_0^2), \end{aligned} \quad (4.41)$$

and the differential of these are proportional to A' and A'_0 respectively. Hence, by similar considerations as in the previous section we get exactly the same estimate as in equation (4.40), i.e.

$$\begin{aligned} |DS^I[\mathbb{A}] \cdot \mathbb{A}' - DS^L[\mathbb{A}] \cdot \mathbb{A}'| &\leq C \left(\|\mathbb{A}'\|_{L^\infty(H^1)} + \|\partial_t \mathbb{A}'\|_{L^\infty(L^2)} \right) \times \\ &\quad \times \begin{cases} h^{1/2} & \mathbb{A} \in L^\infty(H^1), \partial_t \mathbb{A} \in L^\infty(L^2) \\ h & \mathbb{A} \text{ smooth.} \end{cases} \end{aligned} \quad (4.42)$$

We summarize the results in a theorem:

Theorem 3 *Assume \mathbb{M} is a bounded domain in \mathbb{R}^{1+3} . Then the SGT action (2.35), is consistent with the continuous Yang-Mills action (2.4), with respect to the norm*

$$\|\mathbb{A}\| := \|\mathbb{A}\|_{L^\infty(H^1)} + \|\partial_t \mathbb{A}\|_{L^\infty(L^2)}, \quad (4.43)$$

under the assumption that the above-mentioned CFL condition holds.

As a consequence of the above estimates, we get the following estimate for the deviation of the SGT action S^L from the continuous action S ,

$$|S(\mathbb{A}) - S^L(\mathbb{A})| \leq \begin{cases} Ch & , \mathbb{A} \in L^\infty(H^1), \partial_t \mathbb{A} \in L^\infty(L^2) \\ Ch^2 & , \mathbb{A} \text{ smooth.} \end{cases} \quad (4.44)$$

5 Numerical convergence tests

The preceeding sections have defined and proven consistency of the SGT action. However, due to the complexity involved in these quantities, we would like to include some numerical convergence tests as well.

In our computer calculations, we focus on pure gauge theory with $\mathcal{G} = SU(2)$, and used a four-dimensional cubic euclidean domain $[0, 1]^4 \subset \mathbb{R}^4$ with periodic boundary conditions. The lattice structure consisted of a three-dimensional simplicial mesh replicated at each discrete time value. The three-dimensional simplicial mesh consisted of a homogeneous arrangement of N^3 identical cubic building blocks, each building block containing six tetrahedra as shown in figure 5.1. Each such spatial mesh was replicated N times in the time direction to fill the four-dimensional domain, in accordance with the construction detailed in the previous sections. The lattice constant h is defined as the side length of each cubic building block, and also coincides

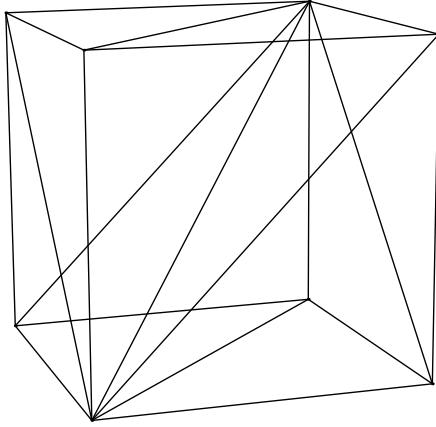


Figure 5.1 Elementary 3d mesh building block containing six tetrahedra, all of which share the single interior diagonal. This particular choice implies an anisotropy in the discretization. This anisotropy disappears in the continuum limit.

with the time discretization interval. In the interest of simplicity we enforced temporal gauge, in which the temporal link matrices reduce to identity matrices.

By defining the distinguished points of all spatial and temporal faces to coincide for as many pairs of faces as possible, we only need the parallel transport matrices for terms in the action involving pairs of temporal faces with no common nodes.

In order to test convergence of the euclidean SGT, we compared the discrete and continuum action for several different choices of gauge fields for which the continuum action S 2.4 can be calculated exactly. We chose the following cases.

1. Gauge field oriented towards the x -direction in space and towards the generator $t^3 = i\sigma^3/2$ within $\mathfrak{su}(2)$, with a sinusoidal time dependence, where σ^3 is a Pauli matrix. The only nonzero component of the gauge field A is

$$A_x^3(t, x, y, z) := \frac{1}{2\pi} \sin(2\pi t), \quad S = 1.$$

2. Gauge field oriented towards the y -direction in space and t^3 within $\mathfrak{su}(2)$, with a sinusoidal x -dependence. The nonzero component of the gauge field in this case was

$$A_y^3(t, x, y, z) := \frac{1}{2\pi} \sin(2\pi x), \quad S = 1.$$

3. A case with two nonzero components,

$$A_x^1 := \frac{1}{2\pi} \sin(2\pi y), \quad A_y^2 := \frac{1}{2\pi} \sin(2\pi x), \quad S = \frac{1}{2} + \frac{1}{8(2\pi)^4}.$$

4. A constant field that only contributes to the nonlinear term in the field strength,

$$A_x^1 := 1, \quad A_y^2 := 1, \quad S = \frac{1}{2}.$$

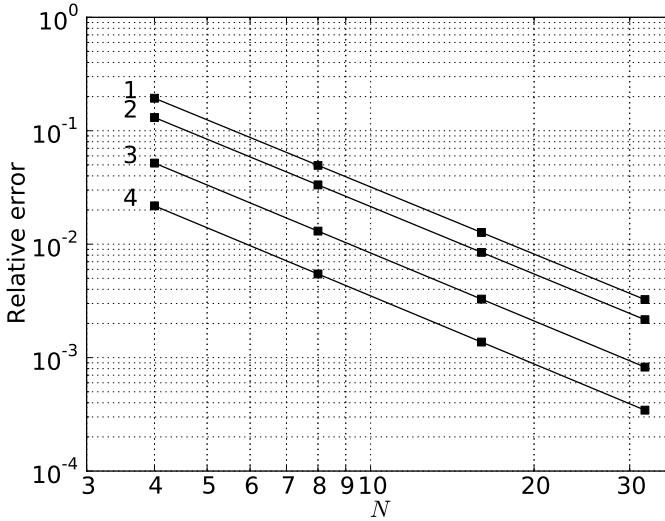


Figure 5.2 The relative error of the action versus the number of lattice sites per side N , for the actions 1, 2, 3, 4 described in section 5. The squares are the simulation data points and the solid lines are the second order polynomial fits. Errors are proportional to h^2 in all cases.

The first case is insensitive to the spatial face mass matrix elements, while the second is insensitive to the spatial edge mass matrix elements which are used in the definition of the temporal mass matrix elements. In the first two cases, the nonlinear contribution to the continuum field strength F vanishes, and the action can be calculated analytically to be unity. In the third case, the nonlinear term survives, and the exact value of the action is

$$S = \frac{1}{2} + \frac{1}{8(2\pi)^4}.$$

In all cases, we measure the relative error of the discretized action, versus the lattice size, for lattices sizes N^4 from $N = 4$ to $N = 32$. The results are displayed graphically in figure 5.2.

We have plotted the absolute value of the relative deviation of the numerical action from the continuum action, as a function of the number of lattice sites per dimension. In all cases, the errors approach zero as the lattice resolution grows, so the action converges towards the correct continuum result. Using least squares second order polynomial fits, we determined that the relative error depends in the following way on the lattice constant h ,

$$\text{Relative error} \approx Ch^2,$$

where C is some constant depending on the choice of gauge field. This is in accordance with the estimate 4.44.

Quantum $SU(2)$ gauge theory Monte Carlo computer simulations using the SGT action was published in a companion article [17].

6 Conclusions

We have proposed a general formulation of lattice gauge theory on simplicial spacetime meshes. For any such constructed spacetime mesh based an arbitrarily shaped 3d simplicial mesh, this action can be used for lattice gauge theory simulations or to study the classical equations of motion. Traditionally, lattice QCD simulations within physics have used a homogeneous mesh. Mesh refinement is a well established concept within the subject of FEM. We feel that it is well worth the effort to investigate the possibility of mesh refinement within classical and quantum gauge theory.

Quantum lattice gauge theory simulations are very computer intensive. Therefore mesh refinement could be beneficial in cases where it makes sense to focus more computational effort on some subset of the simulation domain.

We have shown the consistency of this SGT numerical approximation to the continuous action, in the sense of approximation theory. The lattice gauge theory formalism is of such a complexity, that it makes sense to complement this theoretical proof with numerical “evidence”. This also serves as an interesting test of whether or not the calculated estimates are saturated in practice. We have provided this for a few different cases of $SU(2)$ gauge fields, for which the action was shown to converge towards the continuum result as the grid fineness increased. The generalization to the gauge group $SU(3)$ of QCD or other gauge groups is in principle a simple matter, although the computational complexity will depend strongly on this.

Acknowledgements

T. G. Halvorsen had funding from SPADE-ACE (Project 176891/V30, The Research Council of Norway). We would like to thank S. H. Christiansen for fruitful discussions.

References

1. R. W. B. Ardill, J. P. Clarke, J. M. Drouffe, and K. J. M. Moriarty. Quantum chromodynamics on a simplicial lattice. *Phys. Lett.*, B128:203, 1983.
2. Carl M. Bender and Kimball A. Milton. Approximate determination of the mass gap in quantum field theory using the method of finite elements. *Phys. Rev. D*, 34(10):3149–3155, Nov 1986.
3. Carl M. Bender, Kimball A. Milton, and David H. Sharp. Gauge invariance and the finite-element solution of the schwinger model. *Phys. Rev. D*, 31(2):383–388, Jan 1985.
4. Kevin E. Cahill and Randolph Reeder. Comparison of the simplicial method with Wilson’s Lattice Gauge Theory for $U(1)$ in three-dimensions. *Phys. Lett.*, B168:381, 1986.
5. Snorre H. Christiansen and Tore G. Halvorsen. A gauge invariant discretization on simplicial grids of the Schrödinger eigenvalue problem in an electromagnetic field. *E-print, UiO*, 2009.
6. Snorre H. Christiansen and Tore G. Halvorsen. A simplicial gauge theory. *ArXiv e-prints*, June 2010.
7. Snorre H. Christiansen and Ragnar Winther. On Constraint Preservation in Numerical Simulations of Yang–Mills Equations. *SIAM Journal on Scientific Computing*, 28(1):75–101, 2006.
8. Snorre Harald Christiansen, Hans Z. Munthe-Kaas, and Brynjulf Owren. Topics in structure-preserving discretization. *Acta Numerica*, 20:1–119, 2011.
9. Philippe G. Ciarlet. *The Finite Element Method for Elliptic Problems*, volume 4 of *Studies in mathematics and its applications*. North-Holland Publishing Company, 1. edition, 1978.
10. Michael Creutz. *Quarks, Gluons and Lattices*. Cambridge, Uk: Univ. Pr. (Cambridge Monographs On Mathematical Physics), 1986.

11. J. M. Drouffe and K. J. M. Moriarty. Gauge theories on a simplicial lattice. *Nucl. Phys.*, B220:253–268, 1983.
12. J. M. Drouffe and K. J. M. Moriarty. High-statistics study of the phase transition in U(2) four-dimensional simplicial lattice gauge theory. *Journal of Physics G: Nuclear Physics*, 10(10):L221, 1984.
13. J. M. Drouffe and K. J. M. Moriarty. U(2) four-dimensional simplicial lattice gauge theory. *Z. Phys.*, C24:395, 1984.
14. J. M. Drouffe, K. J. M. Moriarty, and C. N. Mouhas. Monte Carlo simulation of pure U(N) and SU(N) gauge theories on a simplicial lattice. *Comput. Phys. Commun.*, 30:249, 1983.
15. J. M. Drouffe, K. J. M. Moriarty, and C. N. Mouhas. U(1) four-dimensional gauge theory on a simplicial lattice. *J. Phys.*, G10:115, 1984.
16. Brian C. Hall. *Lie Groups, Lie Algebras, and Representations, An Elementary Introduction*. Springer, 2. edition, 2004.
17. Tore Gunnar Halvorsen and Torquil Macdonald Sørensen. Simplicial gauge theory and quantum gauge theory simulation. *Nuclear Physics B*, 854(1):166 – 183, 2012.
18. Ralf Hiptmair. Finite elements in computational electromagnetism. *Acta Numerica*, 11(-1):237–339, 2002.
19. Peter Monk. *Finite Element Methods for Maxwell's Equations*. Oxford Science Publications, reprinted edition, 2006.
20. Hans Munthe-Kaas and Brynjulf Owren. Computations in a free Lie algebra. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 357(1754):957–981, 1999.
21. Steven Weinberg. *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge, UK: Univ. Pr., 1995.
22. Steven Weinberg. *The quantum theory of fields. Vol. 2: Modern applications*. Cambridge, UK: Univ. Pr., 1996.
23. Hassler Whitney. *Geometric integration theory*. Princeton University Press, Princeton, N. J., 1957.
24. Kenneth G. Wilson. Confinement of quarks. *Phys. Rev. D*, 10(8):2445–2459, Oct 1974.
25. Chen-Ning Yang and Robert L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 96:191–195, 1954.

Chapter 10

Simplicial gauge theory and quantum simulation

Published as Nuclear Physics B 854 (2012) 166-183.



Available online at www.sciencedirect.com

SciVerse ScienceDirect

NUCLEAR PHYSICS B

Nuclear Physics B 854 [FS] (2012) 166–183

www.elsevier.com/locate/nuclphysb

Simplicial gauge theory and quantum gauge theory simulation

Tore Gunnar Halvorsen ^a, Torquil Macdonald Sørensen ^{b,*}

^a Department of Mathematical Sciences, Norwegian University of Science and Technology,
NO-7491 Trondheim, Norway

^b Centre of Mathematics for Applications, University of Oslo, NO-0316 Oslo, Norway

Received 8 July 2011; accepted 17 August 2011

Available online 24 August 2011

Abstract

We propose a general formulation of simplicial lattice gauge theory inspired by the finite element method. Numerical tests of convergence towards continuum results are performed for several $SU(2)$ gauge fields. Additionally, we perform simplicial Monte Carlo quantum gauge field simulations involving measurements of the action as well as differently sized Wilson loops as functions of β .

© 2011 Elsevier B.V. All rights reserved.

MSC: 35Q40; 74S05; 81T13; 81T25

Keywords: Lattice gauge theory; QCD; Finite element method; Simplicial mesh; Yang–Mills action

1. Introduction

1.1. General introduction

Gauge quantum field theory (QFT) has been extremely successful in modeling the behaviour of fundamental high energy particle physics. This is done using the standard model of particle physics, which is based on the gauge symmetry group $\mathcal{G} = U(1) \times SU(2) \times SU(3)$. Quantum gauge field theories based on such noncommutative gauge groups are also called Yang–Mills

* Corresponding author.

E-mail addresses: toregha@gmail.com (T.G. Halvorsen), t.m.sorensen@matnat.uio.no, torquil@gmail.com (T.M. Sørensen).

theories [1–4]. Despite the massive successes of this model, there are still large difficulties in calculating low energy properties of quarks and gluons. When restricting to these quantum fields, the standard model reduces to the theory of Quantum Chromodynamics (QCD), with gauge group $SU(3)$. The problem is that through the effect of renormalization, the QCD coupling constant increases as interaction energies decrease, in such a way that perturbation theory breaks down. This phenomenon is the source of confinement in QCD. Direct paper-and-pen calculation of masses and interactions among low energy bound states of quarks is therefore quite problematic.

1.2. Lattice gauge theory

By discretizing QCD onto a lattice, a lot of these difficulties are removed. Lattice gauge theory (LGT) [5,6] has proven itself to be a powerful method of doing nonperturbative gauge theory calculations. It has therefore been, still is, and will for a long time be immensely useful in testing QCD against experimental results at low energy.

Usually LGT models are formulated using a hypercubic lattice on a Euclidean spacetime. Such a mesh preserves some discrete subgroups of the translational, mirror and 4d rotational symmetries. Note that a clever way of retaining continuous symmetries while working on a lattice is to use random lattices [7–9].

The models are almost always defined so as to also preserve a discrete gauge symmetry. This has the beneficial effect of enforcing a vanishing gluon mass in the discrete model.

1.3. Simplicial lattices

Simplicial meshes have been used for QCD simulations before [10–16], with promising numerical results. Here, we construct a simplicial gauge theory (SGT) based on the general mathematical concept of a simplicial complex, while preserving gauge invariance. This allows us to define SGT on a very general class of meshes, without restricting ourselves to a particular type of simplicial lattice.

The construction of the gauge invariant SGT action functional is inspired by the finite element method (FEM) most commonly used for solving partial differential equations, particularly on complicated domains [17–21]. The formalism therefore includes the use of finite element function spaces on simplicial meshes, and the concept of mass matrices. The term “mass matrix” in this context has nothing to do with physical particle masses, and is therefore not to be confused with the usual mass matrix of quantum states within quantum field theory.

Through the use of the FEM formulation, and the massive resources of methods available within that subject area, we hope to gain advantages for QCD simulations in future implementations, in particular with regards to the possibilities of grid refinement. This could be useful in modeling some QCD phenomena, e.g. for highly concentrated gluon flux tubes between quarks where an increase lattice resolution might be desired. An earlier work used FEM inspired methods within QFT, although along a different direction involving solutions of operator equations instead of Monte Carlo simulations [22,23].

1.4. Computer simulation

The mathematical proof of consistency between the SGT and continuous Yang–Mills gauge theory action is described in a companion paper [24], along with a description of the more comprehensive Yang–Mills–Higgs model. In the current article we are content to provide numerical

evidence for convergence towards exact continuum results for several choices of gauge field configurations. In addition, we perform Monte Carlo quantum pure gauge field theory simulations for the gauge group $SU(2)$ in temporal gauge, as a proof-of-concept for SGT. Observable measurements include expectation values of the action density as well as a series of different Wilson loops.

1.5. Outline

Section 2 contains a short repetition of the fundamental definitions of gauge symmetry and the continuous spacetime Yang–Mills action in Section 2.1, the basics of traditional lattice gauge theory in Section 2.2, as well as an introduction to the proposed SGT action in Section 2.3. In Section 3.1, we report on the numerical convergence of the SGT action towards the exact continuum value for several different cases of $SU(2)$ gauge fields, as well as similar results from traditional LGT. Theoretical results proving consistency for general gauge fields can be found in [24]. In Section 3.2, we perform Monte Carlo quantum field theory simulations in order to observe that SGT correctly reproduces the basic aspects of the $SU(2)$ quantum field theory. We draw our conclusions in Section 4. Appendix A contains a short introduction to elementary aspects of simplicial complexes, and some notes about basis functions and mass matrices that are used in our construction of SGT. Appendix B contains a calculation of strong and weak coupling limits for a Wilson triangle and the action density. Lastly, Appendix C contains a short discussion of some aspects of the numerical computer implementation.

2. Construction

2.1. Continuous gauge theory

Consider the spacetime domain $\mathbb{M} = \mathbb{R} \times S$, where \mathbb{R} is time and $S \subset \mathbb{R}^3$. The domain \mathbb{M} represents either Lorentzian or Euclidean spacetime, in each case equipped with the appropriate metric. In the standard orthonormal \mathbb{M} -basis $\{e_\mu\}_{\mu=0,1,2,3}$, a general point $x \in \mathbb{M}$ has components $\{x^\mu\}_{\mu=0,1,2,3}$. Greek indices run from 0 to 3, and Latin indices from 1 to 3.

Furthermore, in this article we shall consider pure $SU(2)$ gauge theory. However, the construction presented is applicable to any gauge theory based on a compact Lie group \mathcal{G} which can be represented by a subgroup of the complex unitary $n \times n$ matrices. We define the real-valued scalar product on \mathcal{G} as

$$g' \cdot g := \Re \operatorname{tr}(g' g^H), \quad (1)$$

where g^H is the hermitian conjugate of a matrix g .

The connection between the continuous theory and the discrete simplicial theory is most easily seen in a coordinate free formulation. Thus, we start with a coordinate free formulation, before we give the more familiar coordinate based one.

The free variable in pure Yang–Mills theory with gauge Lie group \mathcal{G} is a gauge potential or more formally a one-form A on \mathbb{M} , with values in the corresponding gauge Lie algebra \mathfrak{g} . For simplicity of notation, we hereby specify $\mathcal{G} = SU(2)$ and $\mathfrak{g} = \mathfrak{su}(2)$. We split A into temporal and spatial components $A = (A_0, \mathbf{A})$. In this context, A_0 can be thought of as a scalar function,¹ and \mathbf{A} as a spatial vector. The curvature (field strength) of such a one-form is given by

¹ However, not a scalar in the sense of spacetime symmetry transformation properties.

$$F(A) = \text{d}A + \frac{i}{2}[A, A] = d_0\mathbf{A} + dA_0 + d\mathbf{A} + i[\mathbf{A}, A_0] + \frac{i}{2}[\mathbf{A}, \mathbf{A}], \quad (2)$$

where $\text{d} = (d_0, d)$, d_0 and d denote exterior derivative in the temporal and spatial directions respectively, and $[\cdot, \cdot]$ is the commutator between Lie algebra valued one-forms. We choose the basis $\{t^a\}_{a=1,2,3}$, where $t^a := \sigma^a/2$, for $\mathfrak{su}(2)$, where $\{\sigma^a\}_{a=1,2,3}$ are the Pauli matrices. Thus, we can expand the gauge field into components, $A = A^a t^a$. We also have

$$[A, A] = \sum_{ab} A^a \wedge A^b [t^a, t^b] = \sum_{abc} i \epsilon^{abc} A^a \wedge A^b t^c, \quad (3)$$

where ϵ^{abc} is the antisymmetric Levi-Civita symbol with $\epsilon^{123} = 1$ and \wedge is the wedge product (exterior product). For later convenience we split the curvature in a temporal and spatial part

$$F^t(A) = d_0\mathbf{A} + dA_0 + i[\mathbf{A}, A_0], \quad F^s(A) = d\mathbf{A} + \frac{i}{2}[\mathbf{A}, \mathbf{A}]. \quad (4)$$

The action that defines the gauge theory is the functional

$$S[A] = \frac{1}{4e^2} \int_{\mathbb{M}} |F(A)|^2 = \frac{1}{4e^2} \int_{\mathbb{M}} |F^t(A)|^2 + |F^s(A)|^2, \quad (5)$$

where the norms are generated by the metric and e is the dimensionless Yang–Mills coupling constant.

A gauge transformation is defined by a choice of $G(x) \in SU(2)$ for each $x \in \mathbb{M}$, and transforms the gauge field as

$$A_0 \mapsto G(A_0 + d_t)G^{-1}, \quad \mathbf{A} \mapsto G(\mathbf{A} + d)G^{-1}. \quad (6)$$

Note that the action $S[A]$ is invariant under such gauge transformations. For a more precise mathematical exposition, see [24].

A formulation more familiar within physics is obtained by expressing the one form and curvature in coordinates. In other words, one decomposes the one-form A^a in the basis $\{dx^\mu\}$, i.e. $A^a = \sum_\mu A_\mu^a dx^\mu$. The exterior derivative of such a one-form is given by

$$\text{d}A^a = \sum_{\mu\nu} \partial_\nu A_\mu^a dx^\nu \wedge dx^\mu = \sum_{\mu\nu} \frac{1}{2} (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a) dx^\mu \wedge dx^\nu. \quad (7)$$

Furthermore, the curvature is given by $F^a = \sum_{\mu\nu} \frac{1}{2} F_{\mu\nu}^a dx^\mu \wedge dx^\nu$, where

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - \epsilon^{abc} A_\mu^b A_\nu^c. \quad (8)$$

Finally, the action can be expressed as

$$S = \frac{1}{4e^2} \int_{\mathbb{M}} \sum_{\mu\nu a} F_{\mu\nu}^a F^{a\mu\nu} dx, \quad (9)$$

the usual coordinate dependent expression for the Yang–Mills action functional.

2.2. Lattice gauge theory

To see the connection between lattice gauge theory (LGT) and the simplicial gauge theory (SGT), we will in this section give a brief overview of the discretization procedure from LGT. For a more complete description see e.g. [6].

The discretization procedure of both LGT and SGT is based on the following identity. Consider a small surface Σ with area proportional to h^2 , where h is a small positive quantity. Then the following identity holds

$$\oint_{\Sigma} F(A) = \mathcal{H}(A) - 1 + \mathcal{O}(h^3),$$

where $\mathcal{H}(A)$ is the holonomy of the one-form A , i.e. the parallel transport induced by A around the boundary of Σ . This parallel transport is defined as follows. Given a curve $\gamma : [0, 1] \rightarrow \mathbb{M}$, such that $\gamma(0) = x$ and $\gamma(1) = y$, the parallel transport operator along γ is given by

$$U_{\gamma}(x, y) = P \left(\exp \left(i \int_{\gamma} A \right) \right),$$

where P denotes path-ordering, and the subscript γ is attached to U to denote the path-dependence. In LGT, this quantity is known as the Wilson line.

In LGT, spacetime \mathbb{M} is usually discretized by a uniform hypercubic lattice \mathbb{L} . Neighbouring node positions are related through translation vectors $\{a_{\mu}\}$ for which we assume $|a_{\mu}| = h$ for all μ . To each edge e which connects neighbouring nodes, n and $n + a_{\mu}$ for some μ , we attach an approximation of the parallel transport operator along e . Thus,

$$U_{\mu}(n) := \exp \left(i h A_{\mu} \left(n + \frac{1}{2} a_{\mu} \right) \right) \approx U_e(n, n + a_{\mu}) = P \left(\exp \left(i \int_n^{n+a_{\mu}} A \right) \right). \quad (10)$$

In LGT this quantity is called a link variable, link matrix or link group element. Furthermore, given a face f of a cube in the mesh, called a plaquette, we approximate the holonomy associated to this face as the path-ordered product of the link variables along its boundary. In other words, if f lies in the $\mu\nu$ plane, with nodes $n, n + a_{\mu}, n + a_{\nu}$, and $n + a_{\mu} + a_{\nu}$, we approximate the holonomy as

$$\begin{aligned} U_f(n) &:= U_{\mu\nu}(n) := U_{\mu}(n) U_{\nu}(n + a_{\mu}) U_{\mu}^H(n + a_{\nu}) U_{\nu}^H(n) \\ &\approx \mathcal{H}(A) := P \left(\exp \left(i \oint_{\partial f} A \right) \right), \end{aligned} \quad (11)$$

where ∂f denotes the boundary of the plaquette f . In LGT, this quantity is known as the Wilson loop. Moreover, we approximate the curvature as

$$F_{\mu\nu}^f \approx U_f - \mathbb{1}. \quad (12)$$

Finally, the LGT action is defined as

$$S_{LGT} = \beta \sum_f \frac{1}{4} \text{tr}[(U_f - \mathbb{1})(U_f - \mathbb{1})^H] = \beta \sum_f 1 - \frac{1}{4} \text{tr}(U_f + U_f^H), \quad (13)$$

where β is related to the coupling constant by $\beta = 4/e^2$. A discrete gauge transformation is associated with a choice of $G(n) \in SU(2)$ for each node n . Each link variable then transforms as

$$U_\mu(n) \mapsto G(n)U_\mu(n)G(n + a_\mu)^{-1}. \quad (14)$$

By the cyclic invariance of the trace, the action S_{LGT} is discretely gauge invariant.

2.2.1. Remarks

The LGT action can be viewed as a mass lumped FEM action, and this observation is useful to have in mind when we construct the simplicial analogue. In the FEM setting, the gauge potential is assumed to be a lowest order curl-conforming Nédélec element in 4d on hypercubes, with one dimension representing time [19]. The degree of freedom associated to such a gauge potential at an edge e from n to $n + a_\mu$ is

$$A_e = \int_n^{n+a_\mu} A = h A_\mu \left(n + \frac{1}{2} a_\mu \right).$$

The parallel transport operator is as in Eq. (10), i.e. $U_\mu(n) = \exp(i A_e)$. Then, the holonomy is approximated as in Eq. (11), the curvature as in Eq. (12), and one considers $U_f - \mathbb{1}$ as the components of the two-form

$$\sum_f (U_f - \mathbb{1}) \omega_f,$$

where $\{\omega_f\}$ are the Nédélec basis two-forms. The FEM action associated to such a two-form is

$$S := \frac{\beta}{2} \sum_{f, f'} M_{ff'} \text{tr}[(U_f - \mathbb{1})(U_{f'} - \mathbb{1})^H], \quad M_{ff'} := \int_{\mathbb{M}} \omega_f \cdot \omega_{f'},$$

where $M_{ff'}$ is called the mass matrix, and (\cdot) denotes the scalar product of alternating forms w.r.t. the metric. The mass matrix is not diagonal, which means that the discrete curvature at different faces interact. This again implies that the action is not discretely gauge invariant. However, by diagonalizing the mass matrix using numerical quadrature, this action reduces to the LGT action, Eq. (13). The diagonalization procedure can also be shown to be numerically consistent in the sense of approximation theory [24].

2.3. Simplicial gauge theory

In this section we construct the discretely gauge invariant simplicial gauge theory (SGT) action on a simplicial complex, as defined in Appendix A. The construction is the simplicial analogue of the FEM action described above, including additional parallel transport operators to make it discretely gauge invariant.

The curvature associated to the temporal and spatial faces is defined exactly as in LGT. In the notation of Appendix A, consider a temporal and spatial face

$$\begin{aligned} f_t(\tau) &:= \{i_\tau, j_\tau, j_{\tau+\Delta t}, i_{\tau+\Delta t}\}, \\ f(\tau) &:= \{i_\tau, j_\tau, k_\tau\}, \end{aligned} \quad (15)$$

where i_τ denotes node i at time τ . The time-dependency will from here on often be suppressed, unless confusion can arise. The spatial and temporal holonomies associated to these faces, induced by the gauge potential, are approximated as

$$\begin{aligned} U_{f_t}(i_\tau) &= U(i_\tau, j_\tau)U(j_\tau, j_{\tau+\Delta t})U(j_{\tau+\Delta t}, i_{\tau+\Delta t})U(i_{\tau+\Delta t}, i_\tau), \\ U_f(i) &= U(i, j)U(j, k)U(k, i), \end{aligned} \quad (16)$$

where the arguments i_τ and i are included to indicate where the holonomy is located, and the parallel transport operators are defined exactly as in LGT, i.e. Eq. (10). We observe that the holonomies located at different nodes are related through the formulas

$$\begin{aligned} U_{f_t}(i_\tau + \Delta t) &= U(i_{\tau+\Delta t}, i_\tau)U_{f_t}(i_\tau)U(i_\tau, i_{\tau+\Delta t}), \\ U_f(j) &= U(j, i)U_f(i)U(i, j), \end{aligned}$$

which give formulas for parallel transport of curvature. Hence, we have defined the curvature associated to the temporal and spatial faces in our 4d mesh. The distinguished point of f and f_t , i.e. the location of their holonomy, are denoted \dot{f} and \dot{f}_t respectively. Note that under a discrete gauge transformation, the parallel transport operators are transformed as in LGT, i.e.

$$\begin{aligned} U(i_\tau, i_{\tau+\Delta t}) &\mapsto G(i_\tau)U(i_\tau, i_{\tau+\Delta t})G(i_{\tau+\Delta t})^{-1}, \\ U(i, j) &\mapsto G(i)U(i, j)G(j)^{-1}, \end{aligned}$$

for $G(i) \in SU(2)$ for each vertex i .

As in LGT the curvature is approximated as

$$\begin{aligned} F^t &\approx U_{f_t} - \mathbb{1}, \\ F^s &\approx U_f - \mathbb{1}, \end{aligned} \quad (17)$$

considered as components of the two-forms

$$\begin{aligned} \sum_{f_t} (U_{f_t} - \mathbb{1}) \Lambda_{f_t}, \\ \sum_f (U_f - \mathbb{1}) \Lambda_f, \end{aligned}$$

where the Λ are basis functions as described in Appendix A. The associated FEM action is $S = S_t + S_s$, where the temporal part is

$$S_t = \frac{\beta}{2} \Re \sum_{f_t, f'_t} M_{f_t f'_t} \text{tr}[(U_{f_t} - \mathbb{1})(U_{f'_t} - \mathbb{1})^H], \quad M_{f_t f'_t} := \int_{\mathbb{M}} \Lambda_{f_t} \cdot \Lambda_{f'_t}, \quad (18)$$

and the spatial part is

$$S_s = \frac{\beta}{2} \Re \sum_{f, f'} M_{ff'} \text{tr}[(U_f - \mathbb{1})(U_{f'} - \mathbb{1})^H], \quad M_{ff'} := \int_{\mathbb{M}} \Lambda_f \cdot \Lambda_{f'}, \quad (19)$$

where $\beta = 2/e^2$. Note that we have suppressed the dependency of S on A . Again, $M_{f_t f'_t}$ and $M_{ff'}$ are called mass matrices. They depend on the details of the mesh, and are described more in detail in Appendix A. As pointed out in the discussion about the FEM formulation of LGT, the mass matrices are not diagonal. This implies that the action is not discretely gauge invariant. However, this can be resolved by parallel transport of curvature. The temporal and spatial part of the action, S_t and S_s , are now treated separately.

2.3.1. The temporal part

Let $f_t(\tau)$ and $f'_t(\tau)$ be two temporal faces. We now use some properties of the basis functions, which are explained in [Appendix A](#). Since the temporal basis face functions (Λ_{f_t}) are piecewise constant in time, the interactions between the temporal curvature occur only at coinciding time intervals. Also, by properties of the edge basis functions (λ_e), which define the temporal basis face functions, we can connect the curvature at f_t with the curvature at f'_t by parallel transport along at most one edge. Thus, we connect the curvatures by parallel transport along the connecting edge $e = \{\dot{f}_t, \dot{f}'_t\}$ of their distinguished points. In other words, we approximate the temporal part of the action by

$$S_{SGT}^t := \frac{\beta}{2} \Re \sum_{f_t(\tau), f'_t(\tau)} M_{f_t(\tau), f'_t(\tau)} \operatorname{tr}[U(\dot{f}'_t, \dot{f}_t)(U_{f_t(\tau)} - \mathbb{1})U(\dot{f}_t, \dot{f}'_t)(U_{f'_t(\tau)} - \mathbb{1})^H]. \quad (20)$$

2.3.2. The spatial part

Let f and f' be two spatial faces of a tetrahedron T . The curvature associated to the face f at time τ will interact with the curvature associated to the face f' not only at time τ , but also at times $\tau \pm \Delta t$, since the facial basis functions are piecewise affine in time. Thus, to connect the curvature at $f(\tau)$ with the curvature at $f'(\tau')$ we must parallel transport in both space and time. Thus, we replace

$$(U_{f(\tau)} - \mathbb{1})(U_{f'(\tau')} - \mathbb{1})^H$$

by

$$\begin{aligned} & U(\dot{f}'(\tau), \dot{f}(\tau))(U_{f(\tau)} - \mathbb{1})U(\dot{f}(\tau), \dot{f}'(\tau))U(\dot{f}'(\tau), \dot{f}'(\tau'))(U_{f'(\tau')} - \mathbb{1})^H \\ & \times U(\dot{f}'(\tau'), \dot{f}'(\tau)) \end{aligned}$$

in the FEM action [\(19\)](#). In words, we first parallel transport the curvature associated to f , located at the vertex $\dot{f}(\tau)$ to the vertex $\dot{f}'(\tau)$ along the edge $e = \{\dot{f}(\tau), \dot{f}'(\tau)\}$. Then we parallel transport it in the temporal direction from $\dot{f}'(\tau)$ to $\dot{f}'(\tau')$. So, we approximate the spatial part of the action as

$$\begin{aligned} S_{SGT}^s := & \frac{\beta}{2} \Re \sum_{f(\tau), f'(\tau')} M_{f(\tau), f'(\tau')} \operatorname{tr}[U(\dot{f}'(\tau), \dot{f}(\tau))(U_{f(\tau)} - \mathbb{1})U(\dot{f}(\tau), \dot{f}'(\tau)) \\ & \times U(\dot{f}'(\tau), \dot{f}'(\tau'))(U_{f'(\tau')} - \mathbb{1})^H U(\dot{f}'(\tau'), \dot{f}'(\tau))]. \end{aligned} \quad (21)$$

The simplicial gauge theory action is then defined as

$$S_{SGT} := S_{SGT}^t + S_{SGT}^s, \quad (22)$$

and by the cyclic invariance of the trace, this action is discretely gauge invariant. A companion paper [\[24\]](#) contains more details about this construction, as well as mathematical proofs of consistency with the continuous action [\(5\)](#) in the sense of approximation theory.

3. Computer simulation

For our SGT computer simulations, we chose the Euclidean cubic domain $\mathbb{M} = [0, 1]^4 \subset \mathbb{R}^4$ with periodic boundary conditions. We simulated the pure gauge SGT action [\(22\)](#) in temporal

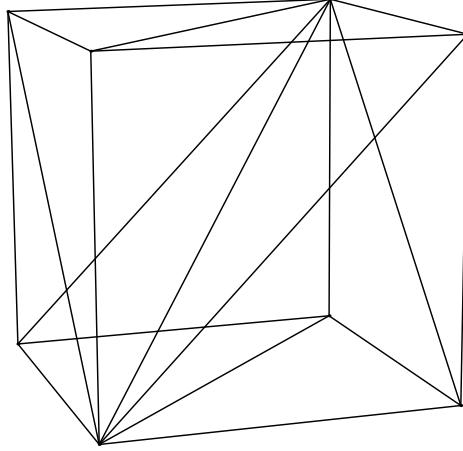


Fig. 1. Elementary 3d mesh building block containing six tetrahedra, all of which share the single interior diagonal. This particular choice implies an anisotropy in the discretization. This anisotropy will of course disappear in the continuum limit.

gauge on a simplicial lattice with the gauge group $SU(2)$. Choice of gauge is not necessary, but it does simplify the algorithm slightly, since all temporal edge matrices then reduce to the identity.

The spatial lattice was constructed using a cubic arrangement of N^3 identical building block cubes of size h^3 , each consisting of six tetrahedra as shown in Fig. 1. The resulting spatial mesh was repeated at N consecutive time steps to form a cubic domain of physical volume $(hN)^4$. As described above, each spatial edge is part of two temporal square-shaped faces, going forward and backward in time.

The SGT action employs parallel transport matrices in order for gauge invariance to be respected. By defining the distinguished points of all spatial and temporal faces to coincide for as many pairs of faces as possible, we only need the parallel transport matrices for terms in the action involving pairs of temporal faces with no common nodes. More details regarding the exact computer implementation are given in Appendix C.

3.1. Convergence of the action

In order to check the continuum limit of the discrete action, we examined four different gauge field configurations for which the exact continuum value S_{cont} of the action is calculable. We did numerical calculations for square meshes with $N = 4, 8, 16, 32$ in order to observe convergence of the numerical values towards the exact values. By the estimates in [24] we expect that the error be of second order in the lattice constant h . We used the following gauge field configuration cases:

1. Gauge field oriented towards the x -direction in space and towards t^3 within $\mathfrak{su}(2)$, with a sinusoidal t -dependence. The only nonzero component of the gauge field A is

$$A_x^3(t, x, y, z) := \frac{e}{2\pi} \sin(2\pi t), \quad S = 1.$$

2. Gauge field oriented towards the y -direction in space and t^3 within $\mathfrak{su}(2)$, with a sinusoidal x -dependence. The nonzero component of the gauge field in this case was

$$A_y^3(t, x, y, z) := \frac{e}{2\pi} \sin(2\pi x), \quad S = 1.$$

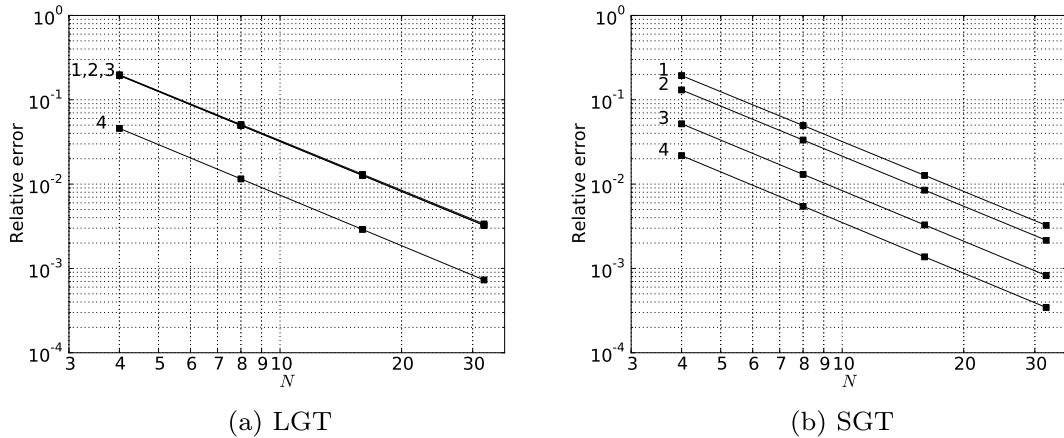


Fig. 2. The relative error of the action versus the number of lattice sites per side N , for the actions 1, 2, 3, 4 described in Section 3.1. The squares are the simulation data points and the solid lines are the second order polynomial fits. Errors are proportional to h^2 in all cases.

3. A case with two nonzero components,

$$A_x^1 := \frac{e}{2\pi} \sin(2\pi y), \quad A_y^2 := \frac{e}{2\pi} \sin(2\pi x), \quad S = \frac{1}{2} + \frac{e^2}{8(2\pi)^4}.$$

4. A constant field that only contributes to the nonlinear term in the field strength,

$$A_x^1 := \sqrt{e}, \quad A_y^2 := \sqrt{e}, \quad S = \frac{1}{2}.$$

In order to provoke a sizable nonlinear contribution in case 3, we chose a small $\beta = 2/e^2 = 1/5$. The link matrices needed to evaluate the SGT action are calculated from these gauge fields by means of the exponential map (10).

The results are displayed using double logarithmic plots in Fig. 2 for traditional Wilson action LGT as well as the SGT results. As expected from the estimates in [24], in all cases the relative error behaves as

Relative error $\sim Ch^2$,

as determined by extracting the linear coefficient of the second order polynomial fits shown in the figures. Note that while the convergence exponent of h is the same in all cases, the prefactor C is smaller in the SGT cases involving time-independent fields, due to its finer spatial discretization for the same N . Where time-dependence is involved, the errors coincide since the time-discretization we have chosen for this SGT simulation is of the same quality as for the LGT simulation.

3.2. Quantum field simulation

Analogous to the traditional lattice QCD simulations, we performed parallel $SU(2)$ quantum field theory Monte Carlo simulations for $N = 8$. In this case, the edge matrices are sampled directly without reference to a gauge field and lattice constant value. Therefore, the physical size of the simulation domain is unknown prior to experimental comparisons. All dimensional observable quantities are automatically calculated in units of powers of the lattice constant h .

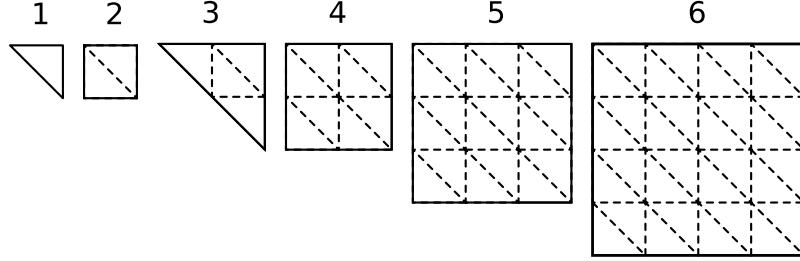


Fig. 3. The simulated Wilson loops shapes correspond to the outer edges of these figures. They lie in the xy , yz and zx planes.

As is customary, it is a Monte Carlo simulation using the Metropolis algorithm to generate a Markov chain of gauge field configurations that are distributed according to the Boltzmann weight $\exp(-S)$. Each Monte Carlo step involves randomization of some edge $SU(2)$ matrices, which is done by multiplication of a small $\mathfrak{su}(2)$ algebra matrix, together with a Metropolis step for acceptance/rejection of the update. The algorithm adapted itself to drive the MC acceptance rate towards 1/2. The use of temporal gauge may slow the convergence of this type of numerical simulation. However, Monte Carlo convergence was ascertained and high quality error estimates were made by the use of data blocking [25]. In addition, convergence was verified subjectively by inspection of the time series for observable values with their accompanying distributions, as well as time series for cumulative averages.

The data blocking error estimates were found to be smaller than the displayed data points in all the plots.

We simulated at different values of β , at each of which we measured the average action density S/N^4 , and a list of different Wilson loops shown in Fig. 3, all of which are gauge-invariant quantities. For each Wilson loop shape, we average over all possible loop positions, as well as loop orientations in the xy , yz and zx planes. For a given closed path \mathcal{C} , the corresponding Wilson loop variable for gauge group $SU(n)$ is defined as

$$W_{\mathcal{C}} := \frac{1}{2} \Re \operatorname{tr} \prod_{e \in \mathcal{C}} U_e, \quad (23)$$

which involves an ordered product of the edge matrices $\{U_e\}$ along the path \mathcal{C} .

Expectation values for any observable quantity \mathcal{O} , e.g. the action density S/N^4 or a Wilson loop $W_{\mathcal{C}}$, is given by

$$\langle \mathcal{O} \rangle = \frac{1}{Z} \int \left(\prod_e dU_e \right) \mathcal{O} \exp(-S), \quad (24)$$

where the partition function Z is defined by

$$Z := \int \left(\prod_e dU_e \right) \exp(-S). \quad (25)$$

The integration measure involved in these expressions is a product of the normalized Haar integration measure for each edge group element in the mesh. Note that the normalized Haar measure satisfies

$$\int_{\mathcal{G}} dU = 1. \quad (26)$$

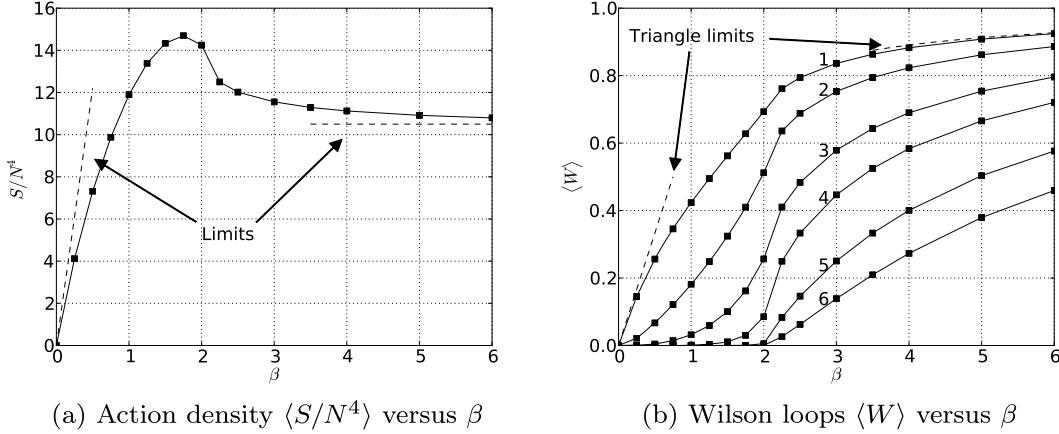


Fig. 4. Plots showing the β -dependency of (a) the average action density $\langle S/N^4 \rangle$ and (b) the various Wilson loops $\langle W \rangle$ from Fig. 3. Solid squares are data points and solid lines are linear interpolations. The strong and weak coupling asymptotes are included for the action density and the elementary triangular loop. Monte Carlo errors are smaller than the data points.

To accompany these measurements, the strong (small β) and weak (large β) coupling asymptotic behaviour were calculated in Appendix B, using methods described in [6]. At strong coupling, this involves various group integrals, while at weak coupling it suffices to use a thermodynamic analogy to determine the limiting behaviour.

The simulated results for the action density and Wilson loops are displayed in Fig. 4. In Fig. 4(a) we can see the characteristic and nontrivial behaviour in the medium coupling range $\beta \in (1, 3)$. This coincides qualitatively with LGT simulations [26]. Only qualitative, not exact, agreement is expected, since the physical lattice constant will differ in each type of simulation. Compared to LGT simulations, the behaviour at small β deviates more from linearity due to the nonlinear aspects of the SGT action. In this region, the actions do not approximate the continuum action, and differences between discrete actions are unphysical.

The Wilson loops in Fig. 4(b) show the same qualitative behaviour as do LGT simulation results, and approaches the calculated asymptotes nicely. Also here, the behaviour is less linear at small β for the same reason as stated above. The typical strong suppression of the Wilson loops as functions of loop area is reproduced, as expected from the area law behaviour that indicates confinement.

4. Conclusions

We have implemented the general SGT action on a particular simplicial mesh, and performed Monte Carlo quantum field theory simulations that show sensible results that are qualitatively consistent with standard LGT simulations, as must be the case for this initial proof-of-concept implementation.

We expect that this method will lend itself nicely to the use of mesh refinement within quantum QCD simulations, and that this will lead to opportunities of novel applications using nontrivial mesh structures, e.g. in the vicinity of gluon flux tubes as mentioned in the introduction.

The nondiagonal nature of the action increases the amount of computer work in the Metropolis step after each proposed update. However, since the number of interactions for each elementary face is finite, the scaling at large meshes for this model will be the same as for traditional QCD.

There might be possibilities of real-time adaptive diagonalization, thereby increasing the algorithm efficiency throughout the initial part of the simulation.

Acknowledgements

T.G. Halvorsen had funding from SPADE-ACE (Project 176891/V30, The Research Council of Norway). We would like to thank S.H. Christiansen for fruitful discussions.

Appendix A. Simplicial complex, finite elements and mass matrices

Consider a collection of vertexes, edges, faces, tetrahedra in 3d space. These elementary objects are called simplexes, and the collection of these a simplicial complex \mathcal{T} . For any k -dimensional simplex T_k for $1 \leq k \leq 3$, the boundary ∂T_k is a union of $(k - 1)$ -dimensional simplexes. Consult [27, Section 5.1] for a precise definition. In our construction, we assume that this spatial simplicial complex spans the spatial domain S . The vertexes, edges, faces, and tetrahedra according to dimension, and are labeled i , e , f , and T respectively. The symbol T will be used for simplexes of any dimension.

In order to expand this to a 4d spacetime simplicial complex \mathbb{T} , consider a uniform time-discretization with a time spacing Δt . The simplicial complex \mathcal{T} is then repeated at each discrete time step value τ . For each such τ , we define additional simplexes for our \mathbb{T} by extruding each simplex of \mathcal{T} along the time interval $[\tau, \tau + \Delta t]$. As the basic building block in classical 3d FEM theory is a tetrahedron T , the basic building block in this extended FEM version is $T \times I_\tau$, where $I_\tau = [\tau, \tau + \Delta t]$, i.e. a time-extrusion of a tetrahedron. Temporal edges are generated by extruding 3d vertices, and temporal faces by extruding 3d edges.

The space of Whitney k -forms on $\mathcal{T}(T)$ is denoted $W^k(\mathcal{T})$ ($W^k(T)$), with canonical basis (λ_T) , T ranging over the set of k -dimensional simplexes in \mathcal{T} [20]. The 0-forms λ_i are the barycentric coordinate maps for each vertex i . In other words, it is the piecewise affine map taking the value 1 at the vertex i and 0 at other vertices. For an edge $e = \{i, j\}$, with orientation $i \rightarrow j$, the associated Whitney 1-form is defined by

$$\lambda_e := \lambda_{ij} := \lambda_i \nabla \lambda_j - \lambda_j \nabla \lambda_i. \quad (\text{A.1})$$

For a face $f = \{i, j, k\}$, whose orientation is $i \rightarrow j \rightarrow k$, the associated Whitney 2-form is defined by

$$\lambda_f := \lambda_{ijk} := 2(\lambda_i \nabla \lambda_j \times \nabla \lambda_k + \lambda_j \nabla \lambda_k \times \nabla \lambda_i + \lambda_k \nabla \lambda_i \times \nabla \lambda_j). \quad (\text{A.2})$$

In the 4d spacetime FEM setting, these basis k -forms are extended to be piecewise affine in time and are denoted $(\Lambda_{T(\tau)})$, i.e.

$$\lambda_T \rightarrow \Lambda_{T(\tau)} = \lambda_T \otimes P_1^t,$$

where P_1^t denotes polynomials in the time variable of degree at most one, and $T(\tau) := (\tau, T)$ denotes the spatial simplex T at temporal node τ . More precisely, $\Lambda_{T(\tau)}$ is the piecewise affine function in time, taking the value λ_T at τ and 0 at the other temporal nodes. In addition, we define temporal basis edge and face functions.

To every vertex i in the spatial mesh there are temporal edges $e_t(\tau) = \{i_\tau, i_{\tau+\Delta t}\}$, where $i_\tau := i(\tau)$. The temporal basis edge function attached to $e_t(\tau)$ is then the piecewise constant function in time defined by

$$\Lambda_{e_t(\tau)}(t) = \begin{cases} \lambda_i \circ \pi \frac{1}{\Delta t} dt, & t \in [\tau, \tau + \Delta t], \\ 0, & \text{otherwise,} \end{cases}$$

where π is the canonical projection onto the space S ,

$$\pi : \mathbb{M} = \mathbb{R} \times S \rightarrow S,$$

and dt is the standard basis one-form in the temporal direction.

To every spatial edge e there are corresponding temporal faces $f_t(\tau) = e \times I_\tau$. The temporal basis face function attached to $f_t(\tau)$ is then the piecewise constant function in time defined by

$$\Lambda_{f_t(\tau)}(t) = \begin{cases} \lambda_e \circ \pi \wedge \frac{1}{\Delta t} dt, & t \in [\tau, \tau + \Delta t], \\ 0, & \text{otherwise.} \end{cases}$$

In addition to these basis functions, we must define mass matrix elements. Let $m_{TT'}$ denote the classical 3d mass matrices for spatial Whitney elements

$$m_{TT'} = \int_S \lambda_T \cdot \lambda_{T'},$$

where T, T' are k -dimensional simplexes, and (\cdot) denotes the scalar product of alternating forms.

In the definition of the SGT action we use the generalization

$$M_{T(t)T'(\tau)} = \int_{\mathbb{M}} \Lambda_{T(t)} \cdot \Lambda_{T'(\tau)}.$$

This generalization can be expressed through the classical mass matrices by performing the time integration explicitly. Thus, let T be a spatial tetrahedron and $I_\tau = [\tau, \tau + \Delta t]$. Considering now only this time interval, the piecewise affine function taking the value 1 at time τ and 0 at time $\tau + \Delta t$ is given by

$$p_\tau(t) = 1 - \frac{t - \tau}{\Delta t}.$$

The analogous function for the temporal node $\tau + \Delta t$ on the same time interval is given by

$$p_{\tau+\Delta t}(t) = \frac{t - \tau}{\Delta t}.$$

Restricted to the basic building block $T \times I_\tau$, we therefore get

$$\begin{aligned} M_{f(\tau)f'(\tau)}(T \times I_\tau) &= \int_{T \times I_\tau} \Lambda_{f(\tau)} \cdot \Lambda_{f'(\tau)} = \int_{I_\tau} p_\tau^2 \int_T \lambda_f \cdot \lambda_{f'} = \frac{1}{3} \Delta t m_{ff'}(T), \\ M_{f(\tau)f'(\tau+\Delta t)}(T \times I_\tau) &= \int_{T \times I_\tau} \Lambda_{f(\tau)} \cdot \Lambda_{f'(\tau+\Delta t)} = \int_{I_\tau} p_\tau p_{\tau+\Delta t} \int_T \lambda_f \cdot \lambda_{f'} = \frac{1}{6} \Delta t m_{ff'}(T), \\ M_{f(\tau+\Delta t)f'(\tau+\Delta t)}(T \times I_\tau) &= \int_{T \times I_\tau} \Lambda_{f(\tau+\Delta t)} \cdot \Lambda_{f'(\tau+\Delta t)} = \int_{I_\tau} p_{\tau+\Delta t}^2 \int_T \lambda_f \cdot \lambda_{f'} = \frac{1}{3} \Delta t m_{ff'}(T). \end{aligned}$$

Similarly, the mass matrix element corresponding to the temporal face basis is given by

$$M_{f_t(\tau), f'_t(\tau)}(T \times I_\tau) = \int_{T \times I_\tau} \Lambda_{f_t(\tau)} \cdot \Lambda_{f'_t(\tau)} = \frac{1}{\Delta t} \int_T \lambda_e \cdot \lambda'_e = \frac{1}{\Delta t} m_{ee'}(T).$$

Appendix B. Strong and weak coupling limits

B.1. Strong coupling limit

Here we will show some details regarding the calculation of the strong coupling limits of the elementary triangular Wilson loop. We will use the following integrals over $SU(2)$ group space [6]

$$\begin{aligned} \int dU U^{\alpha\beta} &= 0, & \int dU U^{\alpha_1\beta_1} U^{\dagger\beta_2\alpha_2} &= \frac{1}{2} \delta^{\alpha_1\alpha_2} \delta^{\beta_1\beta_2}, \\ \int dU U^{\alpha_1\beta_1} U^{\alpha_2\beta_2} &= \frac{1}{2} \epsilon^{\alpha_1\alpha_2} \epsilon^{\beta_1\beta_2}, \end{aligned} \quad (\text{B.1})$$

where the Greek symbols are matrix indices.

In this calculation, the Wilson loop encircles an elementary spatial triangular plaquette P_t at time t . We denote this Wilson loop by W_{P_t} . By Eq. (23), it is given by

$$W_{P_t} := \frac{1}{2} \Re \operatorname{tr}(U_a U_b U_c),$$

where the plaquette P_t is encircled cyclically by the $SU(2)$ edge matrices U_a, U_b and U_c . Due to our choice of distinguished points and plaquette orientations, the spatial SGT action is given by

$$S = \frac{\beta}{2} \sum_{f,f'} M_{ff'} \operatorname{tr}(U_f U_{f'}^H - U_f - U_{f'}^H + \mathbb{1}),$$

where the sum extends over all spatial faces at all times. Since we are interested in small β , consider a first order truncated Taylor expansion of the exponential in Eq. (24), i.e.

$$\langle W_{P_t} \rangle \approx \frac{-\beta}{4Z_\beta} \int \left(\prod_e dU_e \right) \Re \operatorname{tr}(U_a U_b U_c) \sum_{f,f'} M_{ff'} \operatorname{tr}(U_f U_{f'}^H - U_f - U_{f'}^H + \mathbb{1}).$$

By the properties of the $SU(2)$ integration measure, terms involving integration over odd powers of link matrices vanish. Therefore, nonvanishing contributions to the integral only come from terms where either f and/or f' coincide with the plaquette P_t . The $U_f U_{f'}^H$ doesn't contribute. Indeed, if either f or f' differ from P_t , we such a term includes an integral over a single power, which vanishes. If on the other hand $f = f' = P_t$, we have $U_f U_f^H = \mathbb{1}$ which again leads to an integral over a single power and thus vanishes. This is also the case for the constant term in the parenthesis.

We are left with

$$\langle W_{P_t} \rangle \approx \frac{\beta}{4Z_\beta} \Re \int \left(\prod_e dU_e \right) \operatorname{tr}(U_a U_b U_c) \sum_{f,f'} M_{ff'} \operatorname{tr}(U_f + U_{f'}^H),$$

where we have moved the real part operator \Re outside of the integral. Contributions only come when at least one of f, f' coincide with P_t . Therefore, by the properties of the particular mesh we have constructed,

$$\begin{aligned}\langle W_{P_t} \rangle &\approx \frac{\beta}{4Z_\beta} (M_{P_t P_t} + M_{P_t P_{t+1}} + M_{P_t P_{t-1}}) \\ &\times \Re \int \left(\prod_e dU_e \right) \text{tr}(U_a U_b U_c) \text{tr}(U_{P_t} + U_{P_t}^H).\end{aligned}$$

Using $U_{P_t} := U_a U_b U_c$ and the $SU(2)$ integration formulas (B.1), we get

$$\langle W_{P_t} \rangle \approx \frac{\beta}{2} (M_{P_t P_t} + M_{P_t P_{t+1}} + M_{P_t P_{t-1}}) = \frac{2}{3} \beta,$$

where we have used $Z \approx 1$ for small β . The last equality follows from the particular mass matrix element values produced by our choice of simplicial lattice.

A similar calculation, only slightly more involved because several faces are involved, can be performed to determine the strong coupling limit of the action. Approximations of higher order in β can be found by including higher order terms in the Taylor expansion of the exponential.

B.2. Weak coupling

In order to determine the weak coupling limit of the action density, we simple follow a thermodynamic analogy described in [6]. At large β , the system is described well by a Gaussian partition function approximation. This corresponds to a free theory, and we can find the weak coupling limit of the action by distributing an amount $kT/2 = 1/2\beta$ of energy among all the degrees of freedom in the theory. We have seven edges for each building block cube, each of which contributes three degrees of freedom (the number of generators of $SU(2)$). In accordance with our use of temporal gauge in the simulations, we have excluded the unphysical temporal components when counting degrees of freedom. To obtain the action, we multiply by β , which results in

$$S_{SGT} \rightarrow \beta \times \frac{1}{2\beta} \times 7 \times 3 = \frac{21}{2} N^4, \quad \text{as } \beta \rightarrow \infty. \quad (\text{B.2})$$

This result can be used to determine the same limit of the triangular Wilson loop in the $\alpha\beta$ plane. We have

$$\langle W_1 \rangle = 1 - \frac{a^4}{16} \langle \text{tr}(F_{\alpha\beta}^2) \rangle,$$

where there is no sum over the spacetime indices. The antisymmetric field strength has six independent spacetime components. By the equipartitioning of the Euclidean energy among these degrees of freedom, we have

$$\langle \text{tr}(F_{\alpha\beta}^2) \rangle = \frac{1}{6} \langle \text{tr}(F_{\mu\nu} F^{\mu\nu}) \rangle = \frac{2g^2}{6} \left\langle \frac{S_{SGT}}{N^4} \right\rangle = \frac{42g^2}{12}.$$

Now using $\beta = 2/g^2$, we get

$$\langle W_1 \rangle = 1 - \frac{21}{48\beta}. \quad (\text{B.3})$$

Appendix C. Computer implementation

Our computer implementation of the simplicial lattice and accompanying SGT action consists of object-oriented C++ code, using MPICH2 [28] for parallelization, running on a quadruple CPU run-of-the-mill modern workstation computer. The data structures involved are reminiscent of what is used in implementations of the finite element method. This involves different types of mass matrix and connectivity information for elements of the simplicial mesh. The parallelization consisted of running independent simulations on each node, and averaging the results. We used the *yarn2* algorithm from the TINA pseudo-random number generator [29], which is designed for use in parallelized algorithms. Although the edge matrix randomization appeared to perform stably enough for our purposes, we regularly did projections of the edge matrices onto $SU(2)$ as a precautionary measure.

References

- [1] C.-N. Yang, R.L. Mills, Conservation of isotopic spin and isotopic gauge invariance, *Phys. Rev.* 96 (1954) 191–195, [doi:10.1103/PhysRev.96.191](https://doi.org/10.1103/PhysRev.96.191).
- [2] S. Weinberg, *The Quantum theory of Fields*, vol. 1: Foundations, Cambridge Univ. Press, Cambridge, UK, 1995.
- [3] S. Weinberg, *The Quantum Theory of Fields*, vol. 2: Modern Applications, Cambridge Univ. Press, Cambridge, UK, 1996.
- [4] M.E. Peskin, D.V. Schroeder, *An Introduction to Quantum Field Theory*, Addison–Wesley, Reading, USA, 1995.
- [5] K.G. Wilson, Confinement of quarks, *Phys. Rev. D* 10 (8) (1974) 2445–2459, [doi:10.1103/PhysRevD.10.2445](https://doi.org/10.1103/PhysRevD.10.2445).
- [6] M. Creutz, *Quarks, Gluons and Lattices*, Cambridge Monographs on Mathematical Physics, Cambridge Univ. Press, Cambridge, UK, 1986.
- [7] N.H. Christ, R. Friedberg, T.D. Lee, Weights of links and plaquettes in a random lattice, *Nucl. Phys. B* 210 (1982) 337, [doi:10.1016/0550-3213\(82\)90124-9](https://doi.org/10.1016/0550-3213(82)90124-9).
- [8] N.H. Christ, R. Friedberg, T.D. Lee, Gauge theory on a random lattice, *Nucl. Phys. B* 210 (1982) 310, [doi:10.1016/0550-3213\(82\)90123-7](https://doi.org/10.1016/0550-3213(82)90123-7).
- [9] N.H. Christ, R. Friedberg, T.D. Lee, Random lattice field theory: General formulation, *Nucl. Phys. B* 202 (1982) 89, [doi:10.1016/0550-3213\(82\)90222-X](https://doi.org/10.1016/0550-3213(82)90222-X).
- [10] J.M. Drouffe, K.J.M. Moriarty, C.N. Mouhas, Monte Carlo simulation of pure $U(N)$ and $SU(N)$ gauge theories on a simplicial lattice, *Comput. Phys. Commun.* 30 (1983) 249, [doi:10.1016/0010-4655\(83\)90092-9](https://doi.org/10.1016/0010-4655(83)90092-9).
- [11] J.M. Drouffe, K.J.M. Moriarty, C.N. Mouhas, $U(1)$ four-dimensional gauge theory on a simplicial lattice, *J. Phys. G* 10 (1984) 115, [doi:10.1088/0305-4616/10/2/004](https://doi.org/10.1088/0305-4616/10/2/004).
- [12] J.M. Drouffe, K.J.M. Moriarty, Gauge theories on a simplicial lattice, *Nucl. Phys. B* 220 (1983) 253–268, [doi:10.1016/0550-3213\(83\)90040-8](https://doi.org/10.1016/0550-3213(83)90040-8).
- [13] J.M. Drouffe, K.J.M. Moriarty, $U(2)$ four-dimensional simplicial lattice gauge theory, *Z. Phys. C* 24 (1984) 395, [doi:10.1007/BF01410379](https://doi.org/10.1007/BF01410379).
- [14] K.E. Cahill, R. Reeder, Comparison of the simplicial method with Wilson's Lattice Gauge Theory for $U(1)$ in three-dimensions, *Phys. Lett. B* 168 (1986) 381, [doi:10.1016/0370-2693\(86\)91648-5](https://doi.org/10.1016/0370-2693(86)91648-5).
- [15] J.M. Drouffe, K.J.M. Moriarty, High-statistics study of the phase transition in $U(2)$ four-dimensional simplicial lattice gauge theory, *Journal of Physics G: Nuclear Physics* 10 (10) (1984) L221; URL <http://stacks.iop.org/0305-4616/10/i=10/a=001>.
- [16] R.W.B. Ardill, J.P. Clarke, J.M. Drouffe, K.J.M. Moriarty, Quantum chromodynamics on a simplicial lattice, *Phys. Lett. B* 128 (1983) 203, [doi:10.1016/0370-2693\(83\)90391-X](https://doi.org/10.1016/0370-2693(83)90391-X).
- [17] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, 1st edition, Studies in Mathematics and Its Applications, vol. 4, North-Holland Publishing Company, 1978.
- [18] P. Monk, *Finite Element Methods for Maxwell's Equations*, reprinted edition, Oxford Science Publications, 2006.
- [19] J.-C. Nédélec, Mixed finite elements in \mathbb{R}^3 , *Num. Math.* 35 (1980) 315–341.
- [20] H. Whitney, *Geometric Integration Theory*, Princeton University Press, Princeton, NJ, 1957.
- [21] R. Hiptmair, Finite elements in computational electromagnetism, *Acta Numerica* 11 (2002) 237–339, [doi:10.1017/S0962492902000041](https://doi.org/10.1017/S0962492902000041).
- [22] C.M. Bender, K.A. Milton, Approximate determination of the mass gap in quantum field theory using the method of finite elements, *Phys. Rev. D* 34 (10) (1986) 3149–3155, [doi:10.1103/PhysRevD.34.3149](https://doi.org/10.1103/PhysRevD.34.3149).

- [23] C.M. Bender, K.A. Milton, D.H. Sharp, Gauge invariance and the finite-element solution of the Schwinger model, Phys. Rev. D 31 (2) (1985) 383–388, doi:[10.1103/PhysRevD.31.383](https://doi.org/10.1103/PhysRevD.31.383).
- [24] T.G. Halvorsen, T.M. Sørensen, Simplicial gauge theory on spacetime, arXiv:1107.1420.
- [25] H. Flyvbjerg, H.G. Petersen, Error estimates on averages of correlated data, Journal of Chemical Physics 91 (1) (1989) 461–466.
- [26] M. Creutz, Monte Carlo study of quantized $SU(2)$ gauge theory, Phys. Rev. D 21 (1980) 2308–2315, doi:[10.1103/PhysRevD.21.2308](https://doi.org/10.1103/PhysRevD.21.2308).
- [27] S.H. Christiansen, H.Z. Munthe-Kaas, B. Owren, Topics in structure-preserving discretization, Acta Numerica 20 (2011) 1–119, doi:[10.1017/S096249291100002X](https://doi.org/10.1017/S096249291100002X).
- [28] MPICH2, URL <http://www.mcs.anl.gov/mpi/mpich2>.
- [29] H. Bauke, TINA pseudo-RNG library, URL <http://trng.berlios.de>.

Chapter 11

Comments on “The Stochastic Nonlinear Schrödinger Equation in H^1 ,”

Submitted in 2011 to “Stochastic Analysis and Applications”. Awaiting referee report.

Comments on *The Stochastic Nonlinear Schrödinger Equation in H^1*

Torquil Macdonald Sørensen

Centre of Mathematics for Applications

University of Oslo

NO-0316 Oslo, Norway

Email: t.m.sorensen@matnat.uio.no, torquil@gmail.com

Abstract

We point out two mistakes in the paper *The Stochastic Nonlinear Schrödinger Equation in H^1* by de Bouard and Debussche [1], which gives an existence proof for the stochastic nonlinear Schrödinger equation. These mistakes draw the main conclusions of that paper into question.

Keywords (MSC2010): 35Q41 Time-dependent Schrödinger equations, Dirac equations; 35R60 Partial differential equations with randomness, stochastic partial differential equations; 35G20 Nonlinear higher-order equations.

1 Regarding the proof of Theorem 4.1

Consider [1, Theorem 4.1]. The following n -dependent parameter ranges for σ are assumed,

$$\begin{cases} 0 < \sigma & , n = 1, 2 \\ 0 < \sigma < 2 & , n = 3 \\ \frac{1}{2} \leq \sigma < \frac{2}{n-2} \quad \text{or} \quad \sigma < \frac{1}{n-1} & , n \geq 4. \end{cases} \quad (1.1)$$

The theorem essentially states that an *admissible pair* of Lebesgue space exponents (r, p) exists such that a stochastic nonlinear Schrödinger equation has a unique solution in a certain function space characterised by (r, p) . Admissibility is defined as

$$r \geq 2, \quad \frac{2}{r} = n \left(\frac{1}{2} - \frac{1}{p} \right). \quad (1.2)$$

Dual Lebesgue space exponents are denoted by primed quantities, and are defined by the equation

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

The proof given is for the special case $\sigma \geq 1/2$. In the proof a second admissible pair (γ, s) is introduced after [1, Eq.(4.17)]. The parameters s' and p are related through another parameter q which arises in the proof,

$$\frac{1}{s'} = \frac{2\sigma}{q} + \frac{1}{p}, \quad (1.3)$$

which is described prior to [1, Eq.(4.18)]. The parameter q arises due to the use of the Sobolev embedding $H^1(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$. It is claimed that the embedding holds because " $q < 2n/(n-3) < 2n/(n-2)$ ". However, the second part of this inequality is incorrect, and therefore the Sobolev embedding is used without proper justification.

2 Regarding the proof of Lemma 4.3

In the proof of [1, Lemma 4.3], in the second estimate on p.121, the interpolation inequality for L^p -spaces, followed by the Sobolev embeddings $H^1(\mathbb{R}^n), W^{1,p}(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$ were used, where the parameter q was "as above". Therefore, for the same reason, the Sobolev embeddings used here also lack a proper justification.

3 Conclusions

We believe that the errors described above put into question the validity of the existence proof, which is the main result of [1].

References

- [1] A. de Bouard and A. Debussche. The stochastic nonlinear Schrödinger equation in H^1 . *Stochastic Analysis and Applications*, 21(1):97–126, 2003.