University of Southampton

Faculty of Engineering and Physical Sciences
School of Physics
High Energy Physics (HEP) Group

A first study on the infrared properties of holographic models of cosmology using lattice-quantum-field theory simulations

by

Benjamin Thomas Kitching-Morley
ORCiD: 0000-0003-3150-401X

A thesis for the degree of
Doctor of Philosophy

May 2023

University of Southampton

Abstract

Faculty of Engineering and Physical Sciences
School of Physics

Doctor of Philosophy

A first study on the infrared properties of holographic models of cosmology using lattice-quantum-field theory simulations

by Benjamin Thomas Kitching-Morley

Cosmic inflation, in which the early universe undergoes a short, violent expansion, has successfully described many phenomena, including the near scale-invariance of the Cosmic Microwave Background Radiation (CMB). However, inflation is not ultraviolet-complete and suffers from self-consistency issues. Holographic Cosmology is an alternative framework in which the early universe is described by a three-dimensional dual quantum field theory (QFT). Correlations in the CMB are predicted by two-point correlators of the energy-momentum tensor (EMT) of the dual theory. A perturbative treatment of holographic cosmology has proven competitive with inflation in fits to CMB data, however, a non-perturbative treatment is needed to test the theory against all multipoles of the CMB. Lattice QFT provides such an approach by regularizing theories through placing them on a finite spacetime lattice. In this thesis, we tackle three challenges in making predictions of holographic cosmology using a lattice-regulated dual QFT with scalar fields in the adjoint of $SU(N)$ and a ϕ^4 interaction. First, we provide numerical evidence supporting a conjecture that a class of super-renormalizable theories, including holographic cosmology dual theories, is non-perturbatively infrared finite. This is necessary for holographic cosmology to be predictive and implies a resolution of the Big Bang singularity within the holographic cosmology framework. Secondly, we explore a novel approach to regulating ultraviolet divergences appearing in calculations of the two-point EMT correlator. In this approach, Laplace transforms of position-space lattice data offer cancellation of quadratic divergences appearing in momentum-space. We finally investigate the feasibility of using multilevel methods to reduce statistical noise in holographic dual two-point calculations. Holographic dualities necessitate correlator calculations to be done in the critical regime where the correlation length diverges. Through a novel study of the critical-scaling properties of the multilevel algorithm, using the 2D-Ising model, we demonstrate the unsuitability of the technique in this regime.

Acknowledgements

First and foremost, I want to thank my supervisors Andreas Jüttner and Kostas Skenderis for all of the time and energy they have given me over the course of this Ph.D. Thank you for being supportive of me and giving me the freedom to choose my own path - it's been a pleasure to be your student.

Thank you to all of the members of the LatCos for the stimulating discussions and for welcoming me into the collaboration. I would especially like to thank Joseph K. L. Lee and Henrique Rocha for all of the help you have given me when I was unsure, and for sharing the Ph.D. experience with me.

I would like to thank my parents, first of all for welcoming me back to their homes during the COVID-19 lockdowns, but more importantly for their unwavering love and support.

Next I would like to thank my friends, it's not been an easy road at times but you have been a constant source of positivity for me!

Lastly, I want to thank my flatmates Michael Casey and Chris Cho. The hundreds (thousands?) of hours spent discussing life, the universe and everything has been a highlight of the last four years, and helped me grow immeasurably.

Thank you to the Engeneering and Physical Sciences Research Council for funding this project under grant EP/L015382/1.

Publications

I am author on, and have contributed towards, the following publications:

Journal Publications

[1]: Guido Cossu, Luigi Del Debbio, Andreas Jüttner, Ben Kitching-Morley, Joseph K. L. Lee, Antonin Portelli, Henrique Bergallo Rocha, and Kostas Skenderis and. Nonperturbative infrared finiteness in a superrenormalizable scalar quantum field theory. Physical Review Letters, 126(22), jun 2021. https://www.doi.org/10.1103/PhysRevLett.126.221601

[2]: Luigi Del Debbio, Elizabeth Dobson, Andreas Jüttner, Ben Kitching-Morley, Joseph K. L. Lee, Valentin Nourry, Antonin Portelli, Henrique Bergallo Rocha, and Kostas Skenderis. Renormalisation of the energy-momentum tensor in three- dimensional scalar SU(N) theories using the Wilson flow. Phys. Rev. D, 103(11) https://www.doi.org/10.1103/PhysRevD.103.114501

Conference Proceedings

[3]: Ben Kitching-Morley and Andreas Jüttner. A numerical and theoretical study of multilevel performance for two-point correlator calculations. PoS, LATTICE2021: 133, 2022. https://doi.org/10.22323/1.396.0133

[4]: Joseph K. L. Lee, Elizabeth Dobson, Andreas Jüttner, Ben Kitching-Morley, Valentin Nourry, Antonin Portelli, Henrique Bergallo Rocha, Kostas Skenderis, and Luigi Del Debbio. Renormalisation of the 3D SU(N) scalar energy-momentum tensor using the Wilson flow. PoS, LATTICE2021:313, 2022. https://doi.org/10.22323/1.396.0313

[5]: Henrique Bergallo Rocha, Luigi Del Debbio, Elizabeth Dobson, Andreas Jüttner, Ben Kitching-Morley, Joseph K. L. Lee, Antonin Portelli, and Kostas Skenderis. Critical behaviour and phase structure of 3d Scalar+Gauge Field Theories in the adjoint representation. PoS, LATTICE2021:270, 2022. https://doi.org/10.22323/1.396.0270

Codes and Datasets

[6]: Ben Kitching-Morley and Andreas Jüttner. andreasjuettner/finite-size-scaling- analysis: V1.1 code release, November 2020. https://doi.org/10.5281/zenodo.4290508

[7]: Guido Cossu, Henrique Bergallo Rocha, Luigi Del Debbio, Andreas Jüttner, Ben Kitching-Morley, Joseph K.L. Lee, Antonin Portelli, and Kostas Skenderis. Nonperturbative infrared finiteness in super- renormalisable scalar quantum field theory data release, November 2020.https://doi.org/10.5281/zenodo.4266114

[8]: Ben Kitching-Morley. Multievel notebook: V1.0 code release, https://doi.org/10.5281/zenodo.7311081, November 2022.

[9]: Ben Kitching-Morley. Multievel notebook: V1.0 data release, https://doi.org/10.5281/zenodo.7303520, November 2022.

Contents

Declaration of Authorship

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as: [1],[2],[3],[4],[5]. Codes/datasets: [6],[7],[8],[9]

Signed:..... Date:.....

Chapter 1

Introduction

The Cosmic Microwave Background Radiation (CMB) is a remnant of the early universe, formed around 300,000 years after the Big Bang. Its correlations tell us about a time much earlier in the universe's history - less than 10^{-30}s after the Big Bang. The CMB, therefore, acts as a testing bed for theories of the early universe and, by extension, theories of gravity. Cosmic inflation is a component of ΛCDM, the "standard model of cosmology". It proposes a short, violent expansion early in the universe's history, and its predictions match well with CMB data. However, some discrepancies remain, including in predictions of the low-multipole correlations of the CMB. Moreover, inflation is not an ultraviolet-complete (UV-complete) theory, meaning it can only make physical predictions up to a certain energy scale, in this case, the Planck scale. We are thus motivated to seek candidate theories of the early universe. A UV-complete theory must be a quantum theory of gravity, given the incredibly high energy densities of the universe's first moments. In chapter 2, we explore holographic cosmology, an alternative framework to cosmic inflation, where the four-dimensional early universe is described by a dual three-dimensional quantum field theory (QFT) with fields in the adjoint of $SU(N)$. Holographic cosmology is a quantum theory of gravity, and if correct would be a major leap in humanity's understanding of the universe. In this framework, correlations in the CMB correspond to two-point functions of the energy-momentum tensor (EMT) of the dual three-dimensional QFT. Previous work has used perturbation theory to test holographic cosmology against the CMB. In these tests, it performed as well as cosmic inflation in the range of validity of perturbation theory. However, some of the most significant discrepancies between cosmic inflation and the CMB lie in the low-multipole region. In this region, a non-perturbative treatment of holographic cosmology is required.

Lattice QFT is a successful framework for regularizing QFTs non-perturbatively. By placing the quantum theory onto a finite-volume discrete lattice, ultraviolet and infrared divergences can be regulated. In chapter 3, we recount the theoretical underpinning of lattice QFT before exploring some of the methods used to perform calculations using

computers. We present a lattice regularized version of our holographic dual QFT action, which we use to make non-perturbative predictions of holographic cosmology. This goal is the mandate of the LatCos collaboration. Predictions of the holographic dual QFT are made, as in perturbative calculations, by calculating two-point functions of the EMT. While the steps involved in doing this are understood in principle, several challenges remain. In this thesis, we focus on three of the biggest challenges:

1. *Infrared Divergences*: The three-dimensional QFTs dual to the early universe are super-renormalizable theories. Such theories are expected to be well-behaved in the ultraviolet but suffer irreconcilable divergences in the infrared for perturbative calculations. In chapter 4, we test a hypothesis that proposes super-renormalizable theories such as ours are non-perturbatively finite in the infrared. We use the phase structure of the theory to identify critical masses for different couplings g and lattice sizes L. The behavior of the critical mass in the limit $gL \to \infty$ is indicative of the infrared finiteness of the theory and is used to test it. Demonstrating infrared-finiteness is an essential prerequisite to making calculations of the holographic dual theory. Moreover, it would demonstrate a resolution of the Big Bang singularity in the holographic cosmology framework. It would therefore offer significant motivation for further study of holographic cosmology models.

2. *Renormalization*: By regulating the theory using a finite lattice spacing a, we are left with the challenge of extrapolating our results in the $a \to 0$ limit. Since we are interested in the two-point function of the EMT, this needs to be done in two stages. Firstly, the EMT operator needs to be renormalized. This has been done in our paper [2], but was not a focus of the author's research. In chapter 5, we focus instead on the second stage of the renormalization procedure, which deals with additional divergences appearing in the two-point function of the EMT. Specifically, a q^2 term, divergent in the $a \to 0$ limit appears. In the infinite-volume continuum, the Laplace transforms of certain functions are given by the analytic continuation $q \to -iq$ of the corresponding Fourier transforms. We attempt to use this result to cancel q^2 divergences appearing in the EMT two-point function.

3. *Statistical Noise*: Simulating lattice QFTs is an intrinsically probabilistic exercise, giving estimates of physical quantities with associated statistical error bars. In attempting to make predictions of holographic dual QFTs using the lattice, our collaboration has encountered high levels of statistical noise. It is possible to reduce statistical noise through a variety of methods, for example, by achieving larger statistics through more efficient code or more computing resources. However, we already use an efficient simulation code based on the GRID library, and computational resources are limited. Therefore, in chapter 6, we explore an alternative approach - the multilevel algorithm. This algorithm divides lattice simulations into multiple layers and has been successful in reducing the signal-to-noise ratio

for various lattice QFT correlators. While successful, it is suspected that this algorithm's performance suffers as the correlation length of the system increases. This is troubling since we perform our dual QFT simulations in the critical regime, where the correlation length diverges. The bulk of chapter 6 is therefore devoted to measuring the effect of the correlation length on the performance of the multilevel algorithm. To this end, the 2D-Ising model is used as a test system.

I, the author, am part of a collaboration of researchers called LatCos (Lattice Cosmology). The other members of this collaboration at the time of writing are Guido Cossu, Luigi Del Debbio, Lizzie Dobson, Andreas Jüttner, Joseph K. L. Lee, Valentin Nourry, Antonin Portelli, Henrique Bergallo Rocha, and Kostas Skenderis. The collaboration aims to test holographic cosmology models through the simulation and analysis of lattice-regularized holographic dual QFTs. Our collaboration meets regularly, and most of the research produced involves input from all members. In this thesis, I focus on the research led by me or research for which I was a leading contributor. To clarify my individual contribution, I will now run through my contributions presented in this thesis:

- *Chapter 4 - Infrared Finiteness*: Andreas Jüttner, Joseph Lee and I developed the frequentist methods and I developed the Bayesian approaches used to test IR-finiteness and estimate parameters in [1] for $N = 2$ and $N = 4$ data. Following this paper, Andreas and I produced $N = 3$ configurations, and I produced $N = 5$ configurations to extend this study. I have also expanded upon the frequentist methods, including implementing correction-to-scaling contributions in our fits. I have also developed improvements to the reweighting procedure (section 4.6), of which the smoothed reweighting is, to the best of my knowledge, a novel method. I have also led the investigation of the effect of N on the fit range that IR-finiteness can be tested over (section 4.15). Lastly, I have implemented a Bayesian Model Averaging [10] approach to parameter estimation (section 4.12).
- *Chapter 5 - Renormalization*: The method of using Laplace transforms to remove ultraviolet divergences was conceived by Kostas Skenderis and developed and tested in code by me.
- *Chapter 6 - The Multilevel Algorithm*: This multilevel study was led by me and is the first study, to my knowledge, to investigate the scaling of multilevel performance with the correlation length.

Chapter 2

Holographic Cosmology

2.1 An Expanding Universe

Gazing out into space beyond our galaxy we immediately notice that the universe looks similar in whichever direction we look. This is a statement of spatial homogeneity, which, along with the assumption of isotropy, implies the universe has a geometry described by the FriedmannLemaîtreRobertsonWalker (FLRW) metric [11–18],

$$ds^2 = dt^2 - a(t)^2 \left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right], \tag{2.1}$$

where θ and ϕ are the standard spherical coodinates, r is a radial quantity and t is a time coordinate. The constant $k \in -1, 0, 1$ determines the shape of the universe: $k = 0$ being flat, $k = 1$ being spherical and $k = -1$ being hyperbolic. The factor $a(t)$ is called the scale factor, and if $k \neq 0$ then a is usually normalized so that $|k| = 1$.

In 1929 Edwin Hubble found that the speed at which other galaxies move away from us is directly proportional to their distance from us [19]: we are living in a universe that is expanding. This somewhat startling conclusion, implying that $\dot{a}(t_0) > 0$ where t_0 is the current time, was predicted several years prior by Friedmann [20, 21]. By solving Einstein's equations of general relativity (GR) for a universe with FLRW geometry, homogeneous pressure p and density ρ Friedmann arrived at

$$\left(\frac{\dot{a}}{a}\right) - \frac{k}{a^2} = \frac{8\pi G}{3}\rho, \tag{2.2}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p), \tag{2.3}$$

which can be combined to give

$$\dot{\rho} = -3H(\rho + p), \tag{2.4}$$

where $H = \dot{a}/a$ is the Hubble constant. Given positive density and pressure, the second derivative of the scalar factor $\ddot{a}$ must be negative, which together with $\dot{a}(t_0) > 0$ predicts a universe that was once vanishingly small ($a \sim 0$): the *Big Bang Theory*.

Assuming the equation of state $p = w\rho$, the relation $\rho \propto a^{3(w+1)}$ is obtained. For cold matter $w = 0$ and for radiation $w = 1/3$. This leads to one of the problems with this model of cosmology: the *flatness problem*. In a perfectly flat universe, $\rho = \rho_c$, where ρ_c is the density required to yield $k = 0$ in eq. (2.2). Substituting $\rho = \rho_c + \Delta\rho$, with $\Delta\rho_c \ll \rho_c$ into eq. (2.2) and using the scaling of ρ in terms of a yields

$$|\Delta\rho/\rho| \sim a^{1+3w}, \tag{2.5}$$

so any deviations from the critical density grow with time. At the present time however, $\Omega = \Delta\rho/\rho_c$ is observed to be very close to 1. For this to be the case, it must have been even closer to 1 (*fine-tuned*) in the past such that any deviations have not been amplified since.

To understand the causality of events in the universe's history, it is beneficial to change time coordinate. The *conformal time* τ at a given $t = t^*$ is the minimum time it would take a photon to reach the edge of the observable universe, under the assumption that the scale factor remains constant for $t > t^*$. It is given by

$$\tau(t^*) = \int_0^{t^*} \frac{dt'}{a(t')}. \tag{2.6}$$

The *particle horizon* at time t^* is simply the distance a photon travels if it travels for conformal time $\tau(t^*)$, which in natural units is just equal to the conformal time. Every thing we can observe must, by definition, be closer than the particle horizon. By contrast objects outside of our particle horizon are not causally linked with us since there is a mutual lack of information transfer (as evidenced by our inability to observe them). Since $a(t)$ is strictly non-negative, the particle horizon cannot decrease in absolute terms with time.

The Cosmic Microwave Background Radiation (CMB) (first observed in [23]) acts as a window to the time of *recombination* ($t_{\text{rec}} \approx 300,000$ years) when the universe cooled sufficiently to allow electrons and protons to bind to form neutral atoms. This caused photons to decouple from matter, and it is these very photons (red-shifted) that are observed today in the CMB. The CMB is observed to be homogeneous to better than one part in ten thousand, in all directions. In figure 2.1, the measurements of these small fluctuations of the CMB by the Planck telescope are shown [22]. However, the particle horizon at t_{rec} is considerably smaller than the particle horizon today. This implies that regions of the night sky, which are observed to be homogeneous, have in fact never been in causal contact. This is the *Horizon Problem*, and it is shown diagrammatically in figure 2.2, where it is encapsulated by $x_f < x_b$.

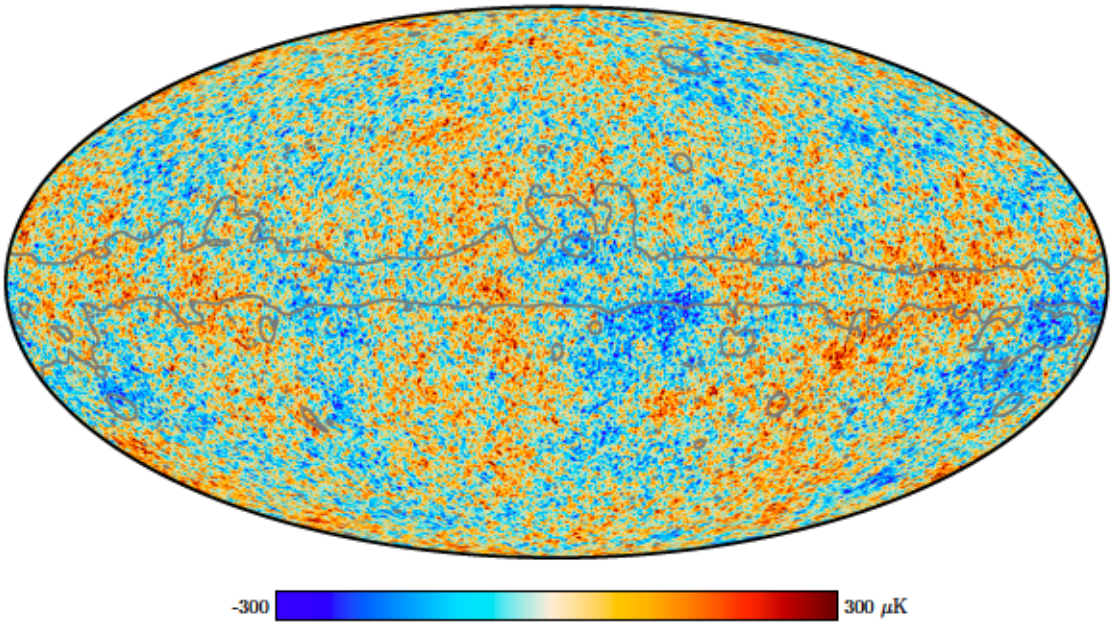

FIGURE 2.1: Measurements fluctuations in the cosmic microwave background radiation by the Planck telescope [22].

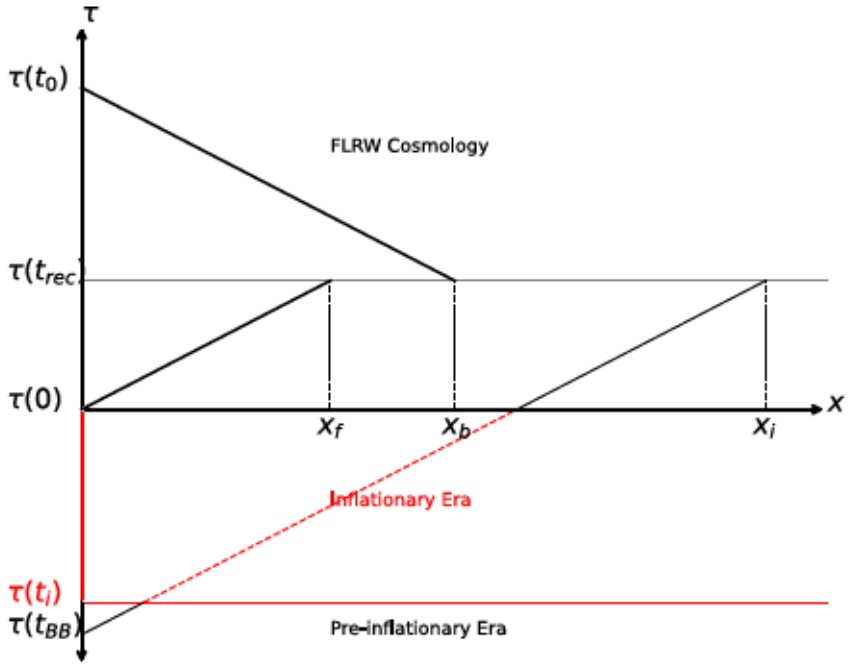

FIGURE 2.2: Diagram of the horizon problem and its resolution from inflation. The solid black lines are the forward and backwards light cones in standard cosmology, where $\tau = t = 0$ is the Big Bang, t_{rec} is the time of recombination and t_0 is the present. The Horizon Problem is the statement that $x_b > x_f$, in contradiction with the observed homogeneity in the CMB. One solution to this is an inflationary phase running from t_i to 0 which dramatically expands the conformal time since the Big Bang (now at $t = t_{BB}$). (Figure inspired by [24])

2.2 Cosmic Inflation

Inflation offers a solution to these issues of standard cosmology [25]. Consider a regime where the dominant energy-matter content has an equation of state $p = -\rho$ (e.g. $w = -1$). Substituting this into equations (2.4) and (2.3) yields $\dot{\rho} = 0$ and $\ddot{a}/a = const$, an exponentially expanding constant density universe. This behavior is postulated to occur during an *inflationary period* starting at $t = t_i$ and ending at $t = t_R$ (the *reheating time*). The duration of this inflationary period is incredibly small in terms of coordinate time t ($\sim 10^{32}s$) and occurred very early in the history of the universe ($\sim 10^{36}s$). In terms of conformal time, however, inflation was very significant. We will use a popular convention shifting t_R to $t = 0$, where the Big Bang and t_i have negative values ($-t_i$ and $t_{\mathrm{BB}} = -t_R$ respectively). Given the tiny size of t_i, this shift to the time coordinate does not affect t_0 or t_{rec} within errors. The effect of adding an inflationary period is shown in figure 2.2.

This $w = -1$ behavior can be realized in Einstein's gravity by a scalar theory with Lagrangian density

$$\mathcal{L}(\varphi) = \frac{1}{2}(\partial_\mu \varphi)(\partial^\mu \varphi) + V(\varphi), \tag{2.7}$$

on a background metric $g_{\mu\nu}$ and Einstein-Hilbert action

$$S_m = \int d^4x \sqrt{-g}\mathcal{L}(x). \tag{2.8}$$

By varying this action with respect to the metric, one can obtain the energy-momentum tensor (EMT), including its diagonal elements ρ and p [24],

$$\rho(\varphi) = \frac{1}{2}\dot{\varphi}^2 + \frac{1}{2a^2}(\nabla\varphi)^2 + V(\varphi), \tag{2.9}$$

$$p(\varphi) = \frac{1}{2}\dot{\varphi}^2 - \frac{1}{6a^2}(\nabla\varphi)^2 - V(\varphi). \tag{2.10}$$

By taking the condition $\dot{\varphi}, a^{-2}(\nabla\varphi)^2 \ll V(\varphi)$ the inflationary equation of state $p = -\rho$ can be recovered. These conditions are called the *slow-rolling* conditions, and for this model, it can be shown that, under the right conditions, they hold throughout the inflationary period [24][1]. Figure (2.2) demonstrates how inflation can resolve the Horizon Problem. In this model, correlations in the CMB correspond to quantum fluctuations which have been expanded as the universe expands. The larger the length scale of the correlations today, the earlier those correlations were formed. Moreover, the substitution of $w = -1$ into eq. (2.5) shows that that during the inflationary period the ratio $\Delta\rho/\rho$ approaches 0, resolving the flatness problem.

[1] We note here that [24] was a useful reference in the writing of sections 2.1 and 2.2 in general.

Many inflationary models suffer from self-consistency issues however. Inflation is built upon perturbations of Einstein's GR in the early universe, yet studies of correlations in the CMB and in galaxy clusters show their origins to be at such early times during inflation that the length scales in question are sub-Planckian. General relativity breaks down at these scales and a quantum theory of gravity is required, in which case the assumptions leading to inflation no longer hold. This is referred to as the *Trans-Planckian problem* [26].

Even in inflationary models that do not suffer from this issue, we still know inflation is not a UV-complete theory. In a universe with inflation the initial Big Bang singularity remains [27]. At times earlier than the inflationary period there is necessarily new physics involving a quantum theory of gravity that inflation cannot describe. This result means the initial value problem of inflation is not well defined. Lastly, it was shown [28] that observations of the magnitude of CMB fluctuations imply a fine-tuning of parameters entering into inflation. Given that one of the motivations for using inflation was to avoid a fine-tuning problem, this is troubling.

These problems among others motivate us to look for a new theory of the early universe to supersede inflation. This theory needs to be UV-complete and ideally would not suffer from the same singularity issues present in inflation. One class of theories which could fill this role is holographic cosmology models, which we will now introduce.

2.3 Holographic Dualities

In popular culture a hologram is a three-dimensional image produced through a two-dimensional array of lights. While an observer may view the construction as an object in its own right, the information necessary to construct it is entirely contained within a two-dimensional surface. An attempt to describe this object from first principles should therefore be founded in the lower dimensional surface that produces it. This situation provides an analogy to the concept of a *holographic duality*. When used in this context holographic dualities relate physical theories including gravity in $(d+1)$-dimensions to quantum field theories (QFTs) without gravity in d-dimensions. Put another way, all physical degrees of freedom of the higher dimensional gravity theory are described by a lower dimensional quantum field theory. In these dualities, the higher dimensional theory is referred to as the *bulk* theory, while the lower dimensional one is referred to as the *boundary* theory. The higher dimensional theory typically contains gravity, while the lower dimensional theory typically contains gauge fields. These dualities are therefore commonly referred to as *gauge-gravity* dualities.

The concept that our universe may in fact be holographic originates in the study of black holes [29, 30]. The argument goes as follows: it is known that the entropy (which directly relates to the number of degrees of freedom (DoF)) of black holes is proportional

to their surface area [31]. Imagine sending matter (information) from infinity into a finite region of empty space. Eventually the density of the matter will reach a point whereby a black hole forms. If black hole formation obeys the second law of thermodynamics then the number of DoF before black hole formation cannot be greater than the number afterwards. Therefore the number of DoFs of the region must be less than or equal to those of the black hole, proportional to the region's area. Specifically the maximum entropy of a region of space is given by the Bekenstein Bound $S_{\max} = A/4G_N$, where A is the area and G_N is Newton's constant [32]. This bound demonstrates that the number of DoF in a $(d+1)$-dimensional theory with gravity match the number of DoF found in a d-dimensional quantum field theory without gravity (where the DoF scale as the volume).

For an outline of holographic dualities it is necessary to introduce some terminology. A *Dp-brane* is a fundamental object in string theory that has p-extensive spatial dimensions, one time dimension and act as places where open strings can end (an open string is a one-dimensional object with two ends, while a closed string has the topology of a circle). As an example a $D1$-brane is a one-dimensional object (distinct from a string). A heuristic argument for holographic dualities was outlined by Maldecena in his famous 1997 paper [33]. He considered a system of N coincident $D3$-branes in a ten-dimensional type IIB string theory in a decoupling/near-horizon limit. In this limit, the system has a weakly coupled description in terms of four-dimensional $\mathcal{N} = 4$ Super Yang-Mills theory when $g_s N \ll 1$, where g_s is the string coupling constant. The dynamics is captured by open strings terminating on the Dp-branes, as shown in the first panel of figure 2.3. In the opposite $g_s N \gg 1$ regime, the system has a weakly coupled description in terms of a type IIB supergravity solution with $AdS_5 \times S^5$ geometry. The Dp-branes impact the dynamics by acting as a source for closed strings, as shown in the second panel of figure 2.3. The AdS/CFT conjecture [33], the most well-known of the holographic conjectures, is that both the Super Yang-Mills description and the string theory description[2] on $AdS_5 \times S^5$ are valid at all values of g_s and N, with a weak/strong duality between the two: when one is strongly coupled the other is weakly coupled and vice versa. This would mean that calculations done in one theory can be done in the other theory. This has been used to better understand strongly coupled gauge theories, as the weakly coupled string theory models of gravity in $AdS_5 \times S^5$ are well understood.

The spacetime AdS_{d+1} here refers to a $(d+1)$-dimensional *Anti-de Sitter* geometry. Anti-de Sitter spaces are maximally symmetric geometries[3] with negative curvature, which can be represented by the metric,

$$ds^2 = d\bar{r}^2 + e^{\frac{2\bar{r}}{L_{\text{AdS}}}} \left[-\frac{d\bar{t}^2}{1 - k\bar{t}^2} + \bar{t}^2 d\Omega_{-}^{(d-1)} \right], \tag{2.11}$$

[2] Well approximated by a supergravity solution when g_s and N are both large.

[3] This means that all points on the manifold have the same local curvature. Physically it means that an observer is unable to determine their location on the manifold by observing their local geometry.

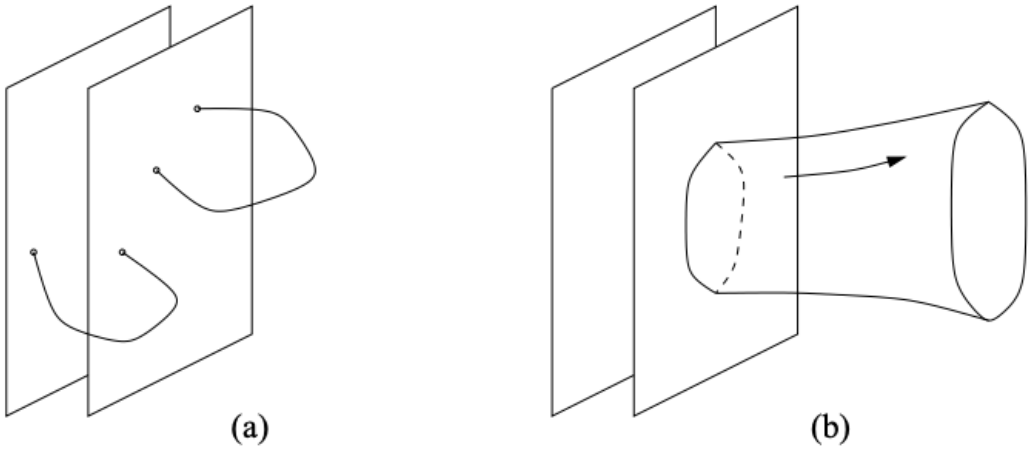

FIGURE 2.3: Two types of dynamics that can result from a stack of coincident Dp-branes. In the first the dynamics is driven by open strings ending on the branes, while in the second the dynamics is driven by branes acting as a source for closed strings. Figure taken from [34].

where $\bar{t}$ is a time coordinate, $\bar{r}$ is a radial coordinate, $k \in \{1, 0, -1\}$, L_{AdS} is the characteristic length scale and Ω_{-}^{d-1} is the metric on a unit hyperboloid with $SO(1, d-1)$ invariance. For the rest of the chapter we take $k = 0$, in which case the metric can be written in a more compact form using Poincaré coordinates,

$$ds^2 = \frac{L_{\text{AdS}}^2}{z^2}(dz^2 + \eta_{\mu\nu}dx^\mu dx^\nu), \tag{2.12}$$

where $\eta_{\mu\nu}$ is the Minkowski metric in d dimensions, $\eta_{\mu\nu}dx^\mu dx^\nu = -d\bar{t}^2 + \delta_{ij}dx^i dx^j$ with δ_{ij} is the $(d-1)$-dimensional Euclidean metric.

A link has therefore been established between a four-dimensional gauge theory (boundary) and five-dimensional theory with gravity (bulk)[4]. This is why these theories are often referred to as gauge-gravity dualities.

In general, stacks of N Dp-branes in the $g_s N \to \infty$ and decoupling/near-horizon limits have associated supergravity solutions conformal to $AdS_{p+2} \times S^{8-p}$ [35]. In the case of an $AdS_4 \times S^6$ near-horizon geometry it was shown that the geometry is locally power law, meaning it has an FLRW form where the scale factor $a \sim \bar{r}^m$, for some integer m [35]. Spacetimes with an *asymptotically-AdS* geometry are referred to as *domain-wall* solutions. These solutions are holographically dual to QFTs that flow towards a *conformal* point in different limits [36]. At these conformal points the theory gains the extra symmetry of invariance under conformal (locally angle preserving) deformations

[4]The bulk theory is five-dimensional rather than ten-dimensional as fields can be expressed as harmonics over the S^5 and the resultant theory lives in remaining five extensive dimensions

of spacetime. On the gravity side these *domain-wall* solutions have the metric

$$ds^2 = d\bar{r}^2 + e^{\frac{2A(\bar{r})}{L_{\rm AdS}}}\eta_{\mu\nu}dx^\mu dx^\nu,, \tag{2.13}$$

where $A(\bar{r}) = \bar{r} + O(1/\bar{r})$. The surface $\bar{r} = \infty$ is then called a conformal boundary, where the spacetime nearby is locally AdS. Note that taking $A(\bar{r}) = \bar{r}$ is another parameterization of a purely AdS metric. In the Poincaré coordinates of eq. (2.12) this conformal boundary corresponds to $z = 0$.

A property common to all holographic dualities is that there is a one-to-one correspondence between operators on the boundary and fields propagating in the bulk. This relation is given by [37, 38]

$$\langle e^{\int d^4\vec{x}\phi_0(\vec{x})O(\vec{x})}\rangle_{\rm boundary} = \mathcal{Z}_{\rm bulk}\left[\phi(\vec{x},z)|_{z=0} = \phi_0(\vec{x})\right], \tag{2.14}$$

where the left-hand side is a path integral expectation value in the boundary theory, while the right-hand side is the full partition function of the bulk theory. The five-dimensional *AdS* coordinates have been broken up into z (see eq. (2.12)) and the remaining $\vec{x}$. A specific example is a correspondence between the energy-momentum tensor (EMT) operator in the boundary theory and the graviton in the bulk. Since the boundary theory is always expected to be a local theory with an EMT, the bulk theory will always contain gravity.

The dynamics in the bulk at a given radial (z) distance are governed by two-point correlators of operators on the boundary at a given energy scale. As the radial distance increases the the energy scale of the boundary correlator decreases. This relationship implies that increasing radial coordinate in the bulk is mapped to renormalization group flow (RG flows) in the boundary theory.

2.4 The Domain-Wall-Cosmology Correspondence

In contrast to AdS, *de Sitter* (dS) geometries are maximally-symmetric positive curvature geometries, which can be expressed by the metric

$$ds^2 = -dt^2 + e^{\frac{2t}{L_{\rm dS}}}\delta_{\mu\nu}dx^\mu dx^\nu,, \tag{2.15}$$

where $L_{\rm dS}$ is a characteristic length scale and $\delta_{\mu\nu}$ is the Euclidean metric in d-dimensions decomposing as $\delta_{\mu\nu}dx^\mu dx^\nu = (dx^0)^2 + \delta_{ij}dx^i dx^j$. This metric is exactly the same as eq. (2.1) taking $a(t) = \exp(2t/L_{\rm dS})$, revealing that dS solutions are a subset of cosmological solutions. Asymptotically-dS solutions can be defined by replacing $\exp(2t/L_{\rm dS})$ by $\exp(2A(t)/L_{\rm dS})$.

The similarities between domain-walls and this class of cosmological geometries are clear to see, and by taking the analytical continuation $(z, \bar{t}, L_{\text{Ads}}) = i(\tau, x^0, L_{\text{dS}})$ one metric can be mapped to the other; locally AdS geometries map to dS ones. This mapping is the *domain-wall-cosmology* (DW/C) correspondence [39]. While this correspondence has not been proven with a strongly-coupled bulk sector, it is conjectured to still hold. In AdS the coordinate $z > 0$, with $z = 0$ being a conformal boundary while for the dS metric $\tau < 0$, with $\tau = 0$ being a conformal boundary. The surface at $\tau = 0$ is most analogous to the time of reheating which marks the end of the inflationary era in cosmic inflation. After this point the universe can be well approximated by a non-holographic model such as Einstein's general relativity. In contrast to the AdS case, approaching the boundary $\tau \to 0$ in the bulk is mapped to an flowing into the ultraviolet on the boundary (or conversely the limit $\tau \to -\infty$ in the bulk corresponds to RG-flow towards the infrared on boundary).

It can also be shown that there is a similar correspondence for asymptotically-power-law solutions, where $a(t/L) \sim c_k(t/L)^k$, with c_k being constants, which result from $D2$-brane solutions. Again this mapping is only a conjecture for a strongly coupled bulk, and therefore it is referred to as a correspondence and not a duality.

2.5 Perturbations of Domain-Wall/Cosmology Solutions

Perturbations can be included in the DW/C correspondence (on either side) by decorating the geometries with scalar fields. In [39] the action

$$S = \frac{1}{\kappa^2} \int d^{d+1}x \sqrt{-g} \left[R - \frac{1}{2}|\partial\phi|^2 - \kappa^2 V(\phi) \right], \tag{2.16}$$

was considered, where g is the determinant of the metric, $V(\phi)$ is a potential term and $\kappa^2 = 8\pi G$ is a constant. This action is extremized by the following equations, which have been derived for domain-wall solutions in a variety of contexts [39–44]:

$$\begin{aligned} \frac{\dot{a}}{a} &= -\frac{1}{2}W, \\ \dot{\phi} &= W', \\ 2\eta\kappa^2 V &= W'^2 - \frac{3}{2}W^2, \end{aligned} \tag{2.17}$$

where $W = W(\phi)$ and $W' = dW/d\phi$ and $\eta = -1$ for cosmological solutions and $\eta = +1$ for domain-wall solutions. By studying linear perturbations of these solutions, equations of

motion can be derived [45]:

$$\ddot{\zeta} + \left(3H + \frac{\dot{\epsilon}}{\epsilon}\dot{\zeta}\right) - \eta\frac{q^2\zeta}{a^2} = 0, \tag{2.18}$$
$$\ddot{\gamma}_{ij} + 3H\dot{\gamma}_{ij} - \eta\frac{q^2\gamma_{ij}}{a^2} = 0,$$

where dotted quantities represent the derivative of those quantities with respect to the time coordinate t on the cosmology side, or with respect to $\bar{r}$ on the domain-wall side. $H = \dot{a}/a$ is the Hubble parameter, $\epsilon(t) = -\dot{H}/H^2$, γ_{ij} is the transverse traceless metric perturbation, $\vec{q}$ is the comoving wavevector of the perturbations and $\zeta = -2\psi(z,\vec{x})\delta_{ij} + (H/\dot{\varphi})\delta\varphi$. These equations can be solved by setting $\gamma_{ij}(q) = \gamma(q)e_{ij}$, where e is a constant polarization tensor. The quantities $\gamma(q)$ and $\zeta(q)$ are known as *cosmological mode functions*. Noting the difference in the sign of η between our solutions, the domain-wall-cosmology correspondence holds as long as the additional analytic continuation

$$q \to -iq, \tag{2.19}$$
$$\kappa^2 \to -\kappa^2,$$

is made. On the domain-wall side these solutions are asymptotically-AdS or asymptotically power-law in the $z \to 0^+$ limit, corresponding to asymptotically dS and asymptotically power-law solutions in the $\tau \to 0^-$ limit on the cosmology side. In this limit, $\zeta(q)$ and $\gamma(q)$ approach constant values ζ_0 and γ_0, leading to scalar- and tensor-power spectra,

$$\Delta_S^2(q) = \frac{q^3}{2\pi^2}|\zeta_0|^2, \tag{2.20}$$
$$\Delta_T^2(q) = \frac{q^3}{2\pi^2}|\gamma_0|^2.$$

By applying the conditions of normality and the Bunch-Davies vacuum condition, $(\zeta(q), \gamma(q) \sim e^{-iq\tau})$, the expressions

$$\Delta_S^2(q) = \frac{-q^3}{4\pi^2 \mathrm{Im}\Omega_0(q)}, \tag{2.21}$$
$$\Delta_T^2(q) = \frac{-q^3}{2\pi^2 \mathrm{Im}E_0(q)},$$

are derived [45], where Ω_0 is the late time limit ($t \to +\infty$) of $2\epsilon a^3/\dot{\zeta}(q)$ and E_0 is the late time limit of $(1/4)a^3\dot{\gamma}(q)/\gamma(q)$. By taking the imaginary part, divergences that appear when taking this limit can be controlled.

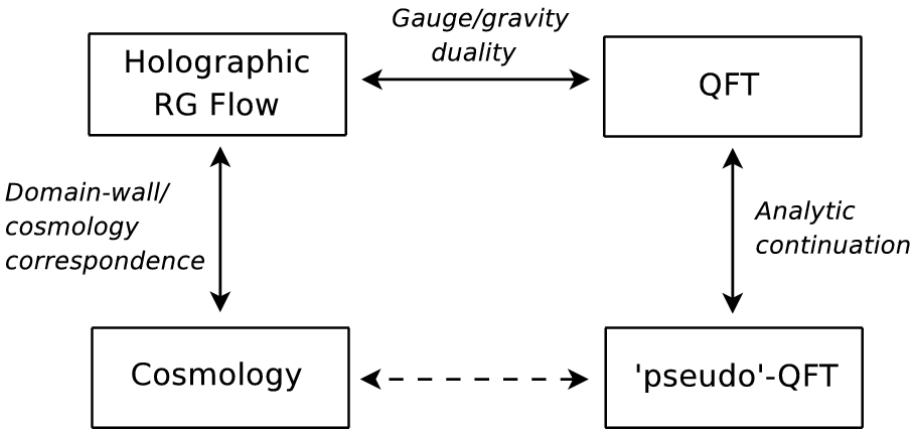

FIGURE 2.4: Proposed holographic dualities of the early Universe [45]

2.6 The Holographic Dictionary

Through combining the two dualities we have discussed thus far, namely the gauge-gravity duality and the DW/C correspondence, the Holographic Cosmology conjecture was born [45]. The conjecture uses domain-wall solutions as a stepping stone between cosmologies and dual QFTs (see figure 2.4). In the case of asymptotically de Sitter cosmological geometry with weak coupling in the bulk we recover inflation-like cosmologies. However, because of the strong-weak nature of the gauge-gravity duality, strongly coupled gravity solutions, which do not fit into a traditional inflationary framework, can also be explored through weakly-coupled dual theories.

One very profound implication of this framework, due to the strongly coupled nature of gravity in the early universe, is that time in this phase is best defined by the inverse-RG flow of the dual theory (see fig. 2.5). Increases in the time coordinate in a cosmological solution correspond to inverse-RG flows in the dual-QFT [45, 46]; the $\tau \to -\infty$ limit corresponds to flowing to the infrared in the dual theory. If the dual theory is IR-finite then the initial Big-Bang singularity, which is unresolved in traditional cosmic inflation, is resolved. A schematic to visualize the holographic cosmology model within the context of our universe is shown in figure 2.5. Correlations in the CMB today originate from primordial correlations in the very early universe. These primordial correlations are connected to n-point functions of operators in the dual theory. The fact that there is correlation between bulk fields, given by the dual theory, is a resolution of the horizon problem [47]. A key example is that fluctuations of the metric in the bulk correspond to expectation values of two-point functions of the Energy-Momentum Tensor (EMT). These two-point functions are responsible for the correlations observed today in the CMB.

After this time the physics of the universe is assumed to be well described by Einstein's general relativity. In order for holographic cosmology to describe physics in later times ($\tau > 0$) irrelevant operators in the dual theory need to be added. Since increasing τ

corresponds to inverse-RG flow, these operators would become more dominant at later times, leading to a change in gravitational phenomenology in the bulk giving physics similar to general relativity.

Focusing on the dual EMT, from symmetry grounds, the two-point function must decompose as

$$\langle T_{ij}(q_1)T_{kl}(q_2)\rangle = (2\pi)^3\delta^3(q_1+q_2)[A(q)\Pi_{ijkl} + B(q)\pi_{ij}\pi_{kl}], \tag{2.22}$$

where $q = |q_1|$, $\Pi_{ijkl} = \pi_{i(k}\pi_{l)j} - \frac{1}{2}\pi_{ij}\pi_{lk}$ and $\pi_{ij} = \delta_{ij} - q_iq_j/q^2$. Through application of the gauge-gravity duality, it is possible to obtain [48, 49]

$$\begin{aligned} 4\bar{E}_0(\bar{q}) &= A(\bar{q}), \\ \bar{\Omega}_0(\bar{q}) &= 4B(\bar{q}), \end{aligned} \tag{2.23}$$

where $\bar{E}_0$ and $\bar{\Omega}_0$ are defined analogously to E_0 and Ω_0 in eq. (2.21), with bars indicating that the quantities are associated with domain-wall solutions.

Taking an $N \to \infty$ limit of these correlators in the dual theory has a leading N^2 behavior [33], while the corresponding bulk limit is the semi-classical limit, which is associated with the leading contribution of $1/(2\kappa^2)$ in the $\kappa \to 0^+$ limit, where $\kappa = 8\pi G$ and G is Newton's constant [33]. Therefore, the analytic continuation of $\kappa^2 \to -\kappa^2$ in the bulk DW/C-correspondence maps to a continuation of $N^2 \to -N^2$ in the dual-correspondence (right hand side of fig. 2.4). Given that momenta in the bulk map to momenta in the dual, the overall dual analytic continuation is

$$\begin{aligned} q &\to -iq, \\ N &\to -iN. \end{aligned} \tag{2.24}$$

Combining this with equations (2.23) and (2.21) yields the *holographic dictionary*

$$\begin{aligned} \Delta_S^2(q) &= -\frac{q^3}{16\pi^2 \mathrm{Im}B(-iq)}, \\ \Delta_T^2(q) &= -\frac{2q^3}{\pi^2 \mathrm{Im}A(-iq)}. \end{aligned} \tag{2.25}$$

Another method using gauge-gravity dualities, that was initially used to study cosmologies, involves predicting the wavefunction of the universe from the dual theory [46, 50–53]. It has since been shown that this method is equivalent to the approach outlined above which uses the DW/C-correspondance [54].

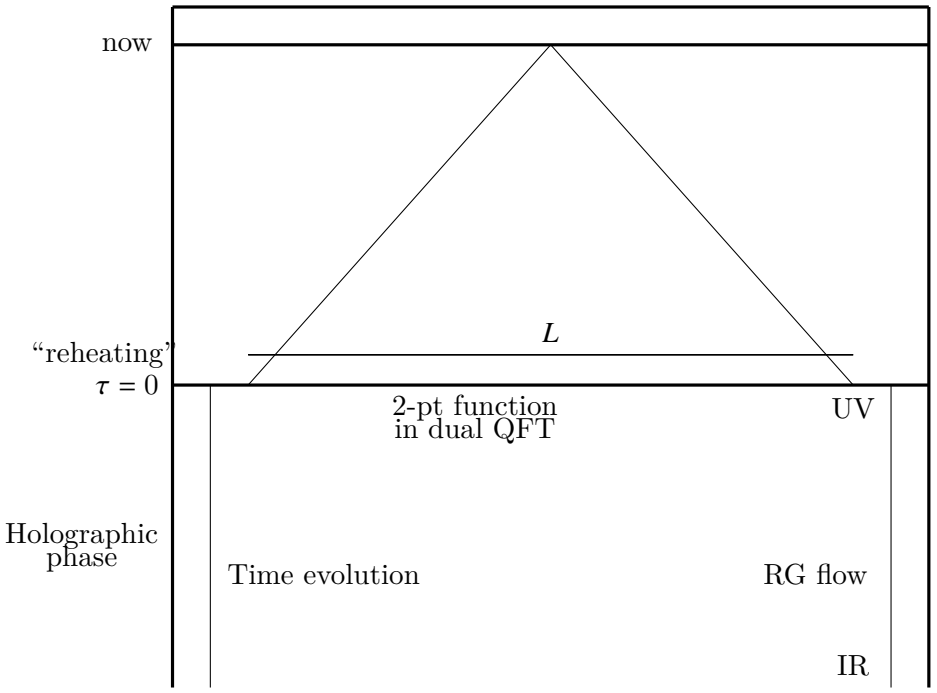

FIGURE 2.5: A schematic of the proposed holographic mapping between four-dimensional cosmology and a three-dimensional dual QFT. The past light cone from now to $\tau = 0$ is represented by diagonal straight lines. In the holographic phase correlations in the bulk are described by n-point functions of operators in the dual theory. Figure taken from [55].

2.7 Strongly Coupled Bulk Solutions

Within the holographic framework, theories that are strongly coupled in the bulk can only be understood by studying their associated weakly-coupled dual theories. It was argued that strongly-coupled asymptotically-de Sitter solutions are unstable [56]. It is therefore asyptotically power-law solutions that we now turn our attention to. On the boundary, these theories are QFTs on a Euclidean metric, with the property of *generalized conformal structure* (GCS) (see [57] and references therein). In standard conformally invariant theories, the action is invariant under conformal transformations. By contrast, in theories with GCS, we can assign a scaling dimension to all of the fields in the theory, such that all terms in the Lagrangian have the same behavior under rescaling. The dual theory should also admit a large-N expansion, where N is the rank of the gauge group. On the gravity side, the universe is known to be four-dimensional, so the dual theory must be three-dimensional. Given these constraints, along with Poincaré invariance and translational invariance, the most general action that can be

written down is [45, 58]

$$S = \frac{1}{g_{\rm YM}^2}\int d^3x\Bigg(\mathrm{Tr}\Bigg[\frac{1}{2}F_{ij}F^{ij} + \delta_{M_1M_2}\mathcal{D}_i\Phi^{M_1}\mathcal{D}^i\Phi^{M_2} + 2\delta_{L_1L_2}\bar{\psi}^{L_1}\gamma^i\mathcal{D}_i\psi^{L_2} \\ + \sqrt{2}\mu_{ML_1L_2}\Phi^M\bar{\psi}^{L_1}\psi^{L_2} + \frac{1}{4!}\lambda^{(1)}_{M_1M_2M_3M_4}\Phi^{M_1}\Phi^{M_2}\Phi^{M_3}\Phi^{M_4}\Bigg] \\ + \frac{1}{4!N}\lambda^{(2)}_{M_1M_2M_3M_4}\mathrm{Tr}\Big[\Phi^{M_1}\Phi^{M_2}\Big]\mathrm{Tr}\Big[\Phi^{M_3}\Phi^{M_4}\Big]\Bigg), \tag{2.26}$$

where $A_\mu = A^a_\mu T^a$, $F_{ij} = F^a_{ij}T^a$ and

$$\mathcal{D}_i\Phi^a = \delta^{ab}\partial_i\Phi^b - i(T^c)^{ab}A^c_i\Phi^b, \tag{2.27}$$

with T^a being the generators of $SU(N)$ with the normalization $\mathrm{Tr}(T^aT^b) = (1/2)\delta^{ab}$. F is the field strength tensor given by

$$F^a_{ij} = \partial_iA^a_jT^a - \partial_jA^a_iT^a + f^{abc}T^bT^c, \tag{2.28}$$

where f^{abc} are the structure constants of $SU(N)$. The fields Φ^M are scalars and ψ^L are fermions, while γ^i are the gamma matrices for the Clifford Algebra $Cl_{3,0}(\mathbb{R})$, and $\bar{\psi} = \psi^\dagger\gamma^0$. We have implicitly summed over the M_i and L_i indices, where $M_i \in \{1, 2, ..., \mathcal{N}_\Phi\}$, $L_i \in \{1, 2, ..., \mathcal{N}_\psi\}$, with $\mathcal{N}_\Phi$ being the number of scalars and $\mathcal{N}_\psi$ the number of fermions. Later in this project, we will use a redefinition of this action with the coupling $g = g^2_{\rm YM}N$, which we show here to refer back to later:

$$S = \frac{N}{g}\int d^3x\Bigg(\mathrm{Tr}\Bigg[\frac{1}{2}F_{ij}F^{ij} + \delta_{M_1M_2}\mathcal{D}_i\Phi^{M_1}\mathcal{D}^i\Phi^{M_2} + 2\delta_{L_1L_2}\bar{\psi}^{L_1}\gamma^i\mathcal{D}_i\psi^{L_2} \\ + \sqrt{2}\mu_{ML_1L_2}\Phi^M\bar{\psi}^{L_1}\psi^{L_2} + \frac{1}{4!}\lambda^{(1)}_{M_1M_2M_3M_4}\Phi^{M_1}\Phi^{M_2}\Phi^{M_3}\Phi^{M_4}\Bigg] \\ + \frac{1}{4!N}\lambda^{(2)}_{M_1M_2M_3M_4}\mathrm{Tr}\Big[\Phi^{M_1}\Phi^{M_2}\Big]\mathrm{Tr}\Big[\Phi^{M_3}\Phi^{M_4}\Big]\Bigg). \tag{2.29}$$

If one were to promote the coupling $g_{\rm YM}$ to a field with an appropriate conformal dimension, then the action (eq. 2.26) would be conformally invariant. That there is a single dimensionful coupling that can be brought outside the integral is another statement of the property of GCS. This condition precludes the inclusion of a mass-term $m^2\phi^2$ in the action since this would introduce a dimensionful coupling that could not be brought out as an overall factor. Motivated by the requirement for the model to have an N^2 large-N limit, a property shared by most well-understood holographic models and assumed in the derivation of eq. (2.24), the scalar and fermion fields are taken to be in the adjoint representation of the gauge group, $(\Phi^M) = (\Phi^M)^aT^a$ and $\psi^L = (\psi^L)^aT^a$. In the large-N limit, the double-trace term in the action is sub-leading, so from here on we take $\lambda^{(2)} = 0$ and refer to $\lambda^{(1)}$ as λ.

In this expression we have implicitly taken the metric tensor to be its known constant value, namely $g = \delta$ the Euclidean metric. However, to define the EMT, variations of this action with the metric must be taken at $g = \delta$. Explicitly,

$$T_{ij} \coloneqq \frac{2}{\sqrt{g}} \frac{\partial S_g}{\partial g_{ij}}\bigg|_{g_{ij}=\delta_{ij}}, \tag{2.30}$$

where S_g is the action as before but with a multiplicative factor $\sqrt{g(x)}$ in the integrand. There is one remaining ambiguity, however, the coupling of the scalar fields to the metric. This contributes a term

$$S_\xi = \frac{1}{2g_{\mathrm{YM}}^2} \int d^3x \sum_{M=1}^{N_\Phi} \xi_M R (\Phi^M)^2, \tag{2.31}$$

to the action, where R is the Ricci Scalar, which has to be incorporated in chapter 5 when we consider the renormalization of the EMT.

2.8 Perturbation Theory in the Boundary Theory

Two-point functions in theories with GCS are of the same form as those of CFTs, except the normalization constant is replaced by a function of a unitless coupling $\tilde{g} = g_{\mathrm{YM}}^2 x^{4-d}$ and the distance between the two points,

$$\langle \mathcal{O}(x)\mathcal{O}(0)\rangle = \frac{f_\Delta(\tilde{g})}{x^{2\Delta}}, \tag{2.32}$$

where Δ is the conformal dimension of the operator [57]. Equivalently in momentum space, we have

$$\langle \mathcal{O}(q)\mathcal{O}(-q)\rangle = q^{2\Delta-d} \tilde{f}_\Delta(\bar{g}), \tag{2.33}$$

where $\bar{g} = g_{\mathrm{YM}}^2/q^{4-d}$. In the case that $\mathcal{O}$ is a component of the EMT (which has scaling dimension 3 in three dimensions), the leading power of q is $2 \times 3 - 3 = 3$. In the $\bar{g} \ll 1$ perturbative limit the function $\tilde{f}_\Delta(\bar{g})$ can be expanded as

$$\tilde{f}_\Delta(\bar{g}) = f_1 + f_2 \bar{g} + ..., \tag{2.34}$$

where the coefficients f_i are calculated from the i^{th} order in the loop-expansion. However, renormalization of the EMT at the two-loop level produces new terms which are of the form $\bar{g} \log \bar{g}$. The largest power of N contributing to $\langle TT \rangle$, calculated by the contraction

over color indices at the one-loop level, is N^2, so that in the large-N limit [58]

$$A(q) \sim N^2 q^3, \qquad B(q) \sim N^2 q^3. \tag{2.35}$$

To make the large-N scaling explicit the two-loop perturbation theory result is expressed using the t'Hooft coupling $\chi = \bar{g}N$:

$$A(q,N) = q^3 N^2 f^T(\chi), \qquad B(q,N) = \frac{1}{4} q^3 N^2 f^S(\chi), \tag{2.36}$$

where for both f^S and f^T can be expanded as

$$f(\chi) = f_0(1 - f_1\chi \log\chi + f_2\chi + O(\chi^2)), \tag{2.37}$$

for small χ. In our three-dimensional model, each power of χ corresponds to a power of $1/q$, implying that each term in the series is more infrared divergent than the last. This divergence in perturbation theory is a result of the three-dimensional theory being super-renormalizable. However, there is a conjecture, discussed further in chapter 4, which states that non-perturbatively these divergences will disappear [59, 60]. The large N expansion is taken by keeping χ constant, while sending $N \to \infty$ [61]. Under the analytic continuation eq. (2.24) $\chi \to \chi$. This result, together with the fact that g^2_{YM}, N and q are real, implies the analytically continued expressions for A and B are:

$$A(q,N) = -iq^3 N^2 f^T(\chi), \qquad B(q,N) = -iq^3 N^2 \frac{1}{4} f^S(\chi). \tag{2.38}$$

These expressions give the perturbative result for the holographic dictionary

$$\Delta^2_S(q) = \frac{1}{4\pi^2 N^2 f^S(\chi)}, \qquad \Delta^2_T(q) = \frac{1}{\pi^2 N^2 f^T(\chi)}, \tag{2.39}$$

where expressions for f_0, f_1 and f_2 can be found in [58].

2.9 Testing Against the Night Sky

Cosmic inflation is one of the constituent theories of the empirical model Λ_{CDM}, also called the standard model of cosmology. Here CDM stands for "Cold Dark Matter" and Λ to the cosmological constant, which quantifies dark energy, the driver behind the observed acceleration of the universe's expansion. One of the great successes of

cosmic inflation, and by extension $\Lambda_{\rm CDM}$, is in predicting the near scale-invariance of the anisotropies in the Cosmic Microwave Background (CMB). Specifically, a scalar power spectrum of the form

$$\Delta_S(q)^2 = \Delta_0^2(q_*)\left(\frac{q}{q_*}\right)^{n_s-1}, \tag{2.40}$$

where q_* is an arbitrary reference scale, and n_s is predicted to be close to 1. If $n_s = 1$ then the spectrum is completely scale-invariant. Predictions of the near scale-invariance of the CMB power spectrum is one of the greatest achievements of the cosmic inflation framework. However, there are unresolved large-angle (small q) anomalies between the predictions of $\Lambda_{\rm CDM}$ and observations made by the WMAP and Planck experiments (see the Introduction of [62] and references therein). In cosmological datasets, two-point functions are measured according to multipoles l, for example $l = 2$ is the dipole of the CMB. Smaller q in momentum space corresponds to larger angles (smaller l) in the observed CMB. Within the model of inflation, these anomalies correspond to the earliest times in the inflationary period, when energy densities were the highest. The discrepancy in the low-multipole region is therefore expected, given that cosmic inflation is not UV-complete. Moreover this discrepancy can be seen as evidence that the energy scales for which cosmic inflation describes the observed universe well do not cover the whole of inflationary period - cosmic inflation is inconsistent.

A UV-complete theory such as holographic cosmology may be able to provide a better fit to the data of the CMB. To test holographic cosmology in the perturbative regime against WMAP [63] and later Planck data [64] a reparameterization of eq. (2.39) was used for coefficients of $f^S(\lambda)$:

$$\begin{aligned} hq_* &= f_1\lambda q, \\ \ln\left(\frac{1}{\beta}\right) &= \frac{f_2}{f_1} + \ln|f_1|, \\ \Delta_0^2 &= \frac{1}{4\pi^2 N^2 f_0}. \end{aligned} \tag{2.41}$$

This gives the prediction of the scalar-power spectrum of the form

$$\Delta_S^2(q) = \frac{\Delta_0^2}{1 + \left(\frac{hq_*}{q}\right)\ln\left(\frac{q}{\beta hq_*}\right) + O\left(\left(\frac{hq_*}{q}\right)^2\right)}, \tag{2.42}$$

which is a universal prediction of the class of holographic models with a strongly coupled-bulk sector. In fits done with this anzatz the higher order terms are dropped. A mixture of frequentist and Bayesian techniques have been used [65] to test both cosmic inflation (eq. 2.40) and holographic cosmology (eq. 2.42) to CMB data. In the case of the frequentist analysis the fit quality was found to be about the same in both cases, while the Bayesian analysis concluded that there was no significant preference for either of

the two models. In (eq. 2.42) it is clear the fitting function only holds for sufficiently large q (small angles), so only multipoles $l \geq 30$ were used. A comparison between the fits from holographic cosmology and cosmic inflation on the scalar-power-spectrum are shown in figure (eq. 2.6). Note that here the shape of the curve appears much more complex than the fitting functions we have seen. This is due to physics which occurs after the reheating-time (or equivalent time in holographic cosmology). The input of this physics is assumed to be the same in both the cosmic inflation and holographic cosmology scenarios, so we will not discuss it further here.

Since the parameters f_i are directly related to the field-contents of the dual theory these fitting parameters, together with the tensor-to-scalar ratio r, can be used to rule out particular theories. A theory with a majority of fermions was ruled out for example (see argument in [66] (version 1)). In a theory with only scalar fields and gauge fields, a choice of $\lambda_{M_1M_2M_3M_4} = \delta_{M_1M_2M_3M_4}$ in eq. (2.26) and $\xi_M = 0.133$ yields best-fit results that imply

$$N = 2995, \tag{2.43}$$

$$\mathcal{N}_\Phi = 23255, \tag{2.44}$$

justifying the use of a large-N limit [64]. Moreover, the value of h obtained implies the perturbative expansion is valid for l larger than roughly 35, meaning the choice of cut was mostly justified, though perhaps it was not conservative enough.

That the perturbative formulas for holographic cosmology cannot be used for the lowest multipoles of the CMB, where the largest deviations between cosmic inflation and observations are, is unfortunate. Moreover, for the scalar only theory the point at which the theory becomes non-perturbative was found to be at a higher multipole value of $l \approx 250$, meaning perturbative calculations are not valid for a wide range of the CMB spectrum in this model. There are however non-perturbative techniques which can be used instead to make predictions across all multipoles. One very successful non-perturbative technique is lattice-quantum-field-theory (lattice QFT), which is a class of quantum theories which are defined on a discrete space-time lattice. The advantage of studying such theories is that they are amenable to simulation using computer programs. This will be discussed in more detail in the next chapter, and the framework of lattice QFT will be used throughout the rest of this thesis.

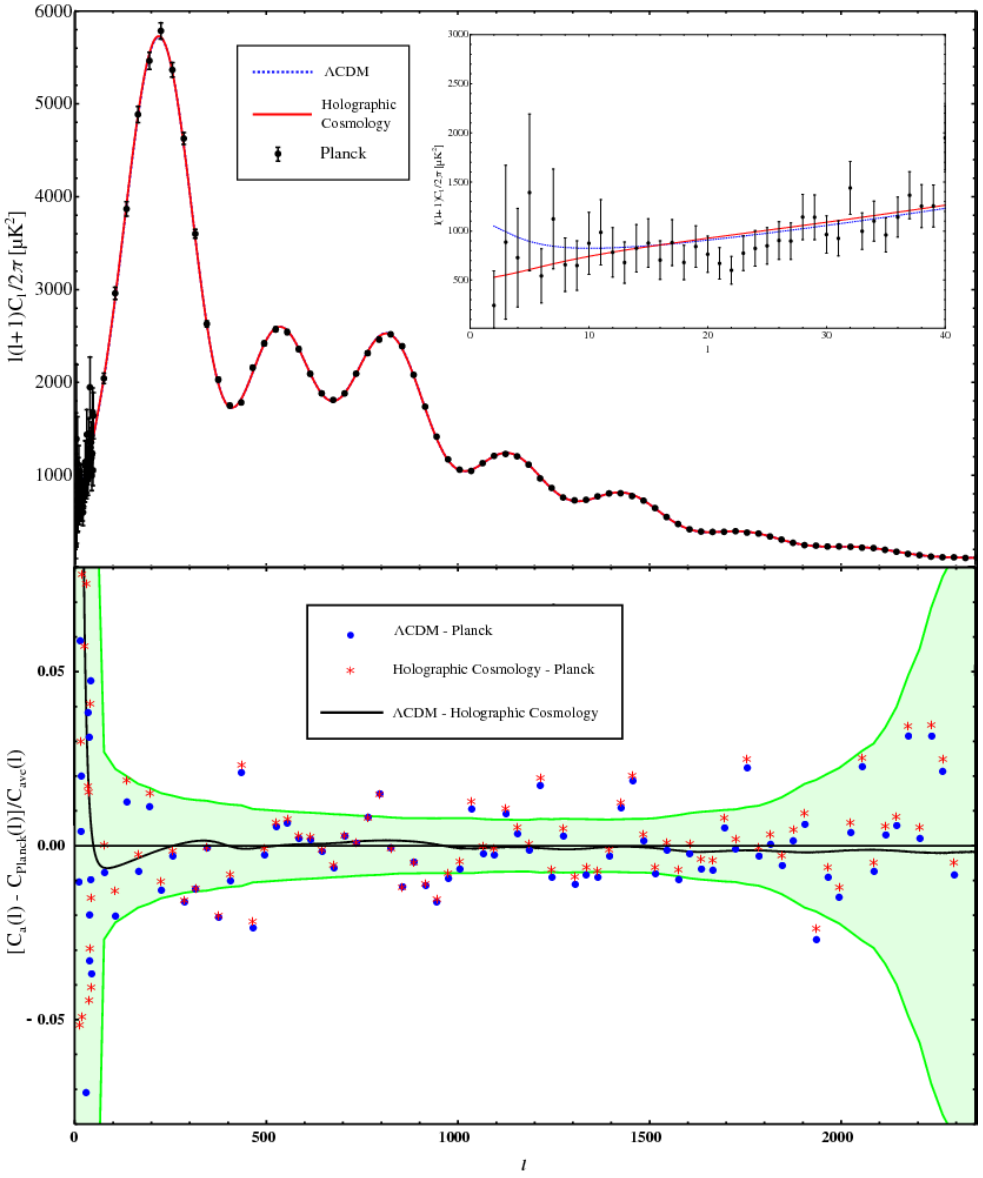

FIGURE 2.6: Present day CMB power spectrum as measured by the Planck telescope [22], with Λ_{CDM} and Holographic cosmology fits. Shown in the bottom figure are the deviations from the fit seen in both models. [65]

Chapter 3

Lattice QFT

3.1 Introduction

Many physical theories of interest, such as quantum chromodynamics (QCD), are studied outside the perturbative regime. These theories therefore suffer from ultraviolet infinities when quantities are calculated using them. To make physical predictions in these theories a regularization procedure must be used. This procedure can involve the use of counter terms and dimensional regularization, for example. An alternative technique that may be used to regulate a theory non-perturbatively is lattice QFT. Here space is discretized into a (hyper-)cubic lattice of separation a, which enforces a maximum momentum of π/a, thereby regulating the ultraviolet. On top of this, by using a lattice of finite extent, the infrared of the theory can also be regulated, with the smallest non-zero momentum being $2\pi/L$.

In this chapter we discuss how our holographic dual QFT can be regulated using a lattice and then simulated and analysed on a computer. In the first few sections, we will follow the standard textbook [67], but with application to our model. We will then outline simulation and analysis techniques which are core to the research done in this thesis. As an intermediate step to studying the full action (eq. (2.30)) we consider the pure-scalar-ϕ^4 action on a Euclidean metric

$$S[\phi] = \frac{N}{g}\int d^3x\, \mathrm{Tr}[(\partial^i\phi(x))(\partial_i\phi(x)) + m^2\phi^2(x) + \phi^4(x)], \tag{3.1}$$

where ϕ is in the adjoint of $SU(N)$ as in (eq. (2.30)) and a bare-mass term $m^2\phi^2$ has been added. By tuning this bare-mass parameter we are able to perform simulations at the non-perturbaitive massless point. How we choose the appropriate bare-mass will be discussed in chapter 4. The fields and the coupling g have been rescaled, such that the constant $\lambda = 1$. We represent the rescaled fields with a lowercase ϕ.

a	Lattice spacing, in units of length
N_L	The number of lattice points in each direction
$\vec{x}$	A two-vector representing position on a 2D lattice Λ'
x	A three-vector representing position on a 3D lattice Λ
$\vec{n}$	A two-vector comprising (n_1, n_2) where $n_i \in \{1, 2, ..., N_L\}$
n	A three-vector comprising (n_1, n_2, n_3) where $n_i \in \{1, 2, ..., N_L\}$
Λ'	The set of all $\vec{x} = (an_1, an_2)$ where $n_i \in \{1, 2, ..., N_L\}$
Λ	The set of all $x = (an_1, an_2, an_3)$ where $n_i \in \{1, 2, ..., N_L\}$
$\vec{i}$	Two-dimensional unit vectors in the i^{th} dimension
e_μ	A three-dimensional unit vector in the μ^{th}-direction
L	The size of the lattice in each direction in units of length
d	The dimensionality of the lattice
δ_i	Discrete derivative in the i^{th} dimension

TABLE 3.1: Notation used in this thesis for topics related to lattice QFT with adjoint scalars.

This lattice action is the simplest candidate model of the larger class of holographic dual QFTs discussed in chapter 2, and will allow us to do a first non-perturbative test of holographic cosmology against CMB data. If the true dual QFT also includes gauge fields, then studying the simpler scalar-only model will be pedagogical in understanding how to make predictions for models that also include gauge fields.

3.2 Placing Quantum Fields on a Spatial Lattice

In this section we discretize our holographic cosmology action. We have chosen to use a Minkowski metric in order to align our conventions with [67], however the calculations can similarly be performed with a Euclidean metric. The action is given by

$$S_M = \frac{N}{g} \int dt dx_1 dx_2 \left[\eta_{\mu\nu} (\partial^\mu \phi)(\partial^\nu \phi) - m^2 \phi^2 - \phi^4 \right], \tag{3.2}$$

where $\eta_{\mu\nu} = \text{diag}(+1, -1, -1)$. This can be obtained from eq (3.1) via the a Wick rotation ($x_0 = it$) and the identification $iS_M = -S_E$. Consider now separating the continuum Lagrangian density of our action into time and space derivatives,

$$\begin{aligned} \mathcal{L}(x) &= \frac{N}{g} Tr \left[\left(\frac{\partial \phi(x)}{\partial t} \right)^2 - \sum_{i=1}^{2} \left(\frac{\partial \phi(x)}{\partial x_i} \right)^2 - m^2 \phi(x)^2 - \phi^4(x) \right] \\ &= \frac{N}{g} Tr \left[\dot{\phi}^2 - (\nabla \phi)^2 - m^2 \phi^2 - \phi^4 \right] \\ &= \frac{N}{g} \left[\frac{1}{2} (\dot{\phi^b})^2 - \frac{1}{2} (\nabla \phi^b)^2 - \frac{1}{2} m^2 (\phi^b)^2 - Tr(\phi^4) \right], \end{aligned} \tag{3.3}$$

where in the second line we have dropped the x dependence as it is assumed. In the third line, terms quadratic in ϕ are separated into components ($\phi = \phi^b T^b$) and the normalization $Tr[T^b T^c] = (1/2)\delta^{bc}$ has been used. We have implicitly summed over

the index b, which runs from 1 to N_b, where $N_b = N^2 - 1$. The ϕ^4 term can also be decomposed into components, however, the expressions obtained by doing so are rather complex and uninformative for our present purpose. From now on we use the three-dimensional coordinate $x = (t, \vec{x})$, where $\vec{x} = (x_1, x_2)$ are the spatial dimensions and t is the time-like dimension,

$$S_M = \int dt \int d^2\vec{x} \mathcal{L}(t, \vec{x}). \tag{3.4}$$

We are now ready to place our theory onto a spatial lattice. For simplicity, we consider systems where the lattice spacing and size are the same in all directions. Explicitly $\vec{x} = a\vec{n}$, where $\vec{n} = (n_1, n_2, ..., n_d)$ and $n_i \in \{1, 2, ..., N_L\}$ where $N_L = L/a$ is the number of lattice sites in each direction. The boundary of the lattice can be defined in many ways, with the most common being periodic ("Pac-Man") boundaries where the points at $n_i = N_L + 1$ are identified with the points at $n_i = 1$. We denote the set of all points on the spatial lattice by Λ'. If a is small compared to the relevant physical length scale then the partial derivatives in the Hamiltonian can be well approximated by discretized derivatives,

$$\frac{\partial \phi^b(\vec{x})}{\partial x_i} = \bar{\delta}_i \phi^b(\vec{x}) + O(a^2), \tag{3.5}$$

where

$$\bar{\delta}_i X(\vec{x}) \coloneqq \frac{X(\vec{x} + a\vec{i}) - X(\vec{x} - a\vec{i})}{2a}, \tag{3.6}$$

is the *central difference* discrete derivative, and $\vec{i}$ is the unit vector in the i^{th} dimension. Other discrete derivatives can be defined, including the *forward* and *backward* discrete derivatives, given by

$$\delta_i^+ X(\vec{x}) \coloneqq \frac{X(\vec{x} + a\vec{i}) - X(\vec{x})}{a} \tag{3.7}$$

and

$$\delta_i^- X(\vec{x}) \coloneqq \frac{X(\vec{x}) - X(\vec{x} - a\vec{i})}{a} \tag{3.8}$$

respectively. These derivatives, however, give errors of order $O(a)$, and unless otherwise stated we will use the central difference derivative throughout the rest of this thesis. The continuum integrals $\int d^2\vec{x}$ are replaced by lattice sums,

$$\int d^2\vec{x} \to a^2 \sum_{\vec{x} \in \Lambda'}. \tag{3.9}$$

Following the standard second quantisation procedure, we define the canonical momenta by

$$\Pi^b(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}^b(x)} = \frac{N}{g}\dot{\phi}^b(x). \tag{3.10}$$

This allows us to define the classical time-dependent Hamiltonian,

$$\begin{aligned} H(t) &= \sum_{\vec{x}\in\Lambda'} \Pi^b((t,\vec{x}))\dot{\phi}^b((t,\vec{x})) - \int d^2\vec{x}\ \mathcal{L}((t,\vec{x})) \\ &= \frac{N}{g}\sum_{\vec{x}\in\Lambda'}\left[\frac{g^2}{2N^2}(\Pi^b)^2 + \frac{1}{2}\sum_{i=1}^{2}\left(\frac{\partial\phi^b}{\partial x_i}\right)^2 + \frac{1}{2}m^2(\phi^b)^2 + Tr(\phi^4)\right], \end{aligned} \tag{3.11}$$

The commutation relations with a lattice regulated theory are the same as the continuum ones, except extra factors of a are needed. These factors keep the units consistent in going from a continuous to a discrete Dirac-delta function. They are

$$\begin{aligned} [\hat{\phi}^b(\vec{x}), \hat{\Pi}^c(\vec{y})] &= ia^{-2}\delta(\vec{x}-\vec{y})\delta^{bc}, \\ [\hat{\phi}^b(\vec{x}), \hat{\phi}^c(\vec{y})] &= 0, \\ [\hat{\Pi}^b(\vec{x}), \hat{\Pi}^c(\vec{y})] &= 0, \end{aligned} \tag{3.12}$$

where we have taken $\hbar = 1$, $\delta(\vec{x}-\vec{y})$ is the delta-function, while δ_{bc} is equal to 1 when $b = c$ and 0 otherwise. These commutation relations imply

$$\hat{\Pi}^b(\vec{x}) = -\frac{i}{a^2}\frac{\partial}{\partial\phi^b(\vec{x})}. \tag{3.13}$$

This expression guides the formulation of a quantum theory with position dependent operators $\hat{\phi}^b(\vec{x})$ and $\hat{\Pi}^b(\vec{x})$, and Hamiltonian operator $\hat{H} = \hat{H}_0 + \hat{U}$ where

$$\hat{H}_0 = a^2\frac{N}{g}\sum_{\vec{x},b}\frac{1}{2}\left(-\frac{i}{a^2}\frac{\partial}{\partial\phi^b(\vec{x})}\right)^2 = a^2\frac{g}{N}\sum_{\vec{x},b}\frac{1}{2}\hat{\Pi}^b(\vec{x})^2, \tag{3.14}$$

$$\hat{U} = a^2\frac{N}{g}\sum_{\vec{x},b}\left[\frac{1}{2}\sum_{i=1}^{2}[\bar{\delta}_i\hat{\phi}^b(\vec{x})]^2 + \frac{1}{2}m^2\hat{\phi}^b(\vec{x})^2 + Tr(\hat{\phi}^4(\vec{x}))\right]. \tag{3.15}$$

To calculate correlators we will work with the orthonormal basis $|\phi\rangle$ where

$$\hat{\phi}^b(\vec{x})\,|\varphi\rangle = \varphi^b(\vec{x})\,|\varphi\rangle\,. \tag{3.16}$$

The states $|\varphi\rangle$ are normalized such that $\langle\phi'|\phi\rangle = \prod_{\vec{x},b}\delta\left(\phi'^b(\vec{x}) - \phi^b(\vec{x})\right)$. We use $|\Pi\rangle$ to donate the eigenvectors of the $\hat{H}_0$, which are defined by the wave-numbers $\Pi(\vec{x})$ at each

lattice site such that

$$\langle \phi | \Pi \rangle = \prod_{\vec{x},b} \sqrt{\frac{a^2}{2\pi}} \exp\left(i a^2 \Pi^b(\vec{x}) \phi^b(\vec{x}) \right), \tag{3.17}$$

where the factors of $\sqrt{\frac{a^2}{2\pi}}$ achieve the normalization

$$\int \prod_{\vec{x},b} d\Pi^b(\vec{x}) \, |\Pi\rangle \, \langle \Pi| = \hat{\mathbb{I}}. \tag{3.18}$$

Applying $\hat{H}_0$ upon the state $|\Pi\rangle$ reveals the eigenvalue of the operator to be

$$\frac{g}{2N} a^2 \sum_{\vec{x},b} (\Pi^b(\vec{x}))^2. \tag{3.19}$$

3.3 The Lattice Path Integral

Our Holographic action can now be discretized using the quantum states and Hamiltonian from the previous section through the Euclidean transfer matrix operator $e^{-t\hat{H}}$. This operator maps a quantum state $|\phi(t_0, \vec{x})\rangle$ at Euclidean time t_0 to a quantum state $|\phi'(t_0 + t, \vec{x})\rangle$ at Euclidean time $t_0 + t$. The partition function Z_T (for a given choice of T) is defined by

$$\begin{aligned} Z_T = \mathrm{Tr}[\exp(-T\hat{H})] &= \prod_{\vec{x},a} \left(\int_{-\infty}^{+\infty} d\phi^a(\vec{x}) \right) \langle \phi | \exp(-T\hat{H}) \, |\phi\rangle \\ &= \int \mathcal{D}\phi \, \langle \phi | \exp(-T\hat{H}) \, |\phi\rangle \,, \end{aligned} \tag{3.20}$$

where in the second line we have defined $\mathcal{D}\phi$ for convenience. It is understood that $|\phi\rangle$ is dependent on the values of $\phi(\vec{x})$ for all $\vec{x}$. To work with this rather complicated expression we use the Trotter product formula [68],

$$\exp(-T\hat{H}) = \lim_{M \longrightarrow \infty} (\hat{W}_\epsilon)^M, \tag{3.21}$$

where

$$W_\epsilon = \exp\left(\frac{-\epsilon \hat{U}}{2} \right) \exp\left(-\epsilon \hat{H}_0 \right) \exp\left(\frac{-\epsilon \hat{U}}{2} \right), \tag{3.22}$$

$$\epsilon = \frac{T}{M}. \tag{3.23}$$

The partition function can then be evaluated by the insertion of $M-1$ complete sets of states,

$$\begin{aligned} Z_T &= \int \mathcal{D}\phi_0 \, \langle \phi_0 | \exp\left(-T\hat{H}\right) | \phi_0 \rangle \\ &= \int \mathcal{D}\phi_0 \, \langle \phi_0 | \prod_{i=1}^{M} \hat{W}_\epsilon \, | \phi_0 \rangle \\ &= \int \mathcal{D}\phi_0 \, \langle \phi_0 | \prod_{i=1}^{M-1} \left(\int \mathcal{D}\phi_i \hat{W}_\epsilon^M \, | \phi_i \rangle \, \langle \phi_i | \right) \hat{W}_\epsilon \, | \phi_0 \rangle \, . \end{aligned} \tag{3.24}$$

Each of the M constituent $\langle \phi_i | \, \hat{W}_\epsilon \, | \phi_{i-1} \rangle$ is given by

$$\begin{aligned} &\langle \phi_i | \exp\left(\frac{-\epsilon \hat{U}}{2}\right) \exp\left(-\epsilon \hat{H}_0\right) \exp\left(\frac{-\epsilon \hat{U}}{2}\right) | \phi_{i-1} \rangle \\ &= \exp\left(\frac{-\epsilon (U[\phi_i] + U[\phi_{i-1}])}{2}\right) \langle \phi_i | \exp\left(-\epsilon \hat{H}_0\right) | \phi_{i-1} \rangle \, , \end{aligned} \tag{3.25}$$

where the remaining inner product can be evaluated by inserting the expression in eq. (3.18),

$$\begin{aligned} \langle \phi_i | \exp\left(-\epsilon \hat{H}_0\right) | \phi_{i-1} \rangle &= \int \mathcal{D}\Pi \, \langle \phi_i | \exp\left(-\epsilon \hat{H}_0\right) | \Pi \rangle \, \langle \Pi | \phi_{i-1} \rangle \\ &= \int \mathcal{D}\Pi \, \langle \phi_i | \Pi \rangle \exp\left(-\epsilon \frac{1}{2} \frac{g}{N} a^2 \sum_{\vec{x},b} \Pi^b(\vec{n})^2 \right) \langle \Pi | \phi_{i-1} \rangle \\ &= \left(\frac{a^2}{2\pi}\right)^{MN_L^2} \int \mathcal{D}\Pi \exp\left(-a^2 \sum_{\vec{x},b} \left(-\frac{1}{2} \epsilon \frac{g}{N} \Pi^b(\vec{x})^2 + i\Pi^b(\vec{x})(\phi_i^b(\vec{x}) - \phi_{i-1}^b(\vec{x})) \right) \right) \\ &= \left(\frac{\sqrt{N} a^2}{2\sqrt{g}\pi\epsilon}\right)^{MN_L^2} \exp\left(-\frac{1}{2\epsilon} a^2 \frac{N}{g} \sum_{\vec{x},b} ((\phi_i^b(\vec{x}) - \phi_{i-1}^b(\vec{x}))^2 \right), \end{aligned} \tag{3.26}$$

where $\mathcal{D}\Pi$ is defined analogously to $\mathcal{D}\phi$, and in the final line we have completed the square and performed the Gaussian integrals. Substituting this result back into Z_T, and remembering the periodic boundaries that identify ϕ_M and ϕ_0, leads to the pleasing result

$$Z_T = C \prod_{i=1}^{N} \left(\int \mathcal{D}\phi_i \right) \exp\left(-S_E[\phi]\right), \tag{3.27}$$

where

$$
\begin{aligned}
S_E &= \frac{1}{2}a^2\frac{N}{g}\sum_{i=1}^{M}\left[\sum_{\vec{x},b}\left(\frac{1}{\epsilon}(\phi_i^b(\vec{x})-\phi_{i-1}^b(\vec{x}))^2\right)+\epsilon(U[\phi_i]+U[\phi_{i-1}])\right] \qquad (3.28)\\
&= \epsilon a^2\frac{N}{g}\sum_{i=1}^{M}\left[\sum_{\vec{x},b}\frac{1}{2}\left(\frac{\phi_i^b(\vec{x})-\phi_{i-1}^b(\vec{x})}{\epsilon}\right)^2+U[\phi_i]\right]\\
&= \epsilon a^2\frac{N}{g}\sum_{i=1}^{M}\sum_{\vec{x},b}\left[\frac{1}{2}\left(\frac{\phi_i^b(\vec{x})-\phi_{i-1}^b(\vec{x})}{\epsilon}\right)^2+\frac{1}{2}\sum_{j=1}^{2}\left(\frac{\phi_i^b(\vec{x}+\vec{j})-\phi_i^b(\vec{x}-\vec{j})}{2a}\right)^2+\frac{1}{2}m^2\phi_i^b(\vec{x})^2\right.\\
&\qquad \left.+Tr(\phi^4(\vec{x}))\right]\\
&= \epsilon a^2\frac{N}{g}\sum_{i=1}^{M}\sum_{\vec{x},b}\left[\frac{1}{2}\left(\frac{\phi_{i+1}^b(\vec{x})-\phi_{i-1}^b(\vec{x})}{2\epsilon}\right)^2+\frac{1}{2}\sum_{j=1}^{2}\left(\frac{\phi_i^b(\vec{x}+\vec{j})-\phi_i^b(\vec{x}-\vec{j})}{2a}\right)^2+\frac{1}{2}m^2\phi_i^b(\vec{x})^2\right.\\
&\qquad \left.+Tr(\phi^4(\vec{x}))+O(\epsilon)\right],
\end{aligned}
$$

where in the final line we identify ϕ_{i+1} and ϕ_1 and change from a forward (time) discrete derivative to a central (time) discrete derivative for reasons that will become apparent shortly.

Inspired by how special relativity places space and time on an equal footing, we follow a standard procedure and set $a = \epsilon$ since they are both small values of equal mass-dimension. We will also here set $M = L$ so that the time extent of the lattice matches the spatial extent. This allows us to define a new 3-dimensional lattice $\Lambda = \{(\vec{x}, ai)|\vec{x} \in \Lambda', i \in \{1, 2, ..., L\}\}$ and a new measure $\mathcal{D}'\phi = \prod_{x\in\Lambda} d\phi^b(x)$ giving the compact result

$$
Z_T = \int \mathcal{D}'\phi \exp\left(S_E[\phi]\right), \qquad (3.29)
$$

where

$$
S_E[\phi] = a^3\frac{N}{g}\sum_{x\in\Lambda,b}\left(\frac{1}{2}\sum_{\mu=1}^{3}\left(\frac{\phi^b(x+ae_\mu)-\phi^b(x-ae_\mu)}{2a}\right)^2+\frac{1}{2}m^2\phi^b(x)^2+Tr(\phi^4(x))\right), \qquad (3.30)
$$

where $e_1 = (\vec{1}, 0)$, $e_2 = (\vec{2}, 0)$ and $e_3 = (\vec{0}, 1)$, we have assumed translation invariance of the potential V. We here on drop the " $'$ " superscript on $\mathcal{D}$. The subscript-E has been used up until this point to represent the fact that the action in eq. (3.31) is the discretized action of

$$
S_E[\phi] = \frac{N}{g}Tr\left[\int d^2x\left(g_{\mu\nu}\partial^\mu\phi\partial^\nu\phi+m^2\phi^2+\phi^4)\right)\right], \qquad (3.31)
$$

where $g_{\mu\nu} = \delta_{\mu\nu}$, the Euclidean metric. Our holographic dual QFT has thus been regularized, both in the infrared (through the lattice spacing a) and in the ultraviolet (through the lattice size L). This factor of $e^{-S_E[\phi]}$ can be seen as a probabilistic weighting of different field configurations ϕ. This probability distribution viewpoint is pivotal to the success of exploring lattice quantum field theories using computational techniques.

3.4 Two-Point Functions

Useful information content of quantum field theories can be extracted by calculating n-point functions of the theories, with lattice QFTs being no exception. Using computers, one can extract physical information of quantum field theories through the numerical evaluation of n-point functions of corresponding lattice QFTs. The most commonly studied type of n-point function are two-point functions. One can show [67] using the energy-Eigenbasis ($\hat{H}|n\rangle = E_m|n\rangle$) that in the limit $T \to \infty$ the two-point function in is given by

$$\langle O_2(t)O_1(0)\rangle \coloneqq \frac{1}{Z_T} Tr\left[e^{-(T-t)\hat{H}}\hat{O}_2 e^{-t\hat{H}}\hat{O}_1\right] = \sum_n \langle 0|\hat{O}_2|n\rangle\langle n|\hat{O}_1|0\rangle e^{-t\Delta E_n}, \tag{3.32}$$

where $\Delta E_n = E_n - E_0$ with E_0 being the energy of the vacuum. Two-point function information can, therefore, be used to study the spectrum of a quantum system. For example, it can be used to calculate the masses of bound states. Here $\hat{O}_i$ are operators that act at a certain value of t (they are diagonal in a time-basis $|t\rangle = \prod_{\vec{x}} |\phi(\vec{x}, t)\rangle$) and $O_i(t)$ are the eigenvalues of those operators on the basis $|t\rangle$. By choosing operators O with quantum numbers relating to the relevant symmetries, the desired physical spectrum can be probed.

Through a very similar calculation to the one outlined in section 3.3, two-point functions can be evaluated as

$$\langle O_2(t)O_1(0)\rangle = \frac{1}{Z_T}\int \mathcal{D}\phi e^{-S_E[\phi]} O_2(\varphi(t))O_1(\varphi(0)), \tag{3.33}$$

where $\varphi(t)$ is the set of all ϕ with a time coordinate value of t. Wall correlators[1] are two-point functions between operators O that simply average the fields ϕ at a given time t. In contrast, O can be localized to some region, e.g.

$$O(x;\epsilon) = O(\tilde{\phi}(x;\epsilon)), \qquad \tilde{\phi}(x;\epsilon) = \{\phi(y) \mid |y - x| \leq \epsilon\}. \tag{3.34}$$

[1] Or zero-momentum projected correlators or slice correlators (as we will use in sec. 6.5)

Given that our Euclidean action treats time and space coordinates in the same way, we can use operators localized in the space and time directions,

$$\langle O_2(t)O_1(0)\rangle = \frac{1}{Z_T}\int \mathcal{D}\phi e^{-S_E[\phi]}O_2(\phi(x_2))O_1(\phi(x_1)), \tag{3.35}$$

where the x_1 and x_2 are any points on the lattice satisfying $|x_2 - x_1| = t$. Under the assumption that field configurations can be generated according to the probability distribution

$$P[\phi] = e^{-S_E[\phi]}/Z_T, \tag{3.36}$$

then the two-point function can be estimated with

$$\langle O_2(t)O_1(0)\rangle_{\text{lat}} = \frac{1}{N_{\text{samples}}}\sum_{i=1}^{N_{\text{samples}}} O_2(\phi_i(x_2))O_1(\phi_i(x_1)), \tag{3.37}$$

where ϕ_i is the i^{th} sample of P, with N_{samples} taken in total. Note that P is also normalized ($\int \mathcal{D}\phi P[\phi] = 1$).

How these probability distributions are sampled is the subject of section 3.5. Alternative methods of integrating these integrals include quadrature techniques, such as Simpson's rule. However, as the integral has $L^3(N^2 - 1)$ integration variables, it is prohibitively costly to use such techniques as they scale exponentially with the number of integration dimensions. Note also that this technique naturally extends to n-point functions,

$$\langle O_1(x_1)O_2(x_2)...O_n(x_n)\rangle_{\text{lat}} = \frac{1}{N_{\text{samples}}}\sum_{i=1}^{N_{\text{samples}}} O_1(\phi(x_1))O_2(\phi_i(x_2))...O_n(\phi_i(x_n)). \tag{3.38}$$

3.5 Monte Carlo Sampling

We now turn our attention to generating samples ϕ that follow the probability distribution $P(\varphi)$. The vast majority of techniques to do this rely on a *Markov Chain Monte Carlo* (MCMC) method. These methods are a class of stochastic algorithm that generate new samples from previous ones in a way that is memoryless of any samples further in the past of the chain. In the context of lattice QFT, individual samples are called *configurations*, and the set of all configurations produced in a given run is called an *ensemble*. Throughout this thesis, we will often use the term "simulation" to refer to the process of producing an ensemble on a computer using lattice QFT and MCMC methods. If there are R unique states that can be sampled from P (call them ξ_i where $i \in \{1, 2, ..., R\}$), the corresponding MCMC process can be represented by an $R \times R$ transition matrix T. The component T_{ij} is the probability of transitioning from state ξ_i to state ξ_j. For the

samples to follow the correct probability distribution T must satisfy

$$\begin{aligned} 0 \leq T_{ij} \leq 1, \\ \sum_j T_{ij} = 1, \\ \sum_i T_{ij} P_i = \sum_i T_{ji} P_j, \end{aligned} \tag{3.39}$$

where $P_i = P(\xi_i)$. These equations are highly non-trivial. One common solution used in lattice research is the condition of *detailed balance*,

$$P_i T_{ij} = P_j T_{ji}. \tag{3.40}$$

This condition still leaves a lot of flexibility and many different algorithms satisfy it. One very popular algorithm is Metropolis-Hastings. Here T is separated into two parts: the selection T^0 and acceptance T^1 contributions, satisfying $T_{ij} = T^0_{ij} T^1_{ij}$. The probability of state j being selected from state i is given by T^0_{ij}, while T^1_{ij} gives the probability of the proposed new state being accepted as the next link in the Markov Chain. In order to satisfy detailed balance it is required that $T^1_{ij} = \min\left(1, \frac{T^0_{ji} P_j}{T^0_{ij} P_i}\right)$:

$$\begin{aligned} T_{ij} P_i &= T^0_{ij} T^1_{ij} P_i \\ &= T^0_{ij} \min\left(1, \frac{T^0_{ji} P_j}{T^0_{ij} P_i}\right) P_i \\ &= \min\left(T^0_{ij} P_i, T^0_{ji} P_j\right) \\ &= \min\left(\frac{T^0_{ij} P_i}{T^0_{ji} P_j}, 1\right) T^0_{ji} P_j \\ &= T^1_{ji} T^0_{ji} P_j \\ &= P_j T_{ji}. \end{aligned} \tag{3.41}$$

In this algorithm, the selection probability is chosen to be symmetric in the sense that $T^0_{ij} = T^0_{ji}$. Given that $P_i = e^{-S[\xi_i]}$, the acceptance probability is

$$T^1_{ij} = \min(1, e^{-\Delta S_{ij}}), \tag{3.42}$$

where $\Delta S_{ij} = S[\xi_j] - S[\xi_i]$. When the fields can take continuous values the transition matrix has infinite rank, however, the above results still hold. Starting with a predetermined initial state φ_0 and repeatedly applying the steps of suggestion and acceptance generates all configurations in the MCMC.

To sample the entire distribution P, the transition matrix cannot be decomposable as $T = R \oplus S \oplus ... \oplus Z$. If this were the case, the system would have disconnected regions of Monte Carlo space. In lattice simulations, effectively disconnected sets of states may

arise when the system being simulated is in a state of spontaneously broken symmetry. For example, in the Ising model at low temperatures there is a broken Z_2 symmetry. These different vacua are not strictly separated because there is a finite probability of transitioning between them. However, the probability of doing so is vanishingly small in the large volume limit, hence there is an effective separation. A Monte Carlo simulation where all possible states are connected (by non-vanishing probabilities) is called *ergodic*.

3.6 Autocorrelations and Binning

The simulation time t of a sampled state is defined by the total number of states in the chain that have preceded it, $\phi_t \xrightarrow{T} \phi_{t+1}$, $\phi_0 = \Phi$, where Φ is the first saved configuration. The first saved configuration usually differs from the initial configuration, as the first T_{init} configurations generated are removed. This is done since the distribution P is only sampled in an unbiased way in equilibrium, and initially the MCMC is not in equilibrium.

States that are separated by small differences in the simulation time are correlated to each other. This correlation is an issue in that the configurations are not independent samples of the distribution P. A two-point correlator in simulation time for the operator $O(\phi)$ can be defined by

$$\Gamma_O(t) = \langle O_i O_{i+t} \rangle - \langle O_i \rangle^2, \tag{3.43}$$

where the angled brackets imply an average over the index i, the index of configurations. Similarly to two-point correlators in physical space, this correlation decays exponentially for large t,

$$\Gamma_O(t) \propto \exp\left(-\frac{t}{\tau_{O,\text{exp}}}\right), \tag{3.44}$$

with an a priori unknown decay length $\tau_{O,\text{exp}}$. In this thesis, autocorrelations are calculated using the framework of [69]. Here the quantity

$$C_O = \sum_{t=-\infty}^{+\infty} \Gamma_O(t), \tag{3.45}$$

is used to estimate the *integrated-autocorrelation time* $\tau_{O,\text{int}}$ by

$$\tau_{O,\text{int}} = \frac{C_O}{2\Gamma_O(0)}. \tag{3.46}$$

Estimates of the mean of O have a variance that goes like $2\tau_{O,\text{int}}\Gamma(0)/N$ [70], a factor of $2\tau_{O,\text{int}}$ larger than expected. Therefore, $N/2\tau_{O,\text{int}}$ can be thought of as the effective number of independent samples in an ensemble (for the estimation of O). In order to estimate the integrated autocorrelation time it is necessary to truncate the expression

in equation (3.45) at $t = t_{\max}$. This is because each contribution to this sum carries with it associated statistical noise, while the signal decays exponentially with autocorrelation time. The truncated expression is then

$$\tau_{\mathcal{O},\text{int}}(t_{\max}) = \frac{1}{2\Gamma_{\mathcal{O}}(0)} \sum_{t=-t_{\max}}^{t_{\max}} \Gamma_{\mathcal{O}}(t). \tag{3.47}$$

This truncation does unfortunately come with an associated systematic error that scales like $\exp(-t_{\max}/\tau)$, while the statistical error in the estimate scales like $\sqrt{t_{\max}/N}$. The truncation time $t_{\max}$ can therefore be chosen at a value that minimizes the sum of the statistical and systematic errors,

$$\exp\left(\frac{-t_{\max}}{\tau(t_{\max})}\right) + 2\sqrt{\frac{t_{\max}}{N}}, \tag{3.48}$$

where the quantity $\tau(t_{\max})$ is the effective autocorrelation time assuming $\tau_{\mathcal{O},\text{int}}$ was calculated by summing over an exponential decay,

$$2\tau_{\mathcal{O},\text{int}}(t_{\max}) = \sum_{-\infty}^{+\infty} \exp\left(\frac{-S|t|}{\tau(t_{\max})}\right), \tag{3.49}$$

for some choice of the constant S. The minimization procedure must, therefore, be done iteratively: alternately calculating $t_{\max}$ and $\tau(t_{\max})$.

The presence of autocorrelation implies that configurations cannot be treated as independent for the sake of calculating statistical errors. Instead data must be *binned* into contiguous chunks of size S_{bin} in simulation time, where $S_{\text{bin}} > \tau_{\mathcal{O},\text{int}}$.

3.7 Hybrid Monte Carlo

The autocorrelation time is highly dependent not only on the physical parameters of the system but also on the specific algorithm used to generate the next state in the Monte Carlo chain. There is no best algorithm appropriate for all lattice QFTs, and the continued creation and development of algorithms remains a core part of lattice research.

To reduce autocorrelation times, large jumps between configurations in the MCMC chain should be made. However, to do this, fields at many different positions on the lattice have to be randomly changed concurrently. This will likely lead to a configuration with a very different action, giving a small probability of acceptance and in turn reduced algorithm efficiency.

In contrast to the heatbath algorithm (which is discussed in section 3.8), *hybrid Monte Carlo* (HMC) is a technique that proposes globally new field states. The algorithm uses the classical equations of motion of the system to suggest states with the same value

of the action [71]. In practice, the procedure is not exact due to numerical integration, so the action will deviate, meaning that a Metropolis-Hastings accept/reject step is still required. For each of our field variables $\phi_i^b(x)$, a conjugate momenta $\Pi_i^b(x)$ is defined, from which a classical Hamiltonian can be defined (as in section 3.2). A random sample of all conjugate momenta variables is made according to the distribution

$$P[\Pi_i^b(x) = p] \propto \exp\left(\frac{1}{2}p^2\right), \tag{3.50}$$

which nicely preserves the $\Pi \to -\Pi$ symmetry. Numerical techniques, for example a leap-frog integrator, are used to evolve the system according to Hamilton's equations of motion. If this evolution was done with infinite precision, a configuration with a Hamiltonian of the same value as that of the original state would be obtained. However, due to using a finite step size δt in the numerical evolution of these equations, there will be a change in the Hamiltonian, ΔH. As δt is decreased, the average size of ΔH should decrease. It can be shown [67], using a proof similar to eq. (3.41), that this method generates the correct distribution of fields if an acceptance probability of

$$\min\left(1, e^{-\Delta H}\right) \tag{3.51}$$

is used. When determining the length of the trajectory to use (e.g. the number of steps times the size of each step) different considerations must be taken into account. If the trajectory is too long ΔH will likely be large, giving a low acceptance probability, dropping the efficiency of the algorithm. This can be made up for by using smaller step sizes, however this comes with a computational cost. On the other hand, repeated HMC steps explore the phase space like a random walk, so it is important to make sure the trajectories are long enough to move through a large region of phase space. Otherwise there is a danger that the algorithm might only explore a small subset of the overall phase space.

3.8 Heatbath Overrelaxation

The global update steps of the hybrid-Monte Carlo algorithm can be expensive. An alternative method that can be used for local theories[2] is the heatbath overrelaxation algorithm [72–75]. Consider our holographic action implemented on the lattice using forward derivatives, (eq.(3.31)),

$$S_{\text{lat}} = \frac{N}{g}a^3 \sum_{x,b}\left[\frac{1}{2}\sum_{\mu}\left(\frac{\phi^b(x + ae_\mu) - \phi^b(x)}{a}\right)^2 + \frac{1}{2}m^2\phi^2(x) + Tr[\phi^4(x)]\right]. \tag{3.52}$$

[2]To simulate lattice QFTs with fermions, these Grassmann-valued fields must be integrated out to give an effective theory of the remaining scalar fields. This effective theory is highly non-local and therefore cannot be simulated using heatbath techniques.

Setting $\tilde{g} = ag$, $\tilde{m} = am$ and $\tilde{\phi}_n = a\phi_x$ gives the action in terms of unitless parameters,

$$S_{\text{lat}} = \frac{N}{\tilde{g}} \sum_{x,b} \left[\sum_{\mu} \left(\tilde{\phi}^b(x + a\hat{\mu}) - \tilde{\phi}^b(x) \right)^2 + \frac{1}{2}\tilde{m}^2 \tilde{\phi}^2(x) + Tr[\tilde{\phi}^4(x)] \right]. \tag{3.53}$$

Generating random numbers to sample new configurations is one of the most computationally intensive parts of a lattice simulation. It is therefore important to phrase the sampling problem in a way that we can utilize the most efficient forms of random number generation. Efficient methods exist for sampling Gaussian probability distributions. Since the kinetic term of the action is quadratic in the fields we can naturally pull out a piece of e^{-S} that is Gaussian distributed in field variables.

In the heatbath algorithm, changes are proposed to the fields at a specific point on the lattice, $x = x'$ (e.g. a completely localized change). We therefore only need to consider terms in the action that depend on the field $\phi(x')$ (as other terms will cancel in ΔS). We call this reduced action $S'[\phi_{x=x'}, \phi_{x\neq x'}]$. It is known that sampling the field at site $x = x'$ according to

$$P_0(\phi') \propto e^{-S'_0[\phi'_{x=x'}, \phi_{x\neq x'}]}, \tag{3.54}$$

and keeping the proposed field with acceptance probability

$$P_1(\phi') = \mu e^{-S'_1[\phi'_{x=x'}, \phi_{x\neq x'}]}, \tag{3.55}$$

where $S'_1 = S - S'_0$, gives a Monte Carlo process with the correct distribution of the fields [76]. We have used an apostrophe to denote the proposed new configuration ϕ'. The constant μ is chosen so that $P_1(\phi')$ is at most 1 over the range of values that the fields can take. Choosing S'_1 to contain the potential terms we have

$$S_1(\phi(x')) = \frac{N}{g} \left(\frac{1}{2} m^2 \phi'^b(x') + Tr[\phi'(x')] \right). \tag{3.56}$$

After some rearrangements, the remaining (kinetic) term can be written as [74]

$$S_{\text{kin}} = \frac{N}{g} \left(d(\phi'^b(x'))^2 - \phi'^b(x') \sum_{x''\in \text{NN}} \phi'^b(x'') \right), \tag{3.57}$$

where NN is the set of nearest neighbor lattice sites to x'. Subtracting constants from the action we have

$$P_0(\phi') \propto \exp\left(-\frac{N}{2g} d_\alpha \sum_a (\phi'^b(x') - C^b)^2 \right), \tag{3.58}$$

$$P_1(\phi') \propto \exp\left(-\frac{N}{g} \left(\phi'^2(x') + \frac{m_\alpha^2}{2\lambda} \right)^2 \right), \tag{3.59}$$

where

$$\begin{aligned} C^b &:= \frac{1}{d_\alpha} \sum_{x'' \in \mathrm{NN}} \phi^b(x''), \\ d_\alpha &:= 2d + \alpha, \\ m_\alpha^2 &:= m^2 - \alpha, \end{aligned} \tag{3.60}$$

and the constant α is tuned to give the best average acceptance probability. The expression for P_0 implies that $N^2 - 1$ samples of the distribution $\sim \exp\left(-(M^b)^2/2\right)$ can be taken to get the $N^2 - 1$ components of ϕ', with

$$\phi'^b = \sqrt{\frac{g}{d_\alpha N} M_\alpha} + C^b. \tag{3.61}$$

To reduce autocorrelation times, it is often useful to perform overrelaxation steps [77]. These are micro-canonical updates of the system which send

$$\phi \to \phi' = 2b - \phi. \tag{3.62}$$

These are accepted with a Metropolis-Hastings acceptance probability of

$$P = \min\left(1, \exp\left(S(\phi) - S(\phi')\right)\right). \tag{3.63}$$

It is not possible to use solely overrelaxation steps as they are not ergodic, so in practice, heatbath and overrelaxation steps must be alternated. These methods are extremely scalable for local actions such as ours through "checkerboarding" the lattice. In this technique the lattice is split into white and black sites, where adjacent sites are opposite color - like the squares of a checkerboard. The black sites only receive interaction contributions from the white sites and likewise the white sites only receive interaction contributions from the black sites. Therefore, new configurations can be proposed by updates done in parallel for all black sites and then updates done in parallel for all white sites.

3.9 Multi-Histogram Reweighting

In our lattice action (eq. (3.31)), the choice of the bare-parameters m^2 and g continuously alters the theory's physical predictions. To extract estimates of operators $\mathcal{O}$ from our lattice ensembles, an average of $\mathcal{O}$ over all configurations in the ensemble is taken. In general, the expectation value of this average $\mathrm{E}\left[\langle \mathcal{O} \rangle\right]$ is a complicated non-linear function of g and m^2 (as well as the lattice size L, due to finite-size scaling effects). Multi-histogram reweighting can be used [78] to explore these functions in detail without running many simulations across a fine comb of m^2 and g. This technique allows lattice averages to be interpolated between different simulation points. The mathematics

behind reweighting is beautiful in its simplicity. First, we define the action S_m as

$$\begin{aligned} S_{m^2}[\phi] &\coloneqq \mathrm{Tr}\left[\frac{N}{g}\int d^3x\left[(\partial_\mu\phi(x))^2 + m^2\phi^2(x) + \phi^4(x)\right]\right] \\ &= (m^2 - m_0^2)\frac{N}{g}\int d^3x \mathrm{Tr}[\phi^2(x)] + S_{m_0^2}[\phi]. \end{aligned} \tag{3.64}$$

This rearrangement of variables is incredibly useful if a lattice simulation has been performed at the mass point m_0^2 and one wishes to study an operator at a different mass point m^2. Let $\langle O\rangle_{m^2}$ represent the lattice one-point estimate for the operator O that we would get from a lattice ensemble at mass m^2.

$$\begin{aligned} \langle O\rangle_{m^2} &= \frac{\int \mathcal{D}\phi e^{-S_{m^2}[\phi]}O}{\int \mathcal{D}\phi e^{-S_{m^2}[\phi]}} \\ &= \frac{\int \mathcal{D}\phi e^{-(m^2-m_0^2)\frac{N}{g}\int d^3x \mathrm{Tr}[\phi^2(x)] - S_{m_0^2}[\phi]}O}{\int \mathcal{D}\phi e^{-(m^2-m_0^2)\frac{N}{g}\int d^3x \mathrm{Tr}[\phi^2(x)] - S_{m_0^2}[\phi]}} \\ &= \frac{\langle e^{-(m^2-m_0^2)\frac{N}{g}\int d^3x \mathrm{Tr}[\phi^2(x)]}O\rangle_{m_0^2}}{\langle e^{-(m^2-m_0^2)\frac{N}{g}\int d^3x \mathrm{Tr}[\phi^2(x)]}\rangle_{m_0^2}} \\ &= \frac{\langle RO\rangle_{m_0^2}}{\langle R\rangle_{m_0^2}}, \end{aligned} \tag{3.65}$$

where the reweighting factor

$$R = e^{-(m^2-m_0^2)\frac{N}{g}\int d^3x Tr[\phi^2(x)]}, \tag{3.66}$$

has been defined for convenience.

The connection of this technique to histograms is as follows: consider the histogram of evaluations of O in an ensemble at mass m_0^2. The one-point function of O can be obtained by taking an average over this histogram. To get to the histogram of O at m^2, each contributing data point O_i to the m_0^2 histogram is scaled by $R[\phi_i]$. The prefix "multi" refers to the possibility of using multiple sources m_0^2 for reweighting to m^2 and performing a weighted average of these estimates. As $|m^2 - m_0^2|$ increases, however, the discrepancy in R between different points in the original m_0^2 histogram will increase, and so our evaluation of $\langle O\rangle_{m^2}$ will be effectively determined by fewer and fewer configurations. This will lead to an increasing standard deviation in our estimation of the operator. The larger the extrapolation, the larger the error on the error becomes, until the method can no longer be trusted. To determine the validity of the reweighting method, we consider choosing $O = \int d^3x Tr[\phi^2(x)]$. We then define the unitless ratio

$$z(m_0^2, m^2) = \frac{|\mu_{m_0^2} - \mu_{m^2}|}{\sigma}, \tag{3.67}$$

where

$$\sigma = \langle O^2 \rangle_{m_0^2} - \langle O \rangle^2_{m_0^2}, \tag{3.68}$$
$$\mu_{m^2} = \langle O \rangle_m.$$

As z increases the reweighting becomes less reliable. Denoting the reweighting from m_0^2 to m^2 by $\langle O \rangle_{m_0^2 \longrightarrow m^2}$, we get an overall estimate of O_m by taking the weighted average

$$\bar{O}_{m^2} = \sum_{m_0^2} \langle O \rangle_{m_0^2 \longrightarrow m^2} w(m_0^2, m^2), \tag{3.69}$$

where $\sum_{m_0^2} w(m_0^2, m^2) = 1$. In [79] the authors recommend using using a cutoff r in the weighted average,

$$r(m_0^2, m^2) = 1 - \Theta(z(m^2, m_0^2) - z_{\max}), \tag{3.70}$$

where Θ is the standard Heavyside Step function, and $z_{\max}$ is chosen to be at half the full width at half maximum (FWHM), a value of ~ 1.18. In the IR-finiteness chapter (Chap. 4) we discuss other choices of w.

3.10 Errors and Biases

Estimates of quantities from lattice data suffer from three types of deviation that need to be considered. The first is the bias introduced by the nature of the theory being defined with finite lattice spacing and extent: cut-off effects, finite-size effects, and effects due to the breaking of rotational symmetry. Lattice physicists attempt to remove these effects when taking an infinite volume continuum limit, however, this process introduces systematic errors. The finite-size effects in our theory are explored in detail in the IR-finiteness chapter (Chap. 4), while the cut-off effects due to the non-zero lattice spacing are discussed in chapter 5. The second kind of deviation that needs to be considered is systematic errors due to the choice of fitting routines and meta-parameters, which we also deal with extensively in chapter 4. The last source of deviation is the statistical error resulting from the finite number of configurations N in a lattice ensemble. If a quantity of interest X is a primary quantity then, for each configuration i, X can be evaluated as X_i. In this case the estimator

$$\sigma_X = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (X_i - \langle X \rangle)^2}, \tag{3.71}$$

gives the unbiased estimate of the standard deviation of X. Assuming zero autocorrelation between estimates X_i then the statistical error of the mean $\langle X \rangle$ scales like $\sigma_X / \sqrt{N}$.

In practice, because the autocorrelation time τ is usually non-zero, data are grouped into bins larger than τ, where the bin averages

$$\tilde{X}_j = \frac{1}{S_{\text{bin}}} \sum_{i=S_{\text{bin}}j+1}^{S_{\text{bin}}(j+1)} X_i \tag{3.72}$$

are treated as independent contributions to the estimator $\langle X \rangle$. This gives a statistical error of $\sigma_{\tilde{X}}/\sqrt{N_B}$, where $N_B = N/S_{\text{bin}}$ is the number of bins[3].

Before proceeding it is useful to define *empirical probability distributions*. Consider the primary quantities X^α, where α indexes the quantities. Each configuration ϕ_i gives us a sample of these primary quantities, $(X_i^0, X_i^1, ..., X_i^n)$, where n is the number of primary quantities. Let $\mathbf{X} = (X^0, X^1, ..., X^n)$. We can define an empirical probability distribution $\rho_{\mathbf{X}}$ of $\mathbf{X}$. Samples of this distribution $S^{(j)}$ are obtained by taking

$$S^{(j)} = \mathbf{X}_{s_j} = (X_{s_j}^0, X_{s_j}^1, ..., X_{s_j}^n), \tag{3.73}$$

where s_j is a randomly selected integer from the set $\{1, 2, ..., N\}$. This is a very complicated way of saying that the empirical probability distribution of $\mathbf{X}$ is defined such that samples of it are random selections from the set of $\mathbf{X}$ in the ensemble. In the $N \to \infty$ limit the empirical distribution of $\mathbf{X}$ approaches the true distribution of $\mathbf{X}$.

Consider instead a quantity Y that is a non-linear combination of primary quantities, calculated using the whole ensemble. It may be impossible to propagate errors through Y due to its functional form and unknown correlations between contributing terms. The true value of Y is a function of the true probability distribution $\rho_{\mathbf{X}}$ of the primary quantities $\mathbf{X}$. With a finite number of configurations in our ensemble, the distribution $\rho_{\mathbf{X}}$ can be approximated by an empirical sample distribution $\rho'_{\mathbf{X}}$. In doing so, information is lost, and an estimate $Y_{\rho'}$ calculated from this sample distribution will differ from the true value $Y_{\rho'}$. We denote the expectation of this deviation by σ_Y, where σ_Y is a priori unknown.

In the bootstrap resampling paradigm [80] σ_Y is estimated by repeatedly resampling the empirical distribution $\rho'_{\mathbf{X}}$, to get new empirical probability distributions $\rho''_{\mathbf{X}^k}$. Each $\rho''_{\mathbf{X}^k}$ is the empirical distribution of $\mathbf{X}^k$, where

$$\mathbf{X}_i^k = \mathbf{X}_{s_i}, \ \forall i \tag{3.74}$$

where again s_j is a randomly selected integer from the set $\{1, 2, ..., N\}$. We can consider $\mathbf{X}^k$ as a *resampling* of the original $\mathbf{X}$. From each sample k we can calculate a value for

[3] If the autocorrelation time is zero, then our estimates of the standard deviation with and without binning will agree on average (because $E[\sigma_{\tilde{X}}] = E[\sigma_X]/\sqrt{S_{\text{bin}}}$, where E represents the expectation value)

Y,

$$Y^k = Y(\rho''_{\mathbf{X}^k}). \tag{3.75}$$

From these samples, it is then possible to calculate a mean and a standard deviation,

$$\mu_{Y'} = \sum_{k=1}^{N_{\text{boot}}} Y^k, \tag{3.76}$$
$$\sigma_{Y'} = \sqrt{\frac{1}{N_{\text{boot}} - 1} \sum_{k=1}^{N_{\text{boot}}} (Y^k - \mu_{Y'})},$$

The key insight in bootstrap methods [80] is the following: in the large-statistics limit $\sigma_{Y'} \to \sigma_Y$. We can therefore use the sample standard deviation $\sigma_{Y'}$ as an estimator of σ_Y. Intuitively, this makes sense because the information lost going from ρ to ρ' is the same size as the information lost in the resampling from ρ' to ρ''.

3.11 Lattice Gauge Theories

While the research presented in this thesis largely concerns pure scalar-matrix-ϕ^4 theory (eq. (3.1)), the intention of the LatCos collaboration is to include gauge fields in the action, in order to test a larger class of holographic models. A basic understanding of lattice gauge theory is also needed to understand the origins of the multilevel algorithm. For these reasons, we briefly outline the method used to define lattice QFTs with gauge fields. To preserve gauge-invariance of the action, it is necessary to introduce $SU(N)$-valued link-variable fields $U_\mu(x)$ for all axes μ of the lattice. These variables transform like

$$U_\mu(x) \to U_\mu(x) = \Omega(x) U_\mu(x) \Omega(x + a e_\mu)^\dagger \tag{3.77}$$

under a gauge transformation Ω [67]. A path of such variables can be defined by taking a product

$$P[U] = U_{\mu_1}(x) U_{\mu_2}(x + a e_{\mu_1}) U_{\mu_3}(x + a e_{\mu_2} + a e_{\mu_1}) \cdots U_{\mu_n}(y), \tag{3.78}$$

where μ_i can take negative values, and $U_{-\mu}(x) = U^\dagger(x - a e_\mu)$. The final point on the path y must be given by

$$y = x + a \sum_{i=1}^{n-1} e_{\mu_i}. \tag{3.79}$$

Variables of the form $P[U]$ have the benefit of transforming simply under gauge transformations Ω as

$$P \rightarrow \Omega(x) P \Omega(y)^{\dagger}. \tag{3.80}$$

Gauge-invariant quantities can be constructed by taking the trace of path variables with $x = y$. It has been shown that [81] the gauge field action,

$$S_G[U] = \frac{2}{g^2} \sum_{x \in \Lambda} \sum_{\nu} \sum_{\mu < \nu} \mathfrak{Re}[Tr(\mathbb{I} - U_{\mu\nu}(x))], \tag{3.81}$$

becomes Yang-Mills theory in the continuum ($a \rightarrow 0$) limit. The fields $U_{\mu\nu}$ are referred to as *plaquettes* have the expression

$$\begin{aligned} U_{\mu\nu}(x) &= U_{\mu}(x) U_{\nu}(x + a e_{\mu}) U_{-\mu}(x + a e_{\mu} + a e_{\nu}) U_{-\nu}(x + a e_{\nu}), \\ &= U_{\mu}(x) U_{\nu}(x + a e_{\mu}) U_{\mu}(x + a e_{\nu})^{\dagger} U_{\nu}(x)^{\dagger}, \end{aligned} \tag{3.82}$$

where $\mathbb{I}$ is the identity matrix.

Chapter 4

Infrared Finiteness

The dual quantum field theory (eq. (3.1)) is a super-renormalizable theory, which is a statement that the coupling constants appearing in the action have positive mass dimension. Naïvely we expect it to make physical predictions that are well-behaved (finite) in the ultraviolet (UV) limit but suffer from divergences in the infrared (IR). If true, the theory would be ill-defined and unable to be used to make physical predictions. This divergent behavior can be seen by studying these models in perturbation theory - each loop order from two-loop onwards is divergent in the infrared, with the power of the divergence increasing at each order. There is however cause for optimism. Two papers published early in the 1980s hypothesize that non-perturbative effects resolve the infrared divergence in this class of super-renormalizable theory [59, 60].

There are two aims in this chapter. The first is to test the hypotheses that the theory is well-defined in the infrared. Confirming this is essential to verify the viability of using holographic models to describe cosmology. The second aim, conditional on IR-finiteness being demonstrated, is to estimate the infinite-volume critical-mass point. This knowledge would allow us to perform simulations at the non-perturbative critical point, where the theory exhibits generalized conformal structure.

To achieve these goals, we have generated ensembles across a range of values of the bare mass m^2, the coupling g (0.1 to 0.6), the lattice size L (8 to 128), and the rank N (2, 3, 4 and 5). In a massless theory, the relevant volume is given by $x = gL$, and the range of g and L we use offers us a large range of volumes x. The infrared finiteness of the theory can be tested through many different quantities. In a finite volume the critical mass is expected to be at a finite value, however, it is possible that as we remove the infrared-cutoff (increase the size of our lattice) the critical-mass point diverges. We therefore perform a finite-size-scaling study on the finite-volume critical mass as we vary g, L, and N. To determine the value of the finite-volume critical mass, we study the phase structure of the theory. In the infinite-volume limit (if the theory is IR-finite), there is a phase transition between symmetric and broken phases. When regulating the

theory using a lattice, a bare mass is introduced in the action. The critical mass $m^2_{c,\infty}(g)$ is the bare mass at the symmetric-broken phase transition in the infinite-volume limit. In a finite volume, the phase transition blurs, and one has to decide how to define a finite-volume critical mass $m^2(g, L)$ that limits towards $m^2_{c,\infty}(g)$ in the infinite-volume limit,

$$m^2(g, L) \xrightarrow{L\to\infty} m^2_{c,\infty}(g). \tag{4.1}$$

We use the Binder Cumulant,

$$B = 1 - \frac{N}{3}\frac{\langle Tr[M^4]\rangle}{\langle Tr[M^2]\rangle^2}, \tag{4.2}$$

to do this as it changes value across the phase transition. We define the finite-volume critical mass $m^2_c(g, L)$ as the bare mass for which a desired transition value of the Binder Cumulant is observed. Doing this across our full range of ensembles allows us to test for infrared-finiteness but also to estimate the infinite-volume critical mass.

4.1 A Phase Transition

A phase transition of a field theory is a change in the symmetry properties of the vacuum state under the change of a physical parameter (or parameters) of the system. For example, as the temperature of a liquid is reduced there is a phase transition to a solid, which has different symmetry properties. The phases present across the full range of these parameters define the *phase diagram* of the system. A point or set of points at which the theory transitions from one phase to another is the *critical point*, critical line, or critical surface, depending on the dimensionality of the phase diagram.

In the case of our scalar lattice action (eq. (3.31)), there is one relevant parameter: the bare mass. We consider decomposing $\phi(x)$ as $\phi(x) = M + \chi(x)$, where $\chi(x)$ represents the local fluctuations of the theory, and M is the *magnetisation* defined by

$$M = \left(\frac{a}{L}\right)^3 \sum_{x\in\Lambda} \phi(x). \tag{4.3}$$

To understand the phases in our theory, it is useful to consider the effective field theory (EFT), where $\chi(x)$ contributions are dropped from the action. This may seem like a radical step to take, however, since we are interested in the regime near the critical point where the correlation length diverges, it will serve as a useful toy model. In this EFT, the effective action is simply

$$S_{\text{eff}} = \left(\frac{L}{a}\right)^3 \frac{N}{g}\left[m^2 \text{Tr}[M^2] + \text{Tr}[M^4]\right]. \tag{4.4}$$

There are at least two phases of our theory, depending on the value of m^2. To see this, recall that are fields M are in the adjoint of $SU(N)$, meaning that M has the properties of being Hermitian[1] and traceless. Denoting the eigenvalues of M by ρ_i, we can diagonalize M,

$$S_{\text{eff}} = \left(\frac{L}{a}\right)^3 \frac{N}{g} \sum_i \left[m^2\rho_i^2 + \rho_i^4\right]. \tag{4.5}$$

Since M is Hermitian, the eigenvalues must be real-valued, implying that ρ^2 and ρ^4 cannot be negative. Therefore we can identify two phases:

- $m^2 \geq 0$: There is a clear minimum point in this limit where $\rho_i = 0 \ \forall \ i$. The fields ϕ are therefore expected to fluctuate around zero. Given that this vacuum is unaffected by the action of group symmetries, this is a *symmetric phase*.
- $m^2 < 0$: A non-zero value for the absolute value of each ρ_i minimizes the effective action. Since M is traceless $\sum_i \rho_i = 0$. The magnetization must be in one of several degenerate minima, so this phase is a *broken phase*. For example with $N = 2$ the vacuum can be $\rho = (K, -K)$ or $\rho = (-K, K)$, where K minimizes $m^2K^2 + K^4$.

Classically, the change between these two phases occurs at $m^2 = 0$, however, when radiative corrections are accounted for, this point is shifted. If the theory is infrared finite, these shifts are finite in magnitude and the infinite-volume critical point is well defined. It is worth noting that there are no true phase transitions in a finite volume, as the transition is smoothed-out over a range of values rather than following step-function behavior. If the transition is present in the infinite-volume theory however, we could still recover the two phases in the $m^2 \to +\infty$ and $m^2 \to -\infty$ limits in a finite volume. The center of the finite-volume transition is shifted by finite-volume effects however, so it will be at $m^2(g, L) \neq m^2_{c,\infty}(g)$. Holographic cosmology models have generalized conformal structures that require the non-perturbative mass to be zero. In a finite-volume lattice QFT, this corresponds to the point $m^2 = m_c^2(g, L)$. This is the non-perturbative critical point where the quantum corrections to the mass cancel exactly with the bare mass.

4.2 A Conjecture of Infrared Finiteness

To understand why the action (eq. (3.31)) is described as super-renormalizable, we redefine the coupling constants $\tilde{m}^2 = (N/g)m^2$ and $\tilde{\lambda} = N/g$, and place the theory on an infinite lattice Λ with spacing a,

$$S = \text{Tr}\left(\int_{x\in\Lambda(a)} d^3x \left[(\partial_x\phi)^2 + \tilde{m}^2\phi^2 + \tilde{\lambda}\phi^4\right]\right). \tag{4.6}$$

[1] using the common convention in physics, as opposed to an anti-Hermitian convention

By sending $a \to \xi a$, for $0 < \xi < 1$ we flow to the UV of the theory,

$$S_\xi = \mathrm{Tr}\left(\int_{x\in\Lambda(\xi a)} d^3x \left[(\partial_x\phi)^2 + \tilde{m}^2\phi^2 + \tilde{\lambda}\phi^4\right]\right). \tag{4.7}$$

By making the substitution $x' = x/\xi$ the action becomes

$$S_\xi = \mathrm{Tr}\left(\int_{x'\in\Lambda(a)} \xi^3 d^3x' \left[\frac{1}{\xi^2}(\partial_{x'}\phi)^2 + \tilde{m}^2\phi^2 + \tilde{\lambda}\phi^4\right]\right). \tag{4.8}$$

To keep the kinetic term unchanged, we rescale the field variables, $\phi'(\xi) = \xi^{-1/2}\phi$, giving an integrand of $\mathrm{Tr}\left[(\partial_{x'}\phi')^2 + \xi^2\tilde{m}^2\phi'^2 + \xi\tilde{\lambda}\phi'^4\right]$. We can redefine the constants $\tilde{m}^2 \to \tilde{m}^2(\xi) = \tilde{m}^2\xi^2$ and $\tilde{\lambda} \to \tilde{\lambda}(\xi) = \tilde{\lambda}\xi$ to recover the initial form of the action. As we flow to the ultraviolet, our coupling constants naively approach a stable value, namely 0. This argument is entirely classical and could have been seen from power counting ($[\tilde{m}^2], [\tilde{\lambda}] > 0$). Instead of considering the $a \to 0$ limit of an infinite lattice, we can instead take the $L \to \infty$ limit of a lattice with $a = 0$. In this case, the exact opposite scaling of the coupling constants is found, warning us that the coupling constants, and by extension Feynman diagrams, are classically divergent in the infrared.

In perturbation theory, there is an IR-finite contribution to the critical mass at one-loop which is linear in the coupling constant g, a logarithmically IR-divergent contribution at two-loop and a polynomial IR-divergence at n-loops ($n > 2$). Explicitly, the perturbative result has the form

$$m^2_{\mathrm{PT}}(\Lambda_{\mathrm{IR}}) = b_1 g\mu + g^2\left[b_2 \log\left(\frac{\Lambda_{\mathrm{IR}}}{\mu}\right) + b_2' + \sum_{i=3}^{\infty} b_i \left(\frac{g}{\Lambda_{\mathrm{IR}}}\right)^{i-2}\right], \tag{4.9}$$

where μ is an arbitrary reference scale, Λ_{IR} is an infrared cut-off and b_i are the coefficients of terms at i-loop order. The constant b_2' accounts for ambiguity in the choice of μ. When we regularize with a lattice, the reference scale $\mu = 1/a$ while the infrared cut-off $\Lambda_{\mathrm{IR}} \propto 1/L$. However, a pair of papers published simultaneously in 1981 [59, 60] argued that these divergences are an artifact of perturbation theory. Non-perturbatively, these theories are IR-finite with the momentum cut-off in the log-term replaced by $g/(4\pi N)$. In this way, the coupling constant takes the role of an IR regulator. The polynomially divergent terms are completely absent non-perturbatively. Indeed the contribution from all higher loop orders combine to produce the $\log(g)$ term. In addition to this term, the authors of [60] predict a residual contribution from higher orders to the b_2' term. The size of this contribution cannot in general be predicted from perturbation theory. Henceforth in this thesis, we refer to this mechanism as the JTAP-mechanism after the names of the four authors of these papers (Jackiw, Templeton, Applequist, and Pisarski).

4.3 Lattice Perturbation Theory

To perform perturbative calculations in finite volumes and with finite lattice spacings, it is necessary to derive relevant Feynman rules. These rules differ from their continuum counterparts in a few key ways. Firstly any integrals over loop momenta are instead replaced by sums over the unique values of momenta on the lattice,

$$\int dp_i \to \frac{2\pi}{L} \sum_{p_i=-\pi/a}^{\pi/a-2\pi/L}. \tag{4.10}$$

The sum is discretized in increments of $2\pi/L$ due to periodicity. The lattice propagator G of our pure-scalar theory is the inverse operator to the quadratic term $m^2\phi^2 + (\partial\phi)^2$. It can be shown [82] that it has the general form

$$\tilde{G}(p;a) = \frac{1}{m^2 + \hat{p}^2}, \tag{4.11}$$

in momentum space, where

$$\hat{p}_\mu = \frac{2}{a}\sin\left(\frac{ap_\mu}{2}\right) \tag{4.12}$$

and $\hat{p}^2 = \sum_\mu \hat{p}_\mu^2$. The index structure to the propagator derives from the Fierz identity

$$(T^a)^i_j (T^a)^k_l = \delta^i_l \delta^k_j - \frac{1}{N}\delta^i_j \delta^k_l, \tag{4.13}$$

where T^a are the generators of $SU(N)$. Overall the propagator is

$$\tilde{G}(p;a)^{ij}_{kl} = \frac{\delta^i_l\delta^k_j - \frac{1}{N}\delta^i_j\delta^k_l}{m^2 + \hat{p}^2}. \tag{4.14}$$

The leading $\delta^i_l\delta^k_j$ contributes to planar Feynman diagrams, while the piece proportional to $1/N$ contributes to non-planar diagrams. Correlators of the theory are therefore dominated by planar contributions in the large-N limit.

The correlator, however, receives quantum corrections, which shift the mass to the renormalized mass $m^2 \to m^2 - m_c^2(g, L)$,[2] giving an expression for the renormalized propagator of

$$\tilde{G}_R(p;a) \sim \frac{1}{(m^2 - m_c^2) + \hat{p}^2}. \tag{4.15}$$

[2]While this shift is finite for finite L, if the theory is infrared divergent then this shift will also be divergent as $L \to \infty$.

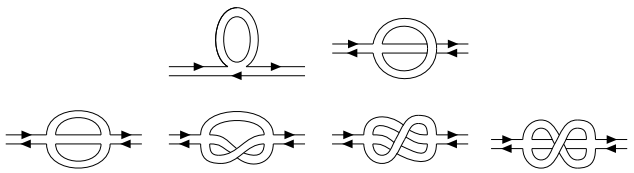

FIGURE 4.1: Feynman Diagrams contribution to the scalar self-energy at one- and two-loop. Figure taken froms [1]

During this project, we are interested in studying the theory near the phase transition (where $m^2 \approx m_c^2$), so we use the approximate form of the renormalized propagator,

$$\tilde{G}_R(p;a) \approx \frac{1}{\hat{p}^2}, \tag{4.16}$$

giving the expression for the one-loop integral proportional to

$$\sum_{p\in\tilde{\Lambda}} \frac{1}{\hat{p}^2}. \tag{4.17}$$

This sum is not strictly well-defined, however, we can calculate the infinite-volume approximation of it using

$$\int_{-\pi/a}^{\pi/a} \frac{d^3k}{(2\pi)^3} \frac{1}{\hat{k}^2} = \frac{Z_0}{a} \text{ with } Z_0 = 0.252731\ldots, \tag{4.18}$$

where $\hat{k}$ is defined analogously to $\hat{p}$. Similarly, the two-loop expression can be well approximated by the integral

$$D(p) = \int_{-\pi/a}^{\pi/a} \frac{d^3k}{(2\pi)^3} \frac{d^3q}{(2\pi)^3} \frac{1}{\hat{k}^2\,\hat{q}^2\,\hat{r}^2}, \tag{4.19}$$

where $r = p - k - q$, and hatted momenta are defined as before. This infinite-volume approximation may seem strange when we are interested in the infrared (large-volume) behavior. However, this calculation teaches us the dependence of the two-loop contribution on the external momentum p, which is easier to calculate in this approximation. When we use a finite-volume lattice regulator, this external momentum is replaced by the lowest non-zero momentum, $2\pi/L$. Unfortunately this integral diverges in the $p \to 0$ limit[3] [1],

$$D(p) \xrightarrow{p\to 0} -\frac{1}{(4\pi)^2}\log(|ap|), \tag{4.20}$$

confirming that the theory is badly behaved in the infrared in perturbation theory. The diagrams in fig. 4.1 yield the N-dependent constants of proportionality for both of these

[3]The calculation in 4D has been done in [83].

contributions. The overall perturbative prediction of the critical mass (using one- and two-loop contributions) is

$$m_c^2(g) = -g\frac{Z_0}{a}\left(2 - \frac{3}{N^2}\right) + g^2 D(\Lambda)\mathcal{N}(N), \tag{4.21}$$

where

$$\mathcal{N}(N) = 1 - \frac{6}{N^2} + \frac{18}{N^4} \tag{4.22}$$

and Λ depends on whether we are considering an IR-finite or IR-infinite anzatz (see also [84]). For the IR-infinite anzatz, the argument of D is the smallest momenta on the lattice, $2\pi/L$, while for the IR-finite anzatz we follow [59, 60] and use $\Lambda = g/(4\pi N)$.

4.4 The Binder Cumulant

To accurately determine the phase transition, we require a quantity to measure which changes across the transition. In [1] and the subsequent work presented in this thesis (as well as other projects in the LatCos collaboration), the Binder Cumulant (defined in equation (4.2)

has been used to serve this role. The Binder Cumulant [85, 86] is an effective metric for differentiating between symmetric and broken phases, and has the advantage that many effects on M due to the process of renormalization cancel between the top and bottom of the fraction. We can use the EFT to make predictions of the Binder Cumulant in the $m^2 \to \pm\infty$ limits. Deep into the broken phase, the ratio σ_{ρ_i}/ρ_i becomes very small, where ρ_i are the eigenvalues of M and σ_{ρ_i} is their standard deviation in an ensemble. This is because the potential well in the $m^2 \to -\infty$ limit becomes infinitely deep. Therefore we can take $\rho_i = \pm K(m^2)$ where $K(m^2)$ is a mass-dependent constant, giving

$$\langle Tr[M^4]\rangle = \sum_i K(m)^4 = NK(m)^4, \tag{4.23}$$

$$\langle Tr[M^2]\rangle^2 = (\sum_i K(m)^2)^2 = N^2K(m)^4, \tag{4.24}$$

and therefore a Binder Cumulant value of 2/3. The value of the Binder Cumulant in the $m^2 \to +\infty$ limit is less trivial and was calculated for $N = 2$ and $N = 4$ in [1]. When we calculate the Binder Cumulant for a set of ensembles created at fixed values of N, g, and L (see fig. 4.2), we observe a smooth monotonic transition between the two values predicted by these EFT calculations.

The Binder Cumulant, therefore, offers us a powerful tool to identify the phase transition (and therefore the critical mass) at each N, g, and L. These critical masses can then be used to test the infrared-finiteness of our theory. If our theory is infrared-finite (and

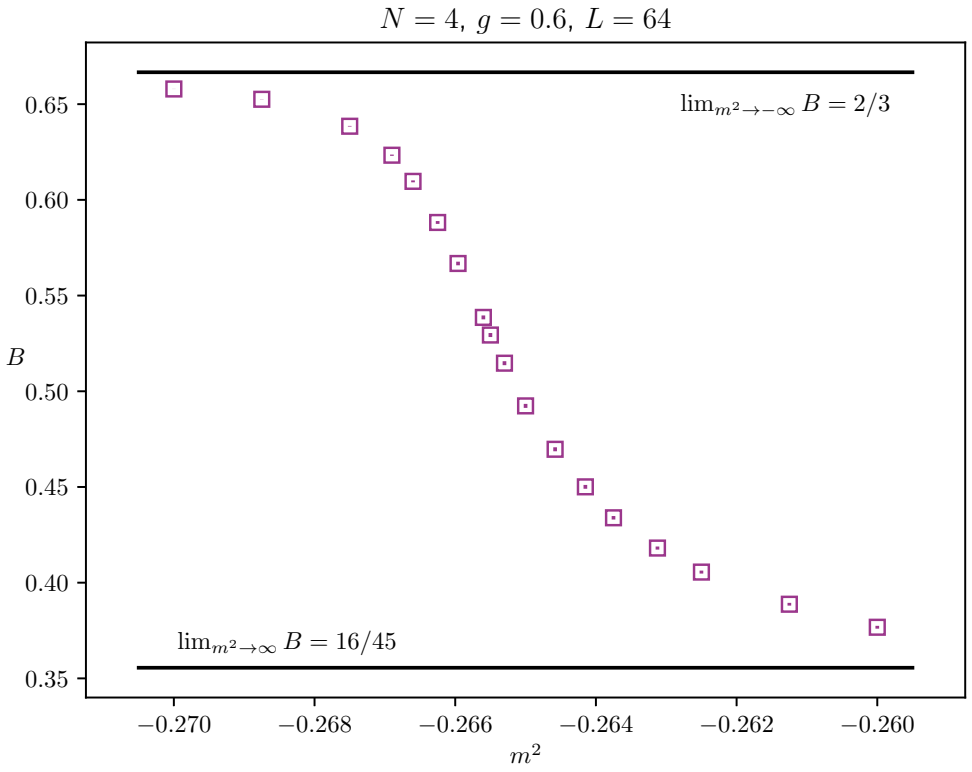

FIGURE 4.2: Plot showing the variation of the Binder Cumulant with bare mass m^2 for $N = 4$, $g = 0.6$, and $L = 64$. Shown at the center of the triangles are the measured values of the Binder Cumulant from lattice simulations with error bars. Shown also are the asymptotic limits calculated in the EFT [1].

therefore can be used to test holographic cosmology), this determination of the critical mass will be essential for simulating the theory at the non-perturbative critical mass point. One way the Binder Cumulant could be used to identify the critical mass point would be to find the point of greatest absolute gradient ($|\partial B/\partial m^2|_{\max}$). In this project, we choose another approach. We use multi-histogram reweighting procedure to calculate the Binder Cumulant continuously with mass. We then choose a fixed value of $\bar{B}$, which lies in the steep region of the transition. The mass at this value of the Binder Cumulant is taken to be the finite-volume critical mass. Note that the choice of $\bar{B}$ is arbitrary, and therefore will contribute to the systematic error in our results.

4.5 Data Generation

Ensembles of 100,000 configurations for a range of values of g, L, and N have been generated using both hybrid-Monte-Carlo [71] and heatbath overrelaxation algorithms [72–75] implemented in the GRID library [87, 88]. Simulations were performed using the Cambridge Service for Data-Driven Discovery (CSD3) with Intel Cascade Lake CPU

N	$\bar{B}$	g
2	0.51, 0.52, 0.53, 0.54, 0.55, 0.56, 0.57, 0.58, 0.59	0.1, 0.2, 0.3, 0.5, 0.6
3	0.40, 0.41, 0.42, 0.43, 0.44, 0.45, 0.46, 0.47, 0.48	0.1, 0.2, 0.3, 0.5, 0.6
4	0.42, 0.43, 0.44, 0.45, 0.46, 0.47	0.1, 0.2, 0.3, 0.5, 0.6
5	0.38, 0.4, 0.42, 0.44, 0.46, 0.48, 0.5, 0.52, 0.54	0.1, 0.2, 0.3, 0.4, 0.5, 0.6

TABLE 4.1: The values of N, g for which ensembles were generated. For each, lattices with size L in each of $\{8, 16, 32, 48, 64, 96, 128\}$ were used to generate ensembles. The values of $\bar{B}$ shown are those that were used to find critical mass estimates.

nodes. From each configuration of these ensembles, three quantities were used:

$$\begin{aligned} O_1 &= \sum_x \mathrm{Tr}[\phi^2(x)], \\ O_2 &= \mathrm{Tr}[M^2], \\ O_3 &= \mathrm{Tr}[M^4], \end{aligned} \tag{4.25}$$

where M is the magnetization. The quantities O_2 and O_3 were recorded in order to calculate the Binder Cumulant, while the quantity O_1 was recorded to perform a reweighting. The specific values of g, L, and N used to generate ensembles are in table 4.1. For each g, L, and N, ensembles were generated for bare masses across the phase transition region (as seen in the curve of the Binder Cumulant). Masses were spaced at a high enough density that a continuous Binder Cumulant curve could be produced across the whole range using reweighting without any gaps. For each ensemble, the autocorrelation times were calculated for O_1, O_2 and O_3, and the data were grouped into bins of $\max(50, 4\tau_{\mathrm{int}})$, where τ_{int} is the largest integrated autocorrelation time (see eq. (3.47)) of the three operators. These autocorrelation times were calculated using the methodology of [69].

4.6 Masses at Binder Cumulant Crossing Points

To get from raw binned data to estimates of the critical mass $\bar{m}^2$ (satisfying $B(\bar{m}^2) = \bar{B}$), a multi-step analysis pipeline is required. Firstly, a multi-histogram reweighting procedure [78] is needed to extend our Binder Cumulant data from a discrete set of mass points for each (N, g, L) to a continuous curve. This is straightforward to implement (as described in section 3.9), however, a bootstrap is required to get errors on these reweighting results.

Let S be a function that randomly samples (with replacement) from a set of binned configuration data. By applying S to our original data we produce samples s_i, where $i \in \{1, 2, ..., N_{\mathrm{boot}}\}$ where N_{boot} is the number of bootstrap samples. In calculating Binder Cumulant crossing points, $N_{\mathrm{boot}} = 500$ was used, while for the finite-size-scaling study 200 bootstrap samples were used. For each sample s_i, we can calculate the functions $B(m^2; s_i, m_0^2)$ using a multi-histogram reweighting [78] (described in section 3.9) for each

mass m_0^2 we reweight from. By finding the mean and standard deviation of the bootstrap samples, we get the functions for the mean and standard deviation ($\mu_B(m^2; m_0^2)$ and $\sigma_B(m^2; m_0^2)$ respectively) according to

$$\mu_B(m^2; m_0^2) = \frac{1}{N_{\text{boot}}} \sum_i B(m^2; s_i, m_0^2), \tag{4.26}$$

$$\sigma_B(m^2; m_0^2) = \sqrt{\frac{\sum_i (B(m^2; s_i, m_0^2) - \mu_B(m^2; m_0^2))^2}{N_{\text{boot}}}}. \tag{4.27}$$

To get a central estimate for $\bar{m}^2$, we combine these functions using a weighting average (see eq. (3.69)) to get an overall estimate $B(m^2)$. The weighting procedure used in [1] was

$$w(m_0^2, m^2) = (1 - \Theta(z(m^2, m_0^2) - z_{max}))/\sigma_B^2(m, m_0^2), \tag{4.28}$$

where z is defined in equation (3.67) and $z_{\text{max}} = 1$. A weighting scheme using just the $1/\sigma_B^2(m, m')$ is called *inverse-variance weighting* and gives the lowest overall error in the results (assuming normality).

The step function prevents inaccuracies due to reweighting too deep into the tails of the $\text{Tr}[\phi^2]$ histograms [79]. However, it also leads to subsequent step changes in the Binder Cumulant. This is an issue since we are interested in the mass at a specific value of the Binder Cumulant. The presence of a step in the vicinity of a constant value of $\bar{B}$ causes many bootstrap samples to get "caught" by the step. This can be seen in the top-left of figure 4.4, where the bootstrap samples are shown by histograms; in the presence of the steps, there are spikes in the histograms. This affects two things: firstly, it leads to an inaccurately low value of the standard deviation of the distribution (seen in the horizontal black bars), and secondly, it affects the normality of the distributions. The normality of the distributions is tested using the Shapiro-Wilkes test [89], with the p-values resulting from this test shown in grey below the histograms. As expected where the steps are present in the curve the p-values are smaller than 0.05 and thus the normality has been significantly impacted by the steps. To mitigate this, the weighted average performed in this project has been designed to smooth out the Binder Cumulant curve. We use the weighting

$$w(m', m) = f(z(m, m'), s, w)/\sigma_B^2(m, m'), \tag{4.29}$$

where f is a C_1-function in z, with a transition centered on s and width w, given by

$$f(z,s,w) = \begin{cases} 1, & z < s - w/2, \\ 1 - (2/w^2)(z - (s - w/2))^2, & s - w/2 \le z < s, \\ (2/w^2)(z - (s - w/2))^2, & s \le z < s + w/2, \\ 0, & z \ge s + w/2. \end{cases} \tag{4.30}$$

As the width w is increased, we indeed observe (see figure 4.4) that the Binder Cumulant curve smooths out, with the histograms of the crossing points becoming significantly more normal. A width of 0.3 is sufficient to produce histograms that pass the Shapiro-Wilks test ($p > 0.05$) in this case. The color of the histograms reveals the mass points that contributed to the reweighting (at that point) with non-zero weight. As the width is increased, the range for which a given mass point is included increases. This is most clearly observed in the upper-left histogram, where the mass represented by the integer 0 (out of the scope of the plot in the positive mass direction) only contributes to the reweighting on the right-hand side of the step. As the width is increased, this mass point contributes to more bootstrap samples in the outer bootstrap, smoothing out the spike.

A very subtle but important feature of the analysis done here is the use of `float128` floating point arithmetic in both the reweighting and root finding (as opposed to using the standard `float64`). This is beneficial because the exponential factor that appears in the reweighting procedure (see equation (3.65)) can be very large or very small. Therefore, it can cross over the maximum and minimum float sizes, causing a `NaN` value to appear and the reweighting to fail (even if the criteria of histogram overlap is satisfied). Using a `float128` float typically increases the maximum absolute value of the exponents available by a factor of 16 and was observed to negate this issue. In the `numpy` module `float128` numbers are handled naturally so this change was easy to implement using `X.astype(numpy.float128)`, where `X` is the data we are changing. For the $N = 2$ data, this difference has no real effect, while for the $N = 4$ data the reweighting is smoothed significantly, with the size of the error bars on the critical masses reduced. This is because the exponential reweighting factor is proportional to N (see eq. (3.66)). For $N = 4$, figure 4.3 serves as an example of the dramatic change to the data this causes. This causes the $N = 4$ data to be far more precise and therefore harder to fit.

Now that the reweighting estimates have been combined to give a single function of $B(m^2)$, a root-finding algorithm can be used to find the value $\bar{m}^2$ that returns $B = \bar{B}$. To get an error on the mass estimate, we do a second layer of bootstrap resampling. First, we take new bootstrap samples of the original data s'_i. Each of these samples is then resampled again by the random function S returning inner samples s'_{ij} where $j \in \{1, 2, ..., N_{\text{boot}}\}$. We then calculate mean and standard deviation functions $\mu_B^{(i)}(m^2; m_0^2)$

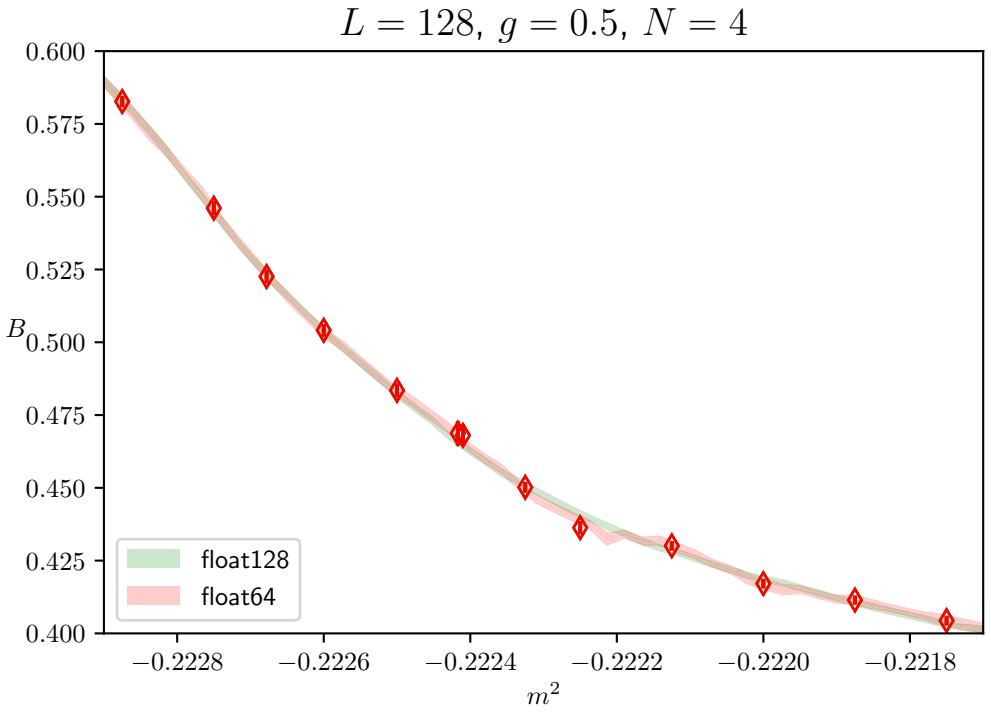

FIGURE 4.3: The data points show Binder Cumulant for $N = 4$, $L = 128$, $g = 0.5$ ensembles. The curve interpolating between them has been calculated using reweighting [78], with float64 arithmetic and with float128 arithmetic. The width parameter is set to 0.

and $\sigma_B^{(i)}(m^2; m_0^2)$ according to

$$\mu_B^{(i)}(m^2; m_0^2) = \frac{1}{N_{\text{boot}}} \sum_j B(m^2; s'_{ij}, m_0^2), \tag{4.31}$$

$$\sigma_B^{(i)}(m^2; m_0^2) = \sqrt{\frac{\sum_j (B(m^2; s'_{ij}, m_0^2) - \mu_B^{(i)}(m^2; m_0^2))^2}{N_{\text{boot}}}}. \tag{4.32}$$

By again combining these functions into overall $B^{(i)}(m^2)$ functions via a weighted average procedure and then performing a root-finding algorithm, we get estimates $(m^2)^{(i)}$ of the mass that returns $B = \bar{B}$. Finally, we can get an error on the mass by

$$\sigma_{m^2} = \sqrt{\frac{(m^2)^{(i)} - \bar{m}^2}{N_{\text{boot}}}}. \tag{4.33}$$

This procedure is called a *double bootstrap*. Since the Binder Cumulant is monotonic, a root-finding algorithm can be used to iteratively choose new m values to find the value of m that gives $B = \bar{B}$.

Before testing the infrared-finiteness of our theory using the Binder Cumulant of our lattice simulation data, it will first be necessary to understand the scaling properties of B that result from the finite size of the lattice.

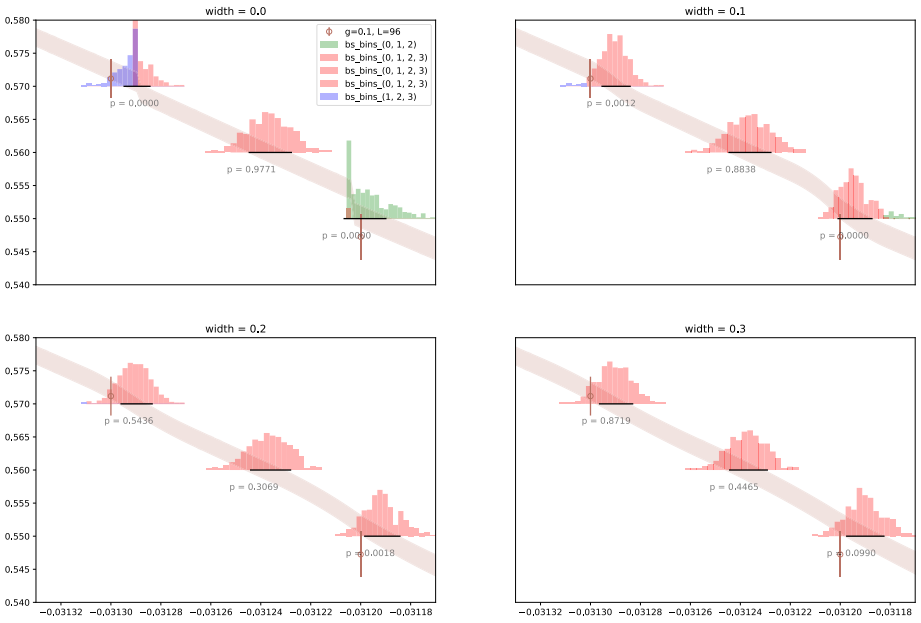

FIGURE 4.4: The Binder Cumulant curve with the double-bootstrap implemented for $L = 96$, $g = 0.1$, $N = 2$ data. The error bars due to the inner bootstrap are represented by the vertical width of the curve while the outer bootstrap is represented by the histograms and horizontal black bars. Below the histograms are p-values from the Shapiro-Wilks test. The histogram bars are colored by the set of masses included in the reweighting. Shown in brown are the raw data points (further points are present outside of the range of the plot).

4.7 Critical Scaling

In the large-volume limit, the finite-volume critical mass will be related to the infinite-volume critical mass and the lattice size L via critical exponents. In our study using the Binder Cumulant to find the critical mass, it will also depend on the choice of $\bar{B}$. The critical exponents represent important physical properties of the system, and two microscopically different theories with the same dimensionality and values of their critical exponents will have the same critical behavior and are said to be in the same *universality class*. An example critical exponent is the parameter ν, which appears in the expression of the correlation length ξ of a system in the approach to the critical point,

$$\xi \propto \left| \frac{m^2 - m_c^2(g, L)}{m_c^2(g, L)} \right|^{-\nu}. \tag{4.34}$$

This result, known as a *scaling-law*, is exact in the large L limit. Under the assumption of IR-finiteness, the Binder Cumulant has a leading scaling behavior

$$B = h\left(\frac{m^2 - m^2_{c,\infty}(g)}{g^2}(gL)^{1/\nu}\right), \tag{4.35}$$

with the critical exponent ν appearing again, $m^2_{c,\infty}(g)$ being the infinite-volume critical mass at coupling g and the function h being a smooth monotonic function depending on N. If the theory is IR-divergent, then the term $m^2_{c,\infty}(g)$ is not well-defined.

These scaling relations are, however, only approximate and suffer from *corrections-to-scaling* that are sub-leading at large lattice volumes. In general, the scaling of a quantity in a finite-size scaling study will be multiplied by a correction-to-scaling function $X(x)$. In our model $x = gL$, however in other theories it may be the unitless volume L/a. This correction-to-scaling function receives contributions from a series of correction-to-scaling coefficients, ω_i, and takes the general form [90]:

$$1 + \sum_i a_i x^{-\omega_i} + \sum_{ij} c_{ij} x^{-(\omega_i+\omega_j)} + ..., \tag{4.36}$$

where the constants a_i, c_{ij}, and those in higher order terms are dependent on the quantity being studied while the coefficients ω_i are universal across different quantities of a given theory. When we later fit our critical mass point data for various g and L, the more correction-to-scaling coefficients included, the higher the statistical errors on the fitting parameters become.

On top of the ω correction-to-scaling coefficients, there is an additional correction term with exponent ϵ (approximately 2 at a 1% level [90]) caused by the breaking of isotropy in a lattice. Other critical exponents can be found in many resources, for example, [91].

4.8 Finite-Size-Scaling Study

Estimates of critical masses and their standard deviations for different values of g, L, N, and $\bar{B}$ are given in table 4.1. To get the expected scaling behavior of the critical mass, it is useful to express m^2 in terms of the unitless ratio

$$\zeta = \frac{m^2 - m^2_{c,\infty}(g)}{g^2}, \tag{4.37}$$

where $m^2_{c,\infty}(g)$ is the infinite-volume critical mass point. The Binder Cumulant can be expressed as a continuous and monotonic function of $f(\zeta)$, where f depends on g, L, and N. We name the inverse of this function $\tilde{f}$ so that $\zeta = \tilde{f}(B)$. In the appendix A the

following expression $\tilde{f}(B)$ is derived:

$$\tilde{f}(B) = A(B)(gL)^{-\frac{1}{\nu}} + \sum_{k=1}^{2} D_1(g,B)(gL)^{-(1/\nu+\omega)} + D_2(g,B)(gL)^{-(1/\nu+2\omega)} \tag{4.38}$$
$$+ \sum_{j=1}^{2} E_j(g,B)(gL)^{-(1/\nu+\epsilon_j)} + O\left((gL)^{-2\epsilon}\right) + O\left((gL)^{-3\omega}\right).$$

We can use the definition of ζ to get an equation for the value of m^2,

$$m^2(g,L,\bar{B}) = m^2_{c,\infty}(g) + g^2\tilde{f}(g,\bar{B},L). \tag{4.39}$$

This expression sets the stage for testing competing anzätze against our data. There are different fit choices both when it comes to the critical mass term and when it comes to the scaling terms. One contribution that is unchanged in all of our fits is the one-loop contribution, as the JTAP-mechanism leaves this contribution unaffected from perturbation theory. Secondly, the one-loop contribution is large compared to other terms in the fit, so not including it is not an option.

At the two-loop order in perturbation theory, we have our first choice: to use an IR-finite or IR-infinite log term. For the IR-infinite anzatz we include the naive expression of the two-loop contribution, which has the form

$$\beta\frac{1}{(4\pi)^2}\mathcal{N}(N)g^2\log\left(\frac{1}{L}\right), \tag{4.40}$$

while the IR-finite anzatz uses the JTAP-mechanism replacing $1/L$ with $g/(4\pi N)$ [59, 60].

All fits also include a term αg^2, accounting for any residual scheme dependence in the UV-regulator and the constant in the log [4]. In the IR-finite scenario, such a term also appears due to non-perturbative effects in the JTAP-mechanism. In some fits we use a single value of alpha while in others we use a different value for each value of $\bar{B}$. There is no a priori justification of why we should use a different value of alpha for each $\bar{B}$ since the choice of $\bar{B}$ is relevant only to the scaling term of eq. (4.39). However, including more alpha terms may be an effective way to offer the fit more fitting power without shifting the values of key parameters such as the scaling coefficient ν.

There is, in general, an infinite tower of terms scaling with coefficients ω_i (eq. (4.36)) which can be included in the scaling term. While including more terms leads to better fit quality, it is also associated with higher statistical errors in the fit parameters. In our study, we found a single coefficient ω sufficient to give a strong fit quality at the level of statistical precision we are working with. Even here, we found that leaving the value of ω free leads to unstable results, so we fix its value to 0.8. We chose this value because it is

[4] Given this ambiguity in the log term, we use $1/L$ (instead of $2\pi/L$) in the argument of the log for the IR-infinite anzatz

a commonly used estimate of the leading correction-to-scaling coefficient in scalar $O(N)$-vector models [90]. The $O(3)$ model is equivalent to taking $N = 2$ in our model, while the $O(8)$ model is in the same universality class as taking $N = 3$ [92]. In $O(N)$ models the value of ω has not been observed to significantly vary with N. We are motivated to believe a similar behavior may be present in our models, so the value of ω for $N = 2$ and $N = 3$ (e.g. 0.8) is likely to be appropriate for choices of $N > 3$. This is however speculative, and further study on the value of ω in this class of models is needed. In addition to these ω-contributions, in some fits we also include a combined term, scaling like $\sim (gL)^{-\epsilon}$, for the breaking of isotropy and due to the magnetic background. Again we find that leaving ϵ free to vary leads to instabilities, so we fix $\epsilon = 2$. The fitting parameters were not found to be particularly sensitive to the specific values used for ω or ϵ. We do allow the leading scaling coefficient ν to vary.

The last kind of fit we refer to in this document is a fit with both $\log(g)$ and $\log(L)$ terms. These are used purely to investigate the relative significance of these two contributions. Putting together all of these fit possibilities, we get table 4.2, where the fits are

$$\bar{m}^2(g, \bar{B}, L) = gZ_0\left(2 - \frac{3}{N^2}\right) + g^2X(\bar{B}) + \frac{1}{(4\pi)^2}\mathcal{N}(N)g^2Y(g, L) + g^2Z(g, L, \bar{B}). \tag{4.41}$$

While the scaling parameters c and e can depend on g as well as $\bar{B}$, in practice, we find that good fit quality is obtained assuming these functions to be independent of g. We therefore do not include any dependence of these parameters on g in our fit anzätze. Throughout this project, we focus on fits that use data with two adjacent values of $\bar{B}$. For example for the $N = 2$ data we have $\bar{B} \in (0.52, 0.53)$, $\bar{B} \in (0.53, 0.54)$, ..., $\bar{B} \in (0.58, 0.59)$. Given we only use two values of $\bar{B}$, it is valid to express the functions a, c, and e as linear forms without losing any information (e.g. $a(\bar{B}) = a_1\bar{B} + a_2$). However, for the Bayesian study, discussed in section 4.11, we want to use the most natural parameterization since it affects the prior distribution, so we used $a(\bar{B}_1) = a_1$ and $a(\bar{B}_2) = a_2$.

4.9 Binder Cumulant Plots

Our finite-size-scaling study produces the necessary data to feed into our fitting anzätze to get out estimates of the fit parameters via a Bayesian model averaging procedure explained in section 4.12. If we work under the assumption that the IR-finite anzatz is correct then we can also predict the infinite-volume critical mass,

$$m^2_{c,\infty}(g) = gZ_0\left(2 - 3/N^2\right) + g^2\left(\alpha + \beta\log\left(\frac{g}{4\pi N}\right)\right). \tag{4.42}$$

Full parameter and critical mass estimates are presented in sections 4.13 and 4.14, respectively. For now, we will use these results to take a first look at the data by plotting the Binder Cumulant on a rescaled mass axis. To do this, we note that the formula

	$X(\bar{B})$	$Y(g,L)$	$Z(g,L,\bar{B})$
A_1	α	$\beta\log\left(\frac{g}{4\pi N}\right)$	$a(\bar{B})(gL)^{-1/\nu}$
A_2	α	$\beta\log\left(\frac{1}{L}\right)$	$a(\bar{B})(gL)^{-1/\nu}$
A_3	α	$\left[\beta_1\log\left(\frac{g}{4\pi N}\right)+\beta_2\log\left(\frac{1}{L}\right)\right]$	$a(\bar{B})(gL)^{-1/\nu}$
B_1	α	$\beta\log\left(\frac{g}{4\pi N}\right)$	$a(\bar{B})(gL)^{-1/\nu}+c(\bar{B})(gL)^{-(1/\nu+\omega)}$
B_2	α	$\beta\log\left(\frac{1}{L}\right)$	$a(\bar{B})(gL)^{-1/\nu}+c(\bar{B})(gL)^{-(1/\nu+\omega)}$
B_3	α	$\left[\beta_1\log\left(\frac{g}{4\pi N}\right)+\beta_2\log\left(\frac{1}{L}\right)\right]$	$a(\bar{B})(gL)^{-1/\nu}+c(\bar{B})(gL)^{-(1/\nu+\omega)}$
C_1	α	$\beta\log\left(\frac{g}{4\pi N}\right)$	$a(\bar{B})(gL)^{-1/\nu}+c(\bar{B})(gL)^{-(1/\nu+\omega)}+e(\bar{B})(gL)^{-(1/\nu+\epsilon)}$
C_2	α	$\beta\log\left(\frac{1}{L}\right)$	$a(\bar{B})(gL)^{-1/\nu}+c(\bar{B})(gL)^{-(1/\nu+\omega)}+e(\bar{B})(gL)^{-(1/\nu+\epsilon)}$
C_3	α	$\left[\beta_1\log\left(\frac{g}{4\pi N}\right)+\beta_2\log\left(\frac{1}{L}\right)\right]$	$a(\bar{B})(gL)^{-1/\nu}+c(\bar{B})(gL)^{-(1/\nu+\omega)}+e(\bar{B})(gL)^{-(1/\nu+\epsilon)}$
D_1	$\alpha(\bar{B})$	$\beta\log\left(\frac{g}{4\pi N}\right)$	$a(\bar{B})(gL)^{-1/\nu}$
D_2	$\alpha(\bar{B})$	$\beta\log\left(\frac{1}{L}\right)$	$a(\bar{B})(gL)^{-1/\nu}$
D_3	$\alpha(\bar{B})$	$\left[\beta_1\log\left(\frac{g}{4\pi N}\right)+\beta_2\log\left(\frac{1}{L}\right)\right]$	$a(\bar{B})(gL)^{-1/\nu}$

TABLE 4.2: A complete list of the different fit anzätze used in this project, along with their names in the left-hand column. The overall fit expression is equation (4.41), which defines X, Y, and Z. ω and ϵ are fixed at values 0.8 and 2 respectively.

for the finite volume critical mass separates into an infinite-volume contribution and a scaling piece,

$$\bar{m}^2(g,\bar{B},L) = m^2_{c,\infty}(g) + g^2(gL)^{-1/\nu}\left(a(\bar{B}) + c(\bar{B})(gL)^{-\omega} + e(\bar{B})(gL)^{-\epsilon}\right), \tag{4.43}$$

where $c(\bar{B}) = 0$ for models A and D and $e(\bar{B}) = 0$ for models A, B and D. This expression can be rearranged for $a(\bar{B})$, the leading scaling behavior of the Binder Cumulant,

$$a(\bar{B}) = \frac{\bar{m}^2(g,\bar{B},L) - m^2_{c,\infty}(g)}{g^2}(gL)^{1/\nu} - c(\bar{B})(gL)^{-\omega} - e(\bar{B})(gL)^{-\epsilon} := m_r^2(g,L,\bar{B}), \tag{4.44}$$

where we have defined the rescaled mass $m_r^2(g,\bar{B},L)$. As discussed in appendix (A) the function a is invertible,

$$\bar{B} = a^{-1}\left(m_r^2(g,\bar{B},L)\right). \tag{4.45}$$

We expect that, if we have captured all the scaling behavior of the theory, the plots of $\bar{B}$ against $m_r^2(g,\bar{B},L)$ will line up on top of each other. At first glance, we see that a simple anzatz with just leading scaling describes the $N = 2$ data well (model A_1, fig. 4.5), while for $N = 3$ a single correction-to-scaling term is required to fit the data well (model B_1, fig. 4.6) and for $N = 4$ and $N = 5$ two correction-to-scaling terms are needed (model C_1, figs. 4.7 and 4.8). To see plots of the Binder Cumulant for all N and all models, please

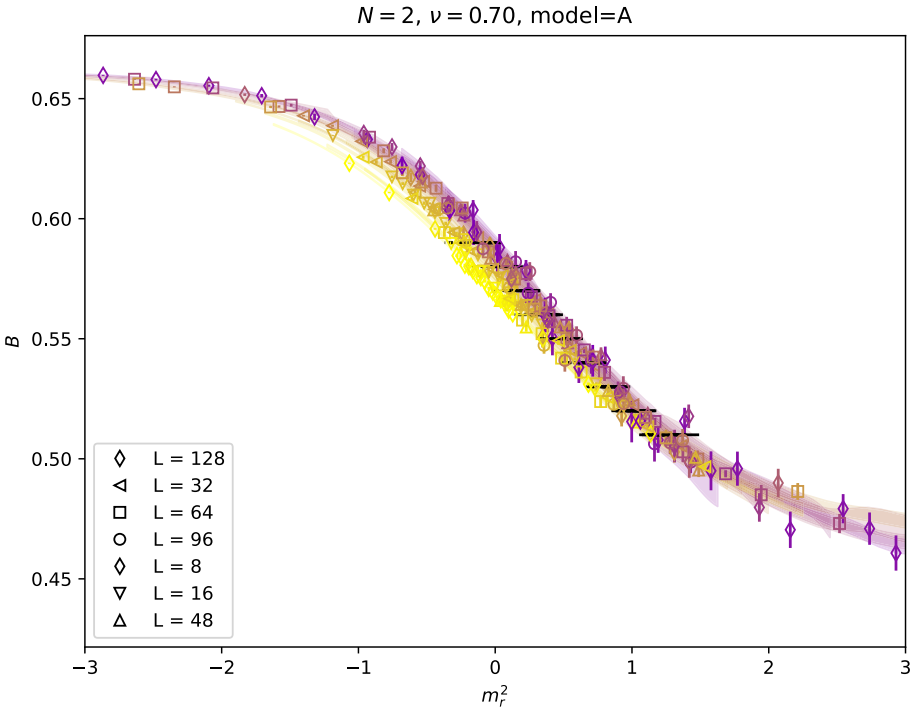

FIGURE 4.5: The Binder Cumulant for $N = 2$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model A_1. All data down to and including $gL_{\text{min}} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

refer to the appendix B. The fact that we can lie the Binder Cumulant curves on top of each other for a large range of gL implies that we have correctly captured the leading scaling of the Binder Cumulant. If there were large IR divergences, this would not be possible. These plots, therefore, provide soft evidence of the IR-finiteness of our model, however, to make stronger conclusions, it will be necessary to do more rigorous data analysis, which is presented in sections 4.10 and 4.11. The fact that more correction-to-scaling terms are required as N increases implies that these contributions become more significant with increasing N. This idea is discussed further in section 4.15.

4.10 Frequentist Analysis

In this section, we discuss the frequentist analysis done to compare the IR-finite and IR-infinite anzätze against each other. This analysis was done first with $N = 2$ and $N = 4$ data and presented in [1]. Since then, the analysis has been refined and applied to $N = 3$ and $N = 5$ data. In this thesis we present this more recent analysis.

In all fits two consecutive values of $\bar{B}$ were used. This is because a single value of $\bar{B}$ leads to redundancies, and fitting with more values of $\bar{B}$ was found not to be practical.

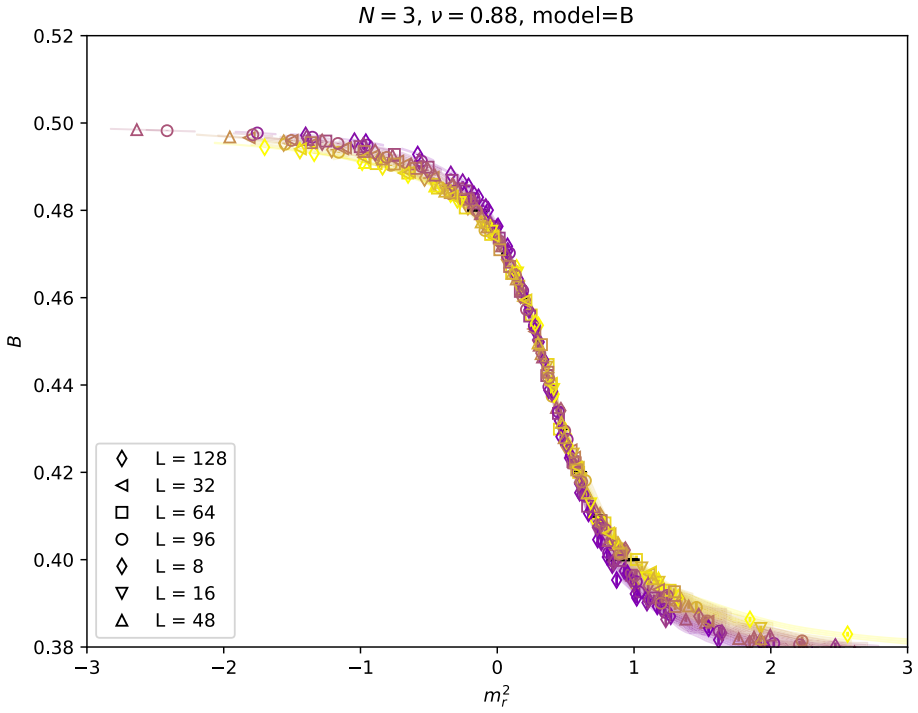

FIGURE 4.6: The Binder Cumulant for $N = 3$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model B_1. All data down to and including $gL_{\min} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

Whenever we refer to the choice of $\bar{B}$, we are actually referring to a fit with that value of $\bar{B}$ and the next value up, e.g. $\bar{B} = 0.52$ for $N = 2$ means data with $\bar{B} = 0.52$ and $\bar{B} = 0.53$ were used. Table 4.1 enumerates the different choices of $\bar{B}$ for each choice of N.

The smaller the value of x, the more the data will be affected by correction-to-scaling effects. We therefore cut out all data with $x < gL_{\min}$. This leaves the question of what a good choice of $gL_{\min}$ is. There is no a priori answer to this question. However, if either the IR-finite or IR-divergent anzatz can fit all data down to a given $gL_{\min}$ without the need to include corrections-to-scaling, then above this $gL_{\min}$ value the data must be sufficiently free from corrections-to-scaling that it can be fit by an anzatz without corrections-to-scaling (at the level of the noise in the data). With this in mind, whichever of the IR-finite or IR-divergent anzatz can fit down to the lowest $gL_{\min}$, are classed the superior anzatz. To this end, we present the results of this frequentist comparison in table 4.3.

Also provided in table 4.3 is the Akaike Information Criterion [93–95]. We use the method of [10] and treat each data point not included in the fit as an extra model

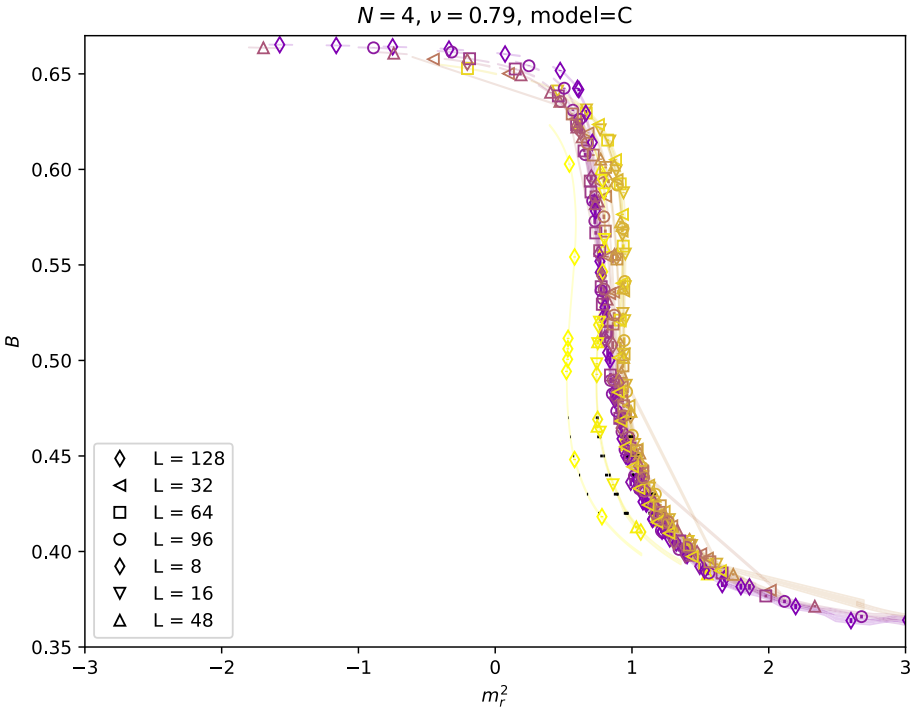

FIGURE 4.7: The Binder Cumulant for $N = 4$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model C_1. All data down to and including $gL_{\text{min}} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

parameter giving,

$$\text{AIC} \approx \chi^2 + 2N_{\text{cut}} + 2k, \tag{4.46}$$

where N_{cut} is the number of data points not included in the fit (due to being cut because $gL < gL_{\text{min}}$) and k is the number of fitting parameters.

The results in table 4.3 tell us many things. Firstly, when we compare the models without corrections-to-scaling, the IR-finite anzatz (model A_1) performs much better than in IR-infinite one (model A_2). Since this anzatz is expected to be correct in the large volume limit this result provides strong evidence in favor of IR-finiteness. Also, the table demonstrates that including corrections-to-scaling in the fit dramatically improves the fitting power as measured by the number of degrees of freedom. This result is especially noticeable for the IR-infinite anzatz, which potentially indicates a conspiracy of the various terms in the fit to mimic $\log(g)$-like behavior for the large gL values. However, it must be noted that the IR-infinite anzätze with corrections-to-scaling perform as well or better than the IR-finite anzätze. This indicates that the data at smaller gL may have $\log(L)$ like behavior. At first glance, this seems very damning for the conclusion of infrared finiteness. However, we argue that this results from $\log(L)$ behavior being

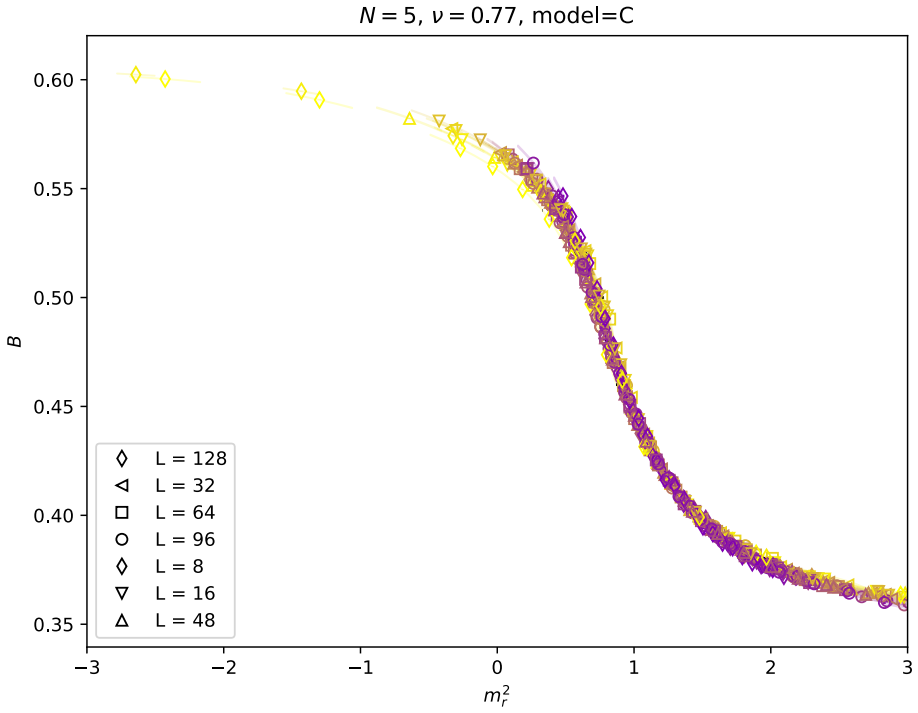

FIGURE 4.8: The Binder Cumulant for $N = 5$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model C_1. All data down to and including $gL_{\text{min}} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

present at small gL where finite-size effects are dominant. This idea is explored further in section 4.15.

4.11 Bayesian Evidence

So far, we have approached the problem of IR finiteness from a frequentist perspective. In this framework, competing models (anzatz) are fit to the data. The quality of the fit is determined by using a test statistic, in this case, the χ^2, to calculate an associated p-value to the fit. Under the null hypothesis that the fit is a true description of the underlying data, the p-value tells us the probability we would see a test statistic as large as we see (purely due to statistical noise). By definition, the p-value must be between 0 and 1. Under the assumption of the null hypothesis, it is sampled from a uniform distribution between 0 and 1. If the p-value is very small, it is highly unlikely the data would be as poorly described by the fit observed by chance. An arbitrary threshold α, on the p-value, called a significance level, must therefore be chosen. If a fit yields a p-value below α, we reject the null hypothesis and the model. This procedure is called a *significance test.*

N	model	model type	$gL_{\min}$	dof	AIC
2	A_1	$\log(g)$	12.8	31	116.0
2	A_2	$\log(L)$	48.0	3	134.8
2	B_1	$\log(g)$	4.0	51	100.5
2	B_2	$\log(L)$	1.6	61	95.0
2	C_1	$\log(g)$	2.4	55	98.6
2	C_2	$\log(L)$	0.8	61	94.0
2	D_1	$\log(g)$	12.8	30	109.0
2	D_2	$\log(L)$	24.0	16	117.6
3	A_1	$\log(g)$	16.0	25	121.4
3	A_2	$\log(L)$	48.0	3	138.2
3	B_1	$\log(g)$	2.4	57	96.2
3	B_2	$\log(L)$	1.6	61	81.7
3	C_1	$\log(g)$	0.8	61	92.2
3	C_2	$\log(L)$	0.8	61	75.7
3	D_1	$\log(g)$	8.0	40	115.1
3	D_2	$\log(L)$	16.0	24	126.8
4	A_1	$\log(g)$	32.0	9	133.6
4	A_2	$\log(L)$	–	–	–
4	B_1	$\log(g)$	16.0	23	126.6
4	B_2	$\log(L)$	3.2	55	96.9
4	C_1	$\log(g)$	8.0	37	111.7
4	C_2	$\log(L)$	0.8	61	92.3
4	D_1	$\log(g)$	25.6	14	134.3
4	D_2	$\log(L)$	28.8	12	134.4
5	A_1	$\log(g)$	32.0	13	161.2
5	A_2	$\log(L)$	–	–	–
5	B_1	$\log(g)$	9.6	47	119.8
5	B_2	$\log(L)$	8.0	49	129.3
5	C_1	$\log(g)$	3.2	67	111.0
5	C_2	$\log(L)$	3.2	67	115.6
5	D_1	$\log(g)$	24.0	22	152.4
5	D_2	$\log(L)$	32.0	12	161.4

TABLE 4.3: Smallest choice of $gL_{\min}$ that has an acceptable (> 0.05) p-value for each model and value of N.

It is worth mentioning that very high p-values (close to 1) can indicate overfitting, so one may also wish to reject the null hypothesis for fits with p-values above a threshold value. In this project, a lower threshold of 0.05 has been used. By calculating the p-values of fits of competing models, evidence in favor of one model over the other can be found. This is the case when one model is rejected under a significance test, but the other is not.

Given the prevalence of significance testing and frequentist statistics frameworks, it may be surprising that the procedure of model selection described above is infact convoluted, and may produce different results than a direct statistical test on which model is favored by the data. The proper way to directly answer such a question lies in Bayes' Theorem.

The probability of a model M with parameters α being true given observed data is

$$\begin{aligned} p(M|\text{data}) &= \int d\alpha\, p(M,\alpha|\text{data})\,, \\ &= \int d\alpha\, \frac{p(\text{data}|M,\alpha)\, p(M,\alpha)}{p(\text{data})} \\ &= \int d\alpha\, \frac{p(\text{data}|M,\alpha)\, p(M) p(\alpha|M)}{p(\text{data})} \\ &= \frac{p(M)}{p(\text{data})} \int d\alpha\, L(M,\alpha)\, p(\alpha|M), \end{aligned} \tag{4.47}$$

where $p(\alpha|M)$ is the *prior* distribution of the model parameters, which captures our belief about what the model parameters should be before looking at any data in this experiment. The factor $p(M)$ accounts for any prior belief of how likely the model is to be true, while $p(\text{data})$ is the probability of obtaining the observed data distribution. In the second line of eq. (4.47), Bayes' theorem is used, and in the final line the definition of the *likelihood function* is used ($L(M,\alpha) = p(\text{data}|M,\alpha)$). When comparing two different models using the same data, the factor $p(\text{data})$ is equal, meaning we can usually ignore it. In the context of model selection, we will usually choose the prior $p(M)$ to be equal for all models. Define data points (x_i, y_i) with $i \in \{1, 2, ..., k\}$ where x_i are values of the predictive (dependent) variables and y_i are the observed values of the (independent) variable. Under the assumption of Gaussian errors, the likelihood function is given by

$$L(M,\alpha) = \frac{1}{\sqrt{(2\pi)^k |\text{Cov}|}} \exp\left(-\frac{1}{2}\chi^2(\text{Cov}, x, y)\right), \tag{4.48}$$

where Cov is the covariance matrix of the output variable, x represents the set of all x_i and y the set of all y_i. The χ^2 statistic is given by

$$\chi^2(\text{Cov}, x, y) = \sum_{i=1}^{k}\sum_{j=1}^{k} (y_i - M(\alpha, x_i))^T \cdot \text{Cov}_{ij}^{-1} \cdot (y_j - M(\alpha, x_j)). \tag{4.49}$$

When comparing two models, we are interested in the ratio of the probabilities of the two models given the data,

$$K(\text{data}) = \frac{p(M_1|\text{data})}{p(M_2|\text{data})} = \frac{\int d\alpha_1\, L(M(\alpha_1))\, p(\alpha_1|M)}{\int d\alpha_2\, L(M(\alpha_2))\, p(\alpha_2|M)}, \tag{4.50}$$

where we have taken $p(M_1)/p(M_2) = 1$. We can interpret the values we get using the Jeffreys' scale [96]. A version of this scale is shown in table 4.5. Similarly to the frequentist case, this interpretation stage is ambiguous, however, we will see later that our results have such large values of K that the exact choice of scale will not affect the conclusions.

Parameter	Uniform Prior Range
α	[-0.4, 0.4]
a_1, a_2	[-30, 30]
β	[-15, 15]
ν	[0.001, 15]

TABLE 4.4: Prior ranges used in calculating the Bayesian evidence comparing the IR-finite model A_1 and the IR-infinite model A_2.

$K \leq -2$	Decisive Evidence for model 2 over model 1
$-2 > K \geq -1$	Strong Evidence for model 2 over model 1
$-1 < K < 1$	Insignificant Evidence
$1 \leq K < 2$	Strong Evidence for model 1 over model 2
$K \geq 2$	Decisive Evidence for model 1 over model 2

TABLE 4.5: A simple version of the Jeffreys' scale [96]

The Bayesian evidence has been calculated using the MULTINEST code [97–99] implemented in the Pymultinest framework [100]. Unlike frequentist methods which rely only on the maximum point of the likelihood function in parameter space, the results of Bayesian analysis depend on the choice of priors used. We choose a uniform prior which is as wide as possible without the MULTINEST integrator running into numerical issues due to under-sampling. The priors are shown in table 4.4.

Due to difficulties with the numerical integration, the Bayesian evidence was only calculated for model A_1 versus model A_2. Other models are too complex in parameter space for the integrator to give stable results. The $\bar{B}$ chosen to make this plot is the central fit value of $\bar{B}$. To calculate this, fits across all possible choices of $\bar{B}$ and $gL_{\min}$ were considered, and the fit with the smallest $gL_{\min}$ which has an acceptable p-value ($p > 0.05$) was chosen. The choice of $\bar{B}$ is simply the value that is associated with this lowest acceptable $gL_{\min}$. If there are multiple choices of $\bar{B}$ which have the same $gL_{\min}$, we choose the one with the larger p-value at this choice of $gL_{\min}$.

The results of this study are shown in combined plots alongside the frequentist results in figures 4.10 for $N = 2$ and $N = 4$ and 4.9 for $N = 3$ and $N = 5$. When sufficient data were included (small enough $gL_{\min}$), the Bayesian Evidence gave decisive evidence for the IR-finite anzatz under the Jeffreys' scale. This is also true in the region of $gL_{\min}$ for which the IR-finite anzätze have acceptable p-value and the IR-infinite anzätze don't. This provides strong evidence for IR-finiteness as these two distinct methods complement each other nicely. The first aim of this project, to provide evidence of infrared-finiteness in our model, is thus achieved. This allows us to turn our attention to extracting estimates for the parameters of our fits, and by extension, the infinite-volume critical mass.

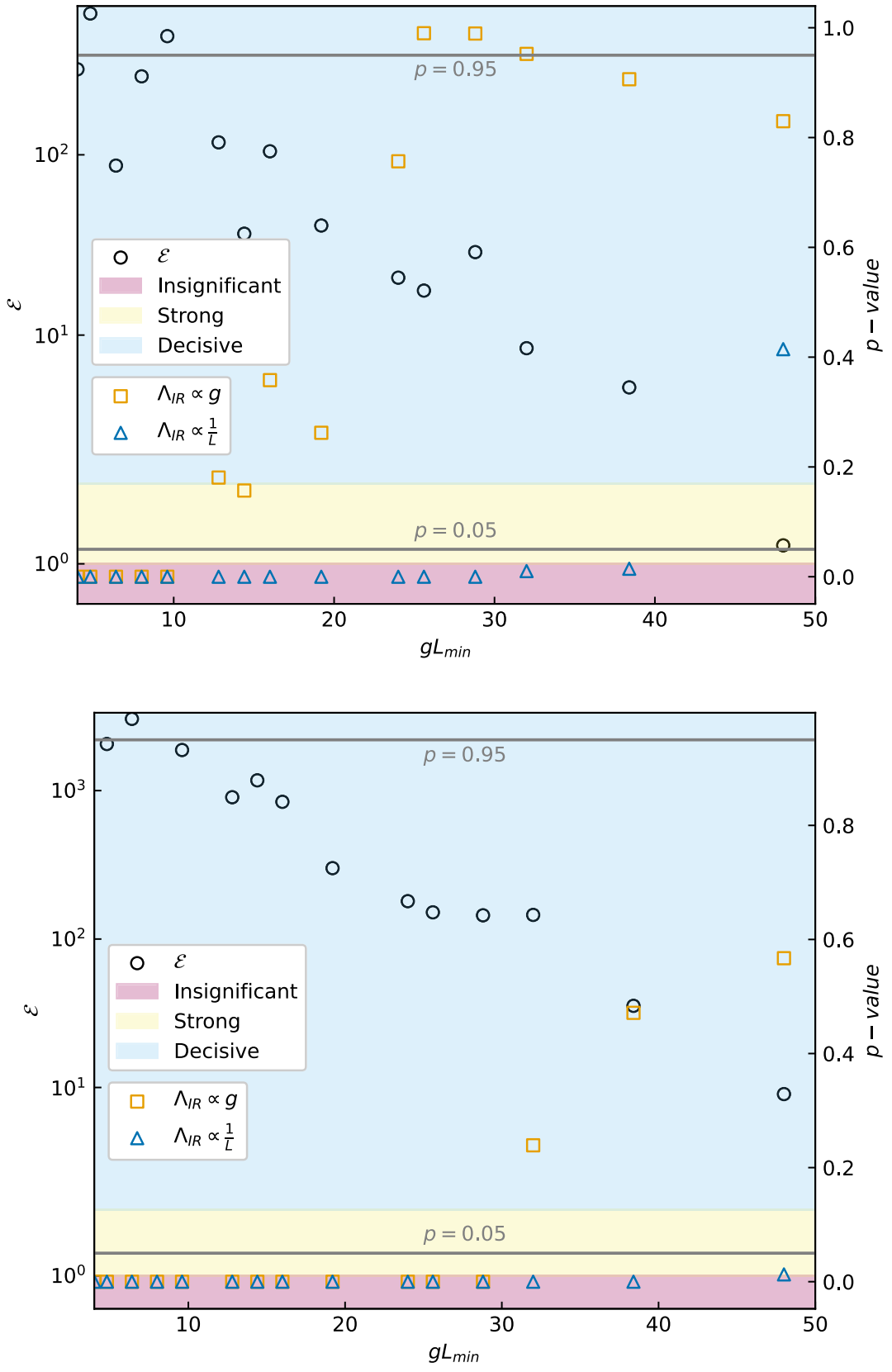

FIGURE 4.9: Combined frequentist and Bayesian analysis plot for $N = 2$ (top) and $N = 4$ (bottom) data. Models A_1 (orange squares) and A_2 (green triangles) were compared. On the x-axis is the $gL_{\min}$ cut while the y-axis shows the p-values for the frequentist fits and $\mathcal{E} = \log_{10}(\mathrm{K})$, where K is the Bayesian evidence for the Bayesian analysis. The graph regions are shaded according to the Jeffreys' scale interpretation of the Bayesian evidence, where the blue region at the top of the graphs is decisive evidence for the IR-finite anzatz over the IR-infinite one.

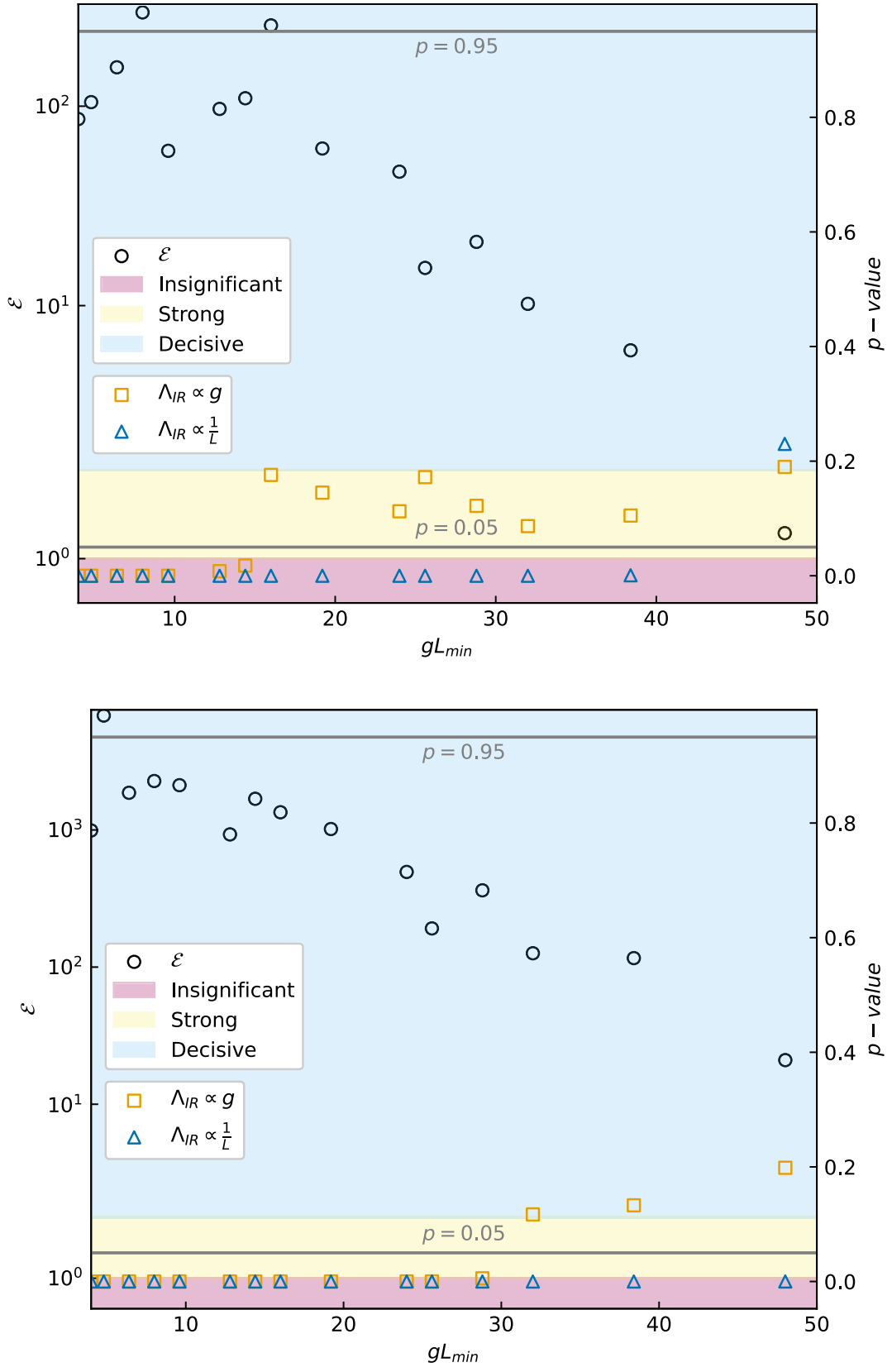

FIGURE 4.10: Combined frequentist and Bayesian analysis plot for $N = 3$ (top) and $N = 5$ (bottom) data. Models A_1 (orange squares) and A_2 (green triangles) were compared. On the x-axis is the $gL_{\min}$ cut while the y-axis shows the p-values for the frequentist fits and $\mathcal{E} = \log_{10}(\mathrm{K})$, where K is the Bayesian evidence for the Bayesian analysis. The graph regions are shaded according to the Jeffreys' scale interpretation of the Bayesian evidence, where the blue region at the top of the graphs is decisive evidence for the IR-finite anzatz over the IR-infinite one.

4.12 Estimating Parameters - Bayesian Model Averaging

We have estimated the parameters of our fit anzätze, for example ν, and through a bootstrap resampling, their errors. However, for each fit anzatz, we have a choice of $\bar{B}$ and $gL_{\min}$. These choices introduce a systematic error into our results since each choice of these meta-parameters gives different predictions of ν. The question is twofold:

- How do we combine the different estimates to give an overall estimate?
- What is the error on this final estimate?

In [1], statistical and systematic errors were calculated separately. The statistical errors of parameters were calculated through a bootstrap resampling of the central fit. The systematic errors were calculated by finding the maximum deviation in the fit parameters among fits deemed acceptable (with $p > 0.05$ and more than 15 degrees of freedom). The minimum number of degrees of freedom is, however, an arbitary choice. If chosen to be a small, then the systematic error bars are very large and likely overly conservative, while if chosen to be large, we may underestimate the systematic error. The method also has the downside that all acceptable fits are included equally, regardless of the number of degrees of freedom they have or their p-values.

An alternative approach is available, however, where all fits can be included, even those with unacceptable p-values! This approach naturally offers more weight to fits with more degrees of freedom and also more weight to fits with better p-values. The approach in question is called Bayesian Model Averaging (BMA) and has been used by the statistics community [101–105] for several decades. However, in the context of lattice physics this technique is more recent[5]. In this section we will be referring to the more recent paper [10] where the authors present a pedagogical introduction to this method.

The basic idea of this method is quite simple: we treat each choice of $gL_{\min}$ and $\bar{B}$ as a different model, which we label by M_i, with $i \in 1, 2, ..., N_M$, and $N_M = |\{gL_{\min}\}| \times |\{\bar{B}\}|$, where $\{gL_{\min}\}$ is the set of all $gL_{\min}$ choices and $\{\bar{B}\}$ is the set of all $\bar{B}$ choices. Parameter results can then be obtained by a weighted average over all M_i,

$$\langle r \rangle = \sum_i \langle r \rangle_{M_i} P(M_i|D), \tag{4.51}$$

where r is a parameter of interest, $\langle r \rangle_{M_i}$ is the estimate of that parameter of interest under model M_i, and $P(M_i|D)$ is the probability of model M_i being true given the data. By definition, this probability must be normalized,

$$\sum_{i=1}^{N_M} P(M_i|D) = 1. \tag{4.52}$$

[5] See for example [106].

In [10], a formula for the variance of this estimate is also given,

$$Var(r) = \sum_{i=1}^{N_M} Var(r|M_i)P(M_i|D) + \sum_{i=1}^{N_M} \langle r \rangle_{M_i}^2 P(M_i|D) - \left(\sum_{i=1}^{N_M} \langle r \rangle_{M_i} P(M_i|D) \right)^2 . \quad (4.53)$$

The probability of the model given the data $P(M|D)$ can be found using Bayes Theorem as in section (4.11). In [10], the authors argue that the naive multivariate Gaussian likelihood function requires a bias correction term. They go on to derive that, in the large statistics limit, $P(M|D)$ approaches $\exp(-\mathrm{AIC_M}/2)$, where $\mathrm{AIC_M}$ is the Akaike information criterion [93–95],

$$\mathrm{AIC_M} = \chi^2(\alpha^*) + 2k, \quad (4.54)$$

where k is the number of degrees of freedom of the fit and α^* are the parameter values at the point of maximum likelihood, and $\chi^2(\alpha)$ is defined in eq. (4.49).

There is a very neat corollary to this: we can consider models which only fit a subset of the data $D' \subset D$ to be models that fit all the data, except they have extra parameters specifically tuned to fit exactly through data points in $D \setminus D'$ [10]. These data points will then contribute nothing to the χ^2, but we must recognise that the model effectively has N_{cut} more parameters, where $N_{\mathrm{cut}} = |D \setminus D'|$, so

$$k = k_0 + N_{\mathrm{cut}}, \quad (4.55)$$

where k_0 is the number of parameters the fit minimises, while N_{cut} is the number of data points the fit ignores. Given we are in the large statistics limit,

$$\langle r_i \rangle_M = r_i^*, \quad (4.56)$$

while the standard deviation of this estimate is given by bootstrap resampling [80].

4.13 Results: Parameter Estimates

It is most appropriate to perform a BMA for both the choice of gL_{min} and the choice of $\bar{B}$ simultaneously. The results for the parameter ν using model B_1 are shown in figure 4.12. The overall average is given by

$$\langle \nu \rangle = \sum_{\bar{B}} \sum_{gL_{\mathrm{min}}} \langle \nu \rangle_{\bar{B}, gL_{\mathrm{min}}} w(\bar{B}, gL_{\mathrm{min}}), \quad (4.57)$$

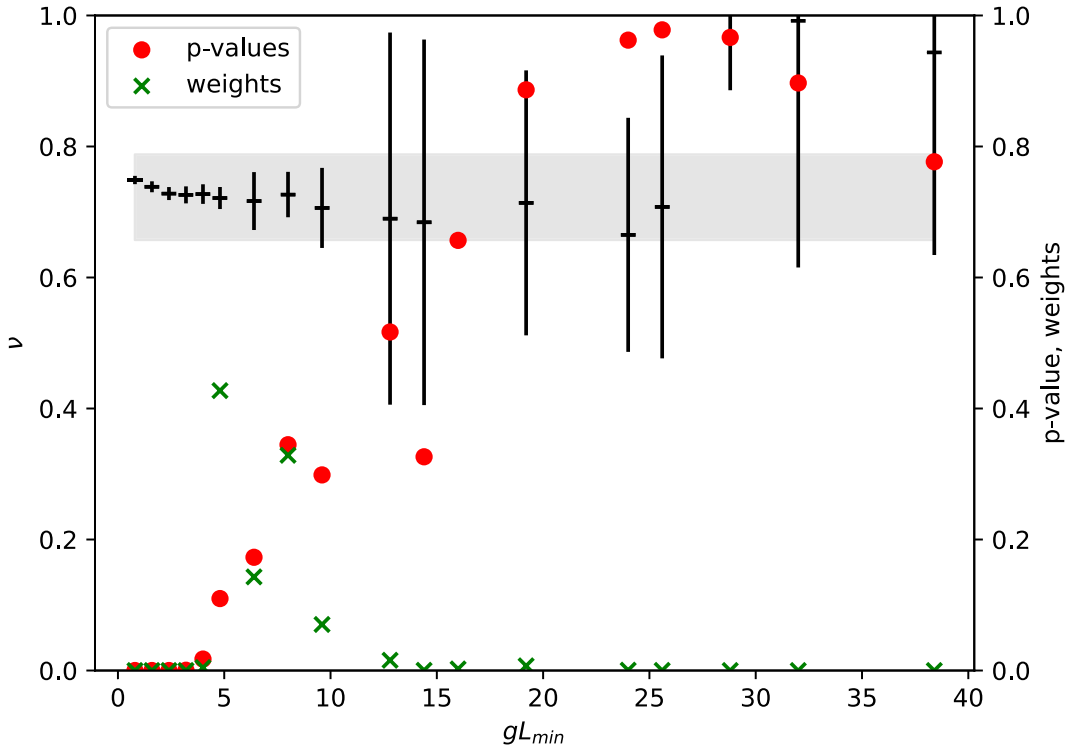

FIGURE 4.11: The parameter ν estimated by a Bayesian model averaging (BMA) procedure for the $N = 2$ data with a $\bar{B} = 0.52$ and $\bar{B} = 0.53$ and model B_1. The p-values of each fit are shown by the red dots, while the central values and statistical errors of each fit are shown by the black error bars. The green crosses show the weight put on each fit in the BMA average. The grey band shows the one-sigma range on the overall parameter estimate in this procedure.

where w is the Bayesian model averaging weight. The green crosses in the figure show the partial sum

$$W(\bar{B}) = \sum_{gL_{\min}} w(\bar{B}, gL_{\min}). \tag{4.58}$$

Shown in table 4.6 are the overall results for all model parameters of all models and choices of N, with a $\log(g)$ term. The error bars on many of the fits appear to be large. One possible cause of this is that *all* fits are included in the average, including those with parameter values compatible with zero. These fits potentially suffer from over-fitting. Therefore, in table 4.7, we present the same results except that fits with parameters compatible with zero are given zero weight in the Bayesian model average.

The overall estimates for α, β, and ν are plotted in figures 4.13, 4.14, and 4.15 respectively. These plots show results for all models (with the IR-finite anzatz) and all N studied. The parameter values are mostly consistent across all model choices. One exception is model A results for $N = 5$ data. In figure B.13, we see that model A provides a

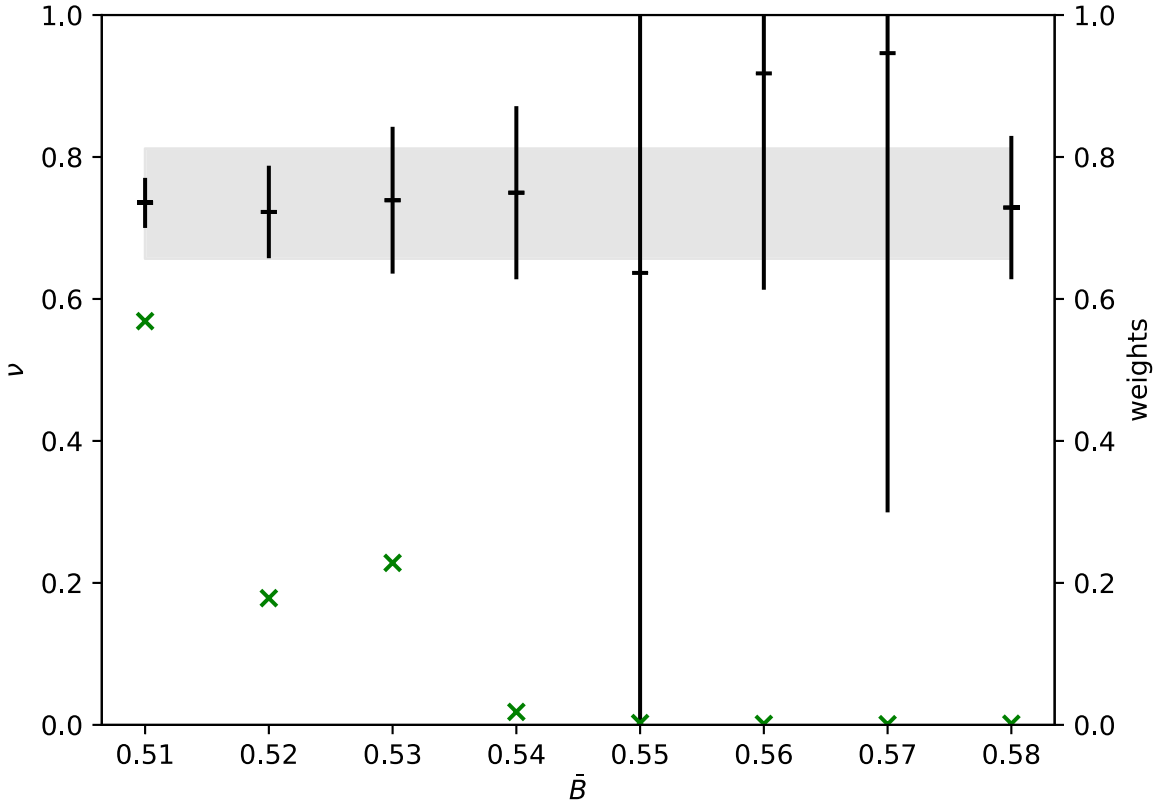

FIGURE 4.12: The parameter ν estimated by a Bayesian model averaging (BMA) procedure for $N = 2$ data with model B_1. The black error bars show, for each choice of $\bar{B}$, the estimate of ν from a BMA procedure across all $gL_{\min}$ choices. The green crosses show the relative weight of each $\bar{B}$ choice in the overall average.

very poor fit to the $N = 5$ data, which likely explains this discrepancy. The estimates of ν for $N = 2$ and $N = 3$ are compatible with estimates from previous lattice determinations in $O(N)$-vector theories. While our $N = 2$ model is equivalent to an $O(3)$-vector model, our $N = 3$ model is in the same universality class as an $O(8)$-vector model [92]. In an $O(3)$-vector model study, a value of $\nu = 0.710(2)$ was obtained [107] which is compatible with our estimates for all models except model D (where it is compatible at a two-sigma level). For all models except A, estimates of ν from $N = 3$ data are compatible with results from a study on the $O(8)$-vector model. In this study, a value of $y = 1.1752(10)$ was estimated (where $y = 1/\nu$) [6]. Model A however does not provide a good fit quality to $N = 3$ data, so this discrepancy is perhaps not surprising.

[6] This gives a value of $\nu = 0.851$ to three decimal places

N	model	α	β	ν
2	A_1	α: 0.0007(15)	β: 1.126(74)	ν: 0.69(85)
2	B_1	α: 0.00150(69)	β: 1.059(44)	ν: 0.734(73)
2	C_1	α: 0.0013(69)	β: 1.074(40)	ν: 0.7(2.2)
2	D_1	α_1: 0.0009(32), α_2: 0.0010(25)	β: 1.079(52)	ν: 0.8(1.4)
3	A_1	α: −0.00018(65)	β: 1.105(45)	ν: 0.764(25)
3	B_1	α: −0.00027(61)	β: 1.092(43)	ν: 0.789(29)
3	C_1	α: −0.00033(68)	β: 1.082(43)	ν: 0.83(12)
3	D_1	α_1: −0.0008(19), α_2: −0.0005(13)	β: 1.098(47)	ν: 0.84(68)
4	A_1	α: 0.00187(91)	β: 1.052(37)	ν: 0.707(60)
4	B_1	α: 0.00190(63)	β: 1.044(17)	ν: 0.68(14)
4	C_1	α: 0.0015(22)	β: 1.04(17)	ν: 0.77(40)
4	D_1	α_1: 0.00148(59), α_2: 0.00160(57)	β: 1.051(24)	ν: 0.791(60)
5	A_1	α: 0.00044(72)	β: 1.043(24)	ν: 0.688(73)
5	B_1	α: −0.00052(26)	β: 1.058(11)	ν: 0.765(20)
5	C_1	α: −0.00055(27)	β: 1.056(11)	ν: 0.772(58)
5	D_1	α_1: −0.00080(65), α_2: −0.00047(33)	β: 1.052(14)	ν: 0.81(51)

TABLE 4.6: Full parameter results obtained with Bayesian model averaging, for each value of N and for IR-finite model anzätze. The width parameter is chosen at 0.5.

N	model	α	β	ν
2	A_1	α: 0.00170(79)	β: 1.066(52)	ν: 0.698(27)
2	B_1	α: 0.00152(66)	β: 1.057(43)	ν: 0.734(33)
2	C_1	α: 0.00126(66)	β: 1.074(41)	ν: 0.728(73)
2	D_1	α_1: 0.00096(80), α_2: 0.00111(78)	β: 1.075(48)	ν: 0.768(41)
3	A_1	α: −0.00080(55)	β: 1.149(39)	ν: 0.775(38)
3	B_1	α: −0.00073(56)	β: 1.101(34)	ν: 0.88(12)
3	C_1	α: −0.00103(45)	β: 1.122(27)	ν: 0.91(11)
3	D_1	α_1: −0.00121(52), α_2: −0.00086(51)	β: 1.119(37)	ν: 0.875(34)
4	A_1	α: 0.00187(91)	β: 1.052(37)	ν: 0.707(60)
4	B_1	α: 0.00132(48)	β: 1.039(18)	ν: 0.810(82)
4	C_1	α: 0.00142(41)	β: 1.039(18)	ν: 0.793(89)
4	D_1	α_1: 0.00148(58), α_2: 0.00160(56)	β: 1.051(24)	ν: 0.792(59)
5	A_1	α: 0.00093(59)	β: 1.036(27)	ν: 0.631(55)
5	B_1	α: −0.00052(26)	β: 1.058(11)	ν: 0.765(20)
5	C_1	α: −0.00055(27)	β: 1.056(11)	ν: 0.773(34)
5	D_1	α_1: −0.00082(34), α_2: −0.00049(30)	β: 1.052(14)	ν: 0.813(61)

TABLE 4.7: Full parameter results obtained with Bayesian model averaging, for each value of N and for IR-finite model anzätze. Only fits with $gL_{\min}$ and $\bar{B}$ choices giving parameter estimates not compatible with zero are included. The width parameter is chosen at 0.5

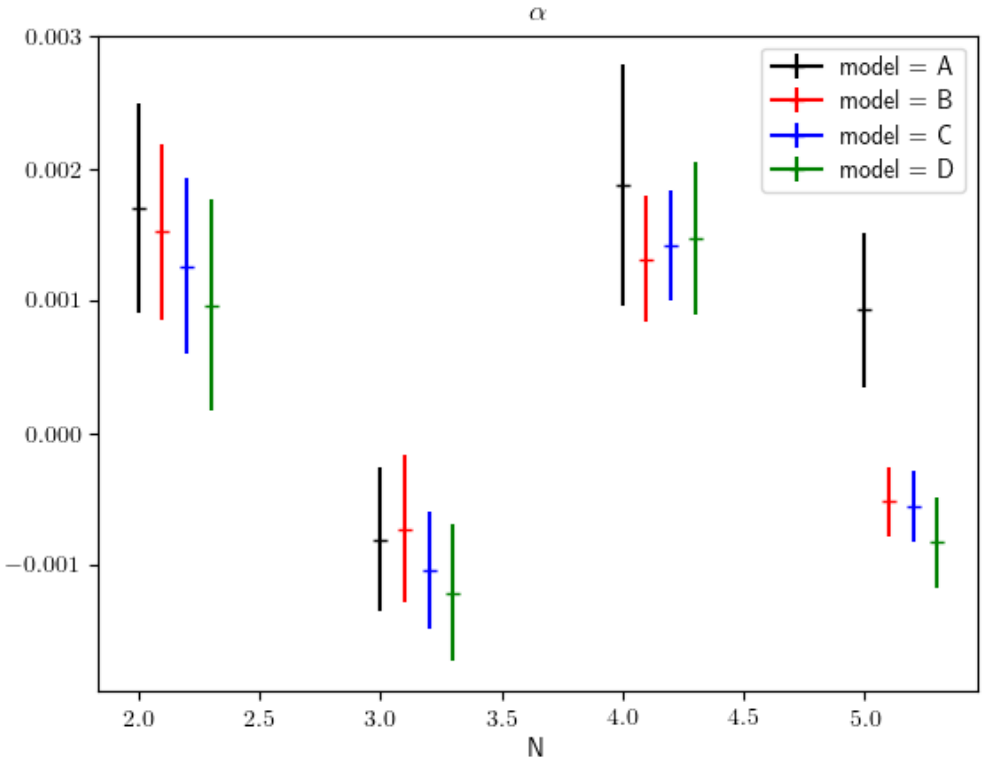

FIGURE 4.13: Bayesian Model Averaging estimates for the parameter α for IR-finite anzätze for each model type and value of N. Only fits with parameter values not compatible with zero at the one-sigma level are included.

4.14 Results: Critical Masses

To estimate the infinite-volume critical masses for each g and N, we also use a Bayesian Model Averaging scheme. The same weights are used as for the parameter estimates. However, the quantity averaged for each choice of gL_{min} and $\bar{B}$ is instead the critical mass. The results of this averaging keeping fits with parameter estimates compatible with zero are shown in table 4.8. While the critical mass estimates from an average not including fits with parameter estimates compatible with zero are shown in table 4.9. In general, the errors on the critical mass estimates are small compared to the size of the critical masses.

4.15 Results: Validity of the Fitting Range

When we compare IR-finite and IR-infinite anzätze without any correction-to-scaling terms, the IR-finite anzatz is heavily favored over the IR-infinite anzatz, both from a

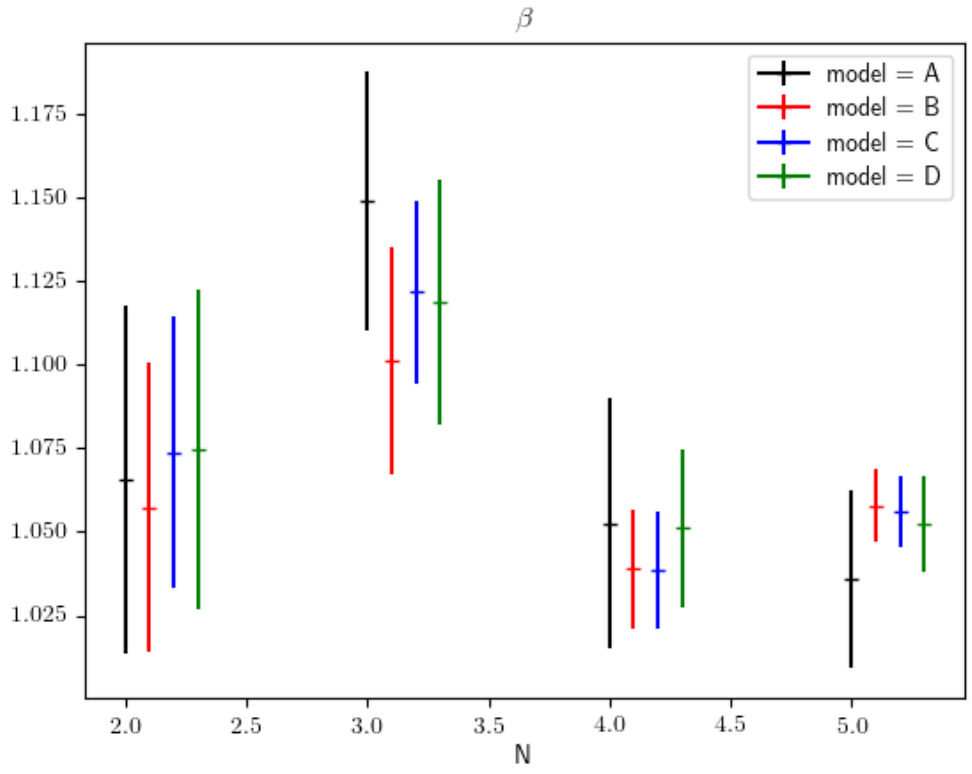

FIGURE 4.14: Bayesian Model Averaging estimates for the parameter β for IR-finite anzätze for each model type and value of N. Only fits with parameter values not compatible with zero at the one-sigma level are included.

N	model	$m_c^2(g = 0.1)$
2	A_1	−0.031336(12)
2	B_1	−0.0313433(38)
2	C_1	−0.031343(69)
2	D_1	−0.031345(34)
3	A_1	−0.0418927(37)
3	B_1	−0.0418962(37)
3	C_1	−0.0418995(50)
3	D_1	−0.041900(20)
4	A_1	−0.0455003(43)
4	B_1	−0.0455027(72)
4	C_1	−0.045508(23)
4	D_1	−0.045505(46)
5	A_1	−0.0471736(47)
5	B_1	−0.0471783(20)
5	C_1	−0.0471792(21)
5	D_1	−0.0471830(61)

TABLE 4.8: Critical masses estimated using IR-finite anzätze and Bayesian model averaging. The width parameter is chosen at 0.5

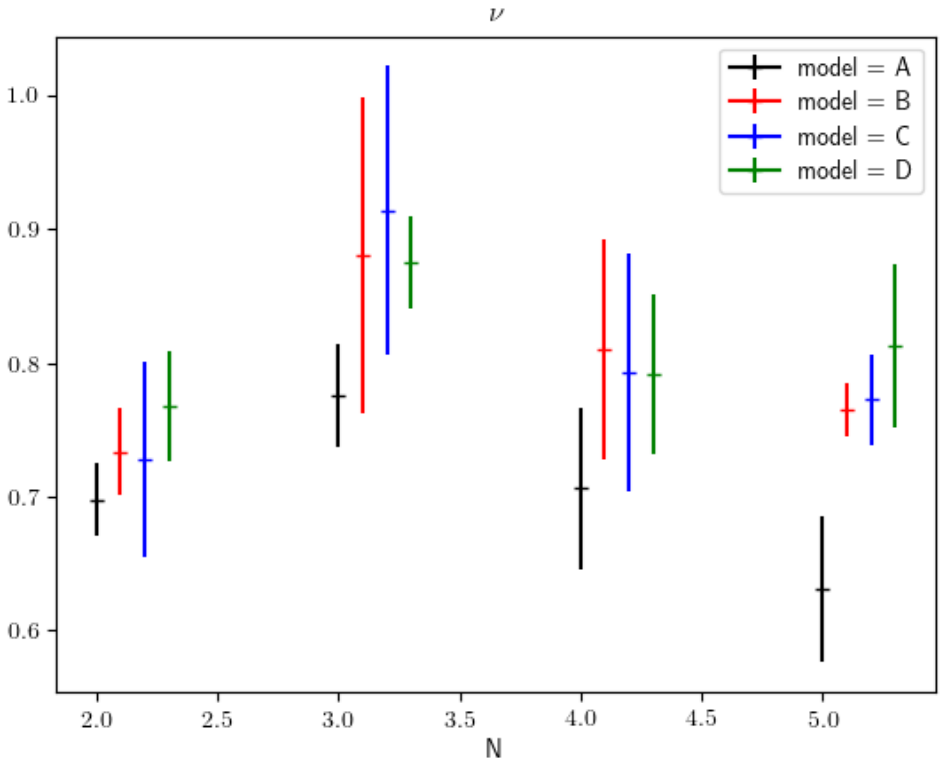

FIGURE 4.15: Bayesian Model Averaging estimates for the parameter ν for IR-finite anzätze for each model type and value of N. Only fits with parameter values not compatible with zero at the one-sigma level are included.

N	model	$m_c^2(g = 0.1)$
2	A_1	−0.0313399(39)
2	B_1	−0.0313433(35)
2	C_1	−0.0313425(36)
2	D_1	−0.0313451(43)
3	A_1	−0.0418899(33)
3	B_1	−0.0418995(45)
3	C_1	−0.0418986(40)
3	D_1	−0.0419004(42)
4	A_1	−0.0455003(43)
4	B_1	−0.0455098(47)
4	C_1	−0.0455089(50)
4	D_1	−0.045505(45)
5	A_1	−0.0471709(44)
5	B_1	−0.0471783(20)
5	C_1	−0.0471792(20)
5	D_1	−0.0471830(26)

TABLE 4.9: Critical masses estimated using IR-finite anzätze and Bayesian model averaging. Only fits with $gL_{\min}$ and $\bar{B}$ choices giving parameter estimates not compatible with zero are included. The width parameter is chosen at 0.5

frequentist and a Bayesian perspective. However, through the addition of correction-to-scaling terms both anzätze can fit a large amount of data. In some fits with correction-to-scaling, the IR-infinite anzätze outperform the IR-finite ones. This suggests that the data contain $\log(L)$-like behavior, but only for small values of gL, where corrections-to-scaling are also present. That these two phenomena are present together is perhaps not so surprising, as they both relate to finite-size effects. An explanation of how $\log(L)$ behavior could manifest is that the infrared cut-off predicted by JTAP cannot propagate on the lattice. This is because the lowest momentum mode on the lattice, other than $p = 0$, is given by $2\pi/L$. Naïvely, we expect that $g^2_{\mathrm{YM}} = g/N$, which sets the energy scale of the theory, will be replaced by $2\pi/L$ in the log when $g/N < 2\pi/L$. This corresponds to a crossover point at

$$\frac{2\pi}{L} = \frac{g}{N} \tag{4.59}$$

$$\implies gL = 2\pi N. \tag{4.60}$$

For gL values less than $2\pi N$, we expect the scaling to have $\log(L)$-like behavior. The constant 2π multiplying N may have a different value, however, as the mechanism for this regime change is not well understood.

We can observe this change in scaling in figure (4.16). As the value of gL_{min} is increased, the superior fit transitions from a $\log(L)$ fit to a $\log(g)$ fit (as can be seen by the transition in color from red to blue). For each value of $\bar{B}$, there is a different transition - shown by the purple bars - where the fit transitions from a fit preferring $\log(L)$ to one preferring $\log(g)$. We are interested in finding the value of the transition point, ρ, where presumably data with $gL_{\mathrm{min}} < \rho$ follow a $\log(L)$ scaling and $gL_{\mathrm{min}} > \rho$ follow a $\log(g)$ scaling. We define two types of crossover point:

- ρ_1: For each $\bar{B}$, calculate the quantity $X = (1/2)(gL_{\mathrm{below}} + gL_{\mathrm{above}})$ where gL_{below} is the largest gL_{min} value below the purple line and gL_{above} is the smallest gL_{min} value above the purple line. Then determine the median value of X across the different choices of $\bar{B}$.

- ρ_2: First define the function $Y(x)$, where Y returns the median value of gL (among the datapoints) in a fit with $gL_{\mathrm{min}} = x$. The motivation is that we expect a transition of preferred fit going from a majority of DoF with $gL < \rho$ to a majority with $gL > \rho$, e.g. when the median value of gL is equal to ρ. We then calculate $Z = (1/2)(Y(gL_{\mathrm{above}}) + Y(gL_{\mathrm{below}}))$ for each value of $\bar{B}$ and then take the median as our estimate.

In both cases, a median across the choice of $\bar{B}$ was used to prevent the result from being sensitive to outliers. The results of this transition study are given in tables 4.10 and 4.11 for ρ_1 and ρ_2 respectively. For models A and D, there is no transition since $\log(g)$

N	A	B	C	D
2	0	5.0	2.4	0
3	0	7.2	2.8	0
4	0	21.6	13.6	0
5	0	27.2	11.2	0

TABLE 4.10: Estimates for the transition value of gL estimated by the ρ_1 estimator as defined in section 4.15. Estimates are given for each model type, as defined in table (4.2). An entry of 0 represents the fact that no transition is observed.

N	A	B	C	D
2	0	14.4	12	0
3	0	17.6	12.8	0
4	0	28.8	24.8	0
5	0	35.2	21.6	0

TABLE 4.11: Estimates for the transition value of gL estimated by the ρ_2 estimator as defined in section 4.15. Estimates are given for each model type, as defined in table (4.2). An entry of 0 represents the fact that no transition is observed.

model is superior across the range of gL_{min}. The results show the expected trend that the transition increases with increasing values of N. By performing a linear fit to the results using `numpy.polyfit` we estimate the transition occurs at $gL = \chi N$, where χ is given by 4.7 and 6.9 for model B with the ρ_1 and ρ_2 metrics respectively and 2.3 and 5.0 for model C similarly. We note that all estimates are of the same order as 2π.

To confirm this picture, fits were done using models A_3, B_3, C_3, and D_3 to investigate the significance of log(g) and log(L) terms when both are included. While the $\log(g)$ terms are significant across the range fitted, the $\log(L)$ term only gains a significant coefficient when the gL_{min} value of the fit is sufficiently small: see figure (4.17). Similar results are seen for other values of N and choices of model, however, the signal is not always as clear.

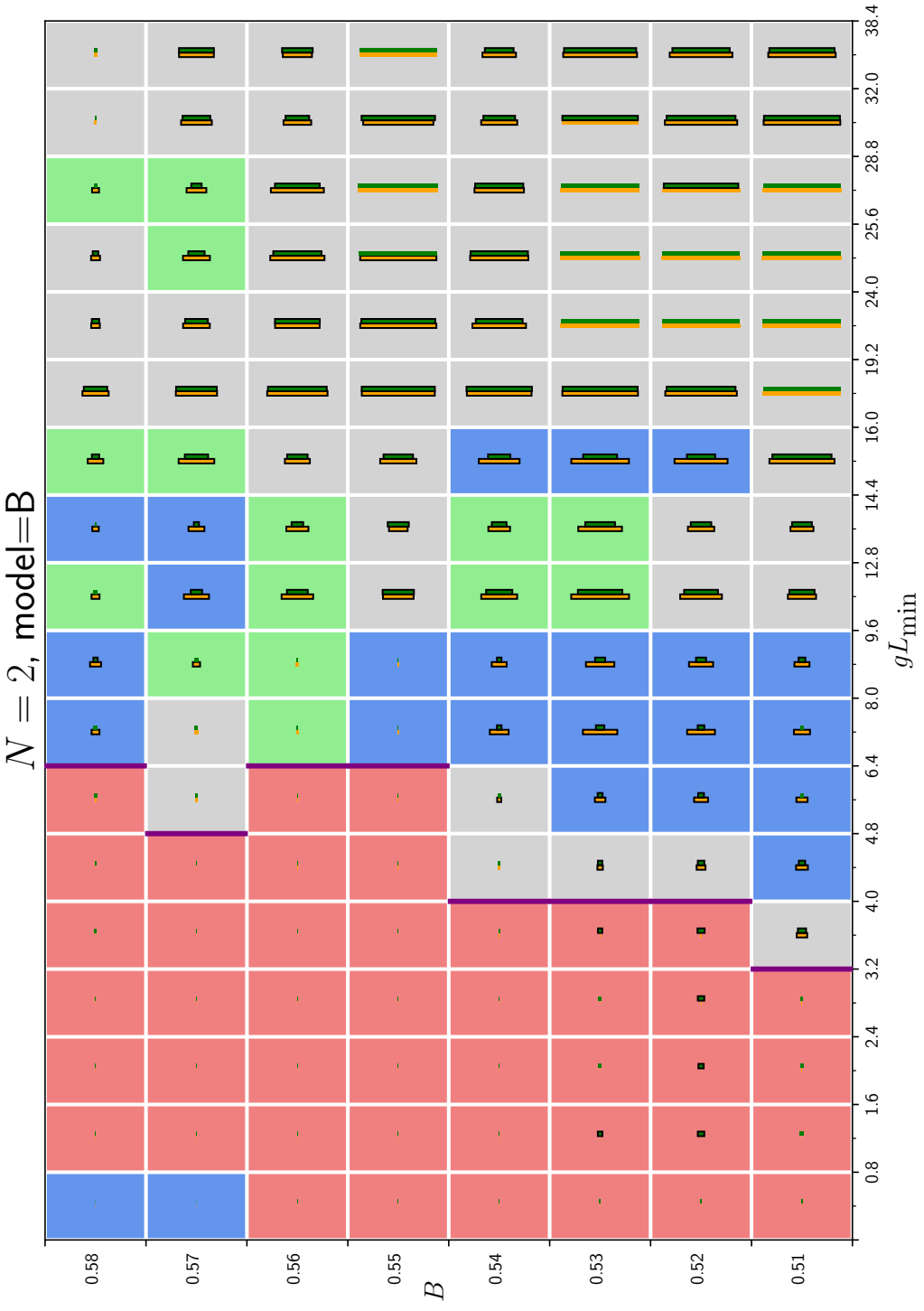

FIGURE 4.16: Comparison between the IR-finite (model B_1) and IR-infinite (model B_2) fits for $N = 2$ data. The orange (IR-finite) and green (IR-infinite) bars show the p-values of the fit. If the p-value is between 0.05 and 0.95, the bar has a black border. The colour scale represents the difference in the χ^2 value of the fits. If χ_1 is the IR-finite χ^2 and χ_2 is the IR-infinite χ^2 then fits are red if $\chi_1 > \chi_2 + 4$, orange if $\chi_2 + 2 < \chi_1 < \chi_2 + 4$, grey if $\chi_2 - 2 < \chi_1 < \chi_2 + 2$, green if $\chi_2 - 2 > \chi_1 > \chi_2 - 4$ and blue if $\chi_1 < \chi_2 - 4$.

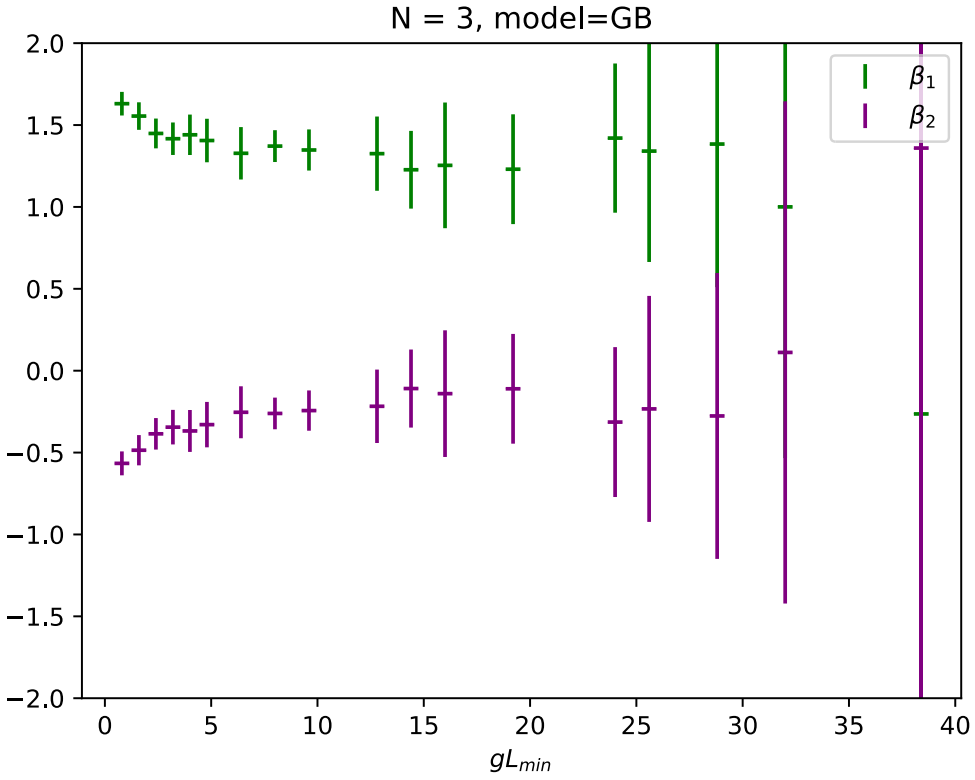

FIGURE 4.17: Parameter estimates of β_1 and β_2 in model B_3 with $N = 3$ data. For each value of $gL_{\min}$ shown, the parameter estimates are calculated by performing a Bayesian Model Averaging procedure over all values of $\bar{B}$.

Chapter 5

Renormalization

In the previous chapter, we demonstrated that the dual theory of our holographic model of the early universe is well-behaved in the infrared limit, validating the use of holographic cosmology models to model our early universe. We have also demonstrated how to make predictions of this model non-perturbatively through a lattice regulator and computer simulation. We are therefore empowered to once again turn our attention to the holographic dictionary (eq. (2.25)) and the EMT two-point function (eq. (2.22)).

In the continuum, the correlation functions of the EMT satisfy certain relations called Ward-Takahashi identities which result from Noether's theorem. These conditions result from the EMT's role as the collection of Noether currents of spacetime symmetries. On the lattice, the discretization of spacetime causes these relations to be violated, due to mixing of the EMT with other operators and the presence of contact terms. To define a renormalized EMT, it is therefore necessary to enforce the continuum Ward-Takahashi identities.

While this allows us to define a renormalized EMT operator, additional UV-divergent contributions arise when considering the two-point function of the (renormalized) EMT. These divergences are due to additional contact terms and include a q^2/a-divergence. The focus of this chapter will be looking at removing these divergences from the data using Laplace transforms. If successful, this will take us one step closer to testing holographic cosmology against CMB data.

5.1 Lattice Symmetries and the EMT

Quantum field theories that do not incorporate gravity are usually defined in a spacetime with the symmetries of special relativity. All operators in the theory transform covariantly under representations of *Poincaré* transformations, as well as spatial translations.

Scalar operators for example, are invariant under the action of Poincaré transformations,

$$\langle \mathcal{O}_1(x)\mathcal{O}_2(0)\rangle = \langle \mathcal{O}_1(Tx+y)\mathcal{O}_2(y)\rangle \qquad T \in O(1,d-1), \qquad x,y \in R^d. \tag{5.1}$$

Performing a Wick-Rotation transforms the symmetry group to rotations ($SO(d)$) and translations. However, due to the lattice being hypercubic, only a subset of this symmetry group survives, namely the Hypercubic symmetry group and the set of discrete translations. Because of this, lattice actions can include operators that do not transform covariantly under rotations. For example, scalar operators can be present that are invariant under the action of the hypercubic group, but not $O(d)$. While these terms are not physically relevant, they may arise during the renormalization procedure and so must be removed.

In order to make predictions of the cosmic microwave background radiation in holographic cosmology, the object of interest is the energy-momentum tensor (EMT), or more specifically it's two-point function (see eqs. (2.25) and (2.22)). Setting $q = (0,0,q)$ in this equation and looking at $\langle T_{22}T_{22}\rangle(q)$ we see that the $\pi_{22} = 0$ and $\Pi_{2222} = 0$, so the physical q^3 term we would expect from perturbation theory is zero. However, there are $O(3)$ breaking terms which appear at one-loop order, for example a non-zero δ_{2222} contribution [108].

As well as contributions from operators which are not Poincaré covariant, other discretization effects may remain. As the lattice spacing a is sent to zero, the maximum momentum on the lattice approaches infinity ($p_{\max} = \pi/a$), causing ultraviolet divergences to become problematic. To remove these infinities, we first have to identify them (through the scaling of terms that diverge as the lattice spacing gets smaller) and subtract them from our result. This method is considerably more direct than other regularization methods. For example, in a dimensional regularization scheme the dimension of the system is set to a small deviation ϵ away from the physical dimension. This deviation causes ultraviolet-divergences to be expressible as powers of $1/\epsilon$. By adding counter terms to the action, which scale in the same way with ϵ, but with opposite sign, the infinities can be removed.

Before we tackle the two-point function of the EMT, it is instructive (and necessary) to renormalize the operator itself. This means regulating the operator $T_{\mu\nu}$ such that the one-point expectation value of the operator is finite in the ultraviolet limit. As touched on in section 2.7, it receives a correction from the improvement term with associated non-minimality parameter ξ. For the action, eq. (2.30), the EMT is [109–111]

$$T_{\mu\nu} = \frac{N}{g}\mathrm{Tr}\Bigg[2(\partial_\mu\phi)(\partial_\nu\phi) - \delta_{\mu\nu}\left(\sum_\rho(\partial_\rho\phi)^2 + (m^2 - m_c^2)\phi^2 + \phi^4\right) + \xi\left(\delta_{\mu\nu}\sum_\rho(\partial_\rho\phi)^2 - (\partial_\mu\phi)(\partial_\nu\phi)\right)\Bigg]. \tag{5.2}$$

As the EMT is the generator for spacetime symmetries, which differ on the lattice from the continuum, it may mix with other operators not present in the continuum are covariant under hypercubic symmetries but not $O(3)$ symmetries. In three dimensions, there is only one such operator,

$$O_3 = \delta_{\mu\nu}(N/g)\mathrm{Tr}\phi^2. \tag{5.3}$$

As these holographic models are massless, it is not known a priori what the physical lattice spacing a is for a given set of simulation parameters. To find this out requires fitting the predictions of our dual QFT, regulated on the lattice, to the CMB. We can, however, identify the contributions divergent as $a \to 0$, which need to be regularized before a continuum limit can be taken.

5.2 Ward-Takahashi Identities

One effective technique for removing UV-divergent lattice artifacts is to use Ward-Takahashi identities. These identities have the advantage of being true at all energy scales in the continuum. For the EMT, the relevant Ward-Takahashi identity is

$$\langle \partial^\mu T_{\mu\nu}(x)P(y)\rangle = -\langle \frac{\delta P(y)}{\delta\phi(x)}\partial_\nu\phi(x)\rangle, \tag{5.4}$$

where $P(y)$ is a composite operator located at y. On the lattice, this identity is broken by the addition of terms of the form $\langle X_\nu(x)P(y)\rangle$ on the right-hand side of this equation. These contributions scale like $1/a$ due to radiative corrections, so must be regularized. This is exactly the constraint considered in [2]. The non-minimality parameter can be taken to be zero ($\xi = 0$) as this term in the EMT trivially satisfies eq. (5.4) both in the continuum and on the lattice. The resulting lattice EMT is

$$T^0_{\mu\nu} = \frac{N}{g}\mathrm{Tr}\left[2(\bar\delta_\mu\phi)(\bar\delta_\nu\phi) - \delta_{\mu\nu}\left(\sum_\rho(\bar\delta_\rho\phi)^2 + (m^2 - m_c^2)\phi^2 + \phi^4\right)\right], \tag{5.5}$$

where the derivatives have been replaced by discrete derivatives and a superscript '0' has been added to indicate that this is the non-renormalized EMT. In four dimensions, there are five operators that the EMT can mix with in general [109]. However, dimensional counting implies that only one of these operators (O_3) can mix with the EMT in three dimensions. Therefore we choose $P(y) = O_3(y)$ for the probe operator. If P is such that the right-hand side of Eq. (5) is finite when $x \neq y$, the left-hand side, which contains the divergence of the EMT, is finite up to contact terms. In momentum space, the Ward-Takahashi identity that needs to be satisfied becomes simply

$$\bar q^\mu \langle T^R_{\mu\nu}\mathrm{Tr}\phi^2\rangle(\bar q) = 0, \tag{5.6}$$

where $\bar{q} = (1/a)\sin(aq)$ is the lattice momentum and the momentum space two-point function is defined by the Fourier transform

$$\langle T_{\mu\nu}\mathrm{Tr}\phi^2\rangle(q) = \left(\frac{N}{g}\right)^2 \sum_x e^{-iq\cdot x}\langle T_{\mu\nu}(x)\mathrm{Tr}\phi^2(0)\rangle. \tag{5.7}$$

We use T^R to indicate this condition needs to be satisfied by the renormalized EMT.

There are two divergent contributions that must be removed to satisfy eq. (5.6). One of them is a term due to a contact term where the operators are at the same point on the lattice ($x = y$) in eq. (5.4). This contributes a delta-spike in position space which corresponds to a constant factor in momentum space. The second divergent term results from the operator mixing. This operator mixing can be canceled by the addition of a $\mathrm{Tr}\phi^2$ term to the EMT,

$$T^R_{\mu\nu} = T^0_{\mu\nu} - \frac{g}{a}c_3\delta_{\mu\nu}\frac{N}{g}\mathrm{Tr}\phi^2, \tag{5.8}$$

where the constant c_3 needs to be tuned to the correct value. The procedure for doing this is outlined in the following section.

5.3 The Wilson Flow

Wilson Flow methods [112–115] have been applied in many renormalization contexts in lattice QFT, including application to renormalizing the EMT [116–123] and for composite operators [124–128] such as $\langle T_{\mu\nu}\mathrm{Tr}\phi^2\rangle$. The basic idea of the technique is to add an additional dimension, the *flow-time* t dimension. Operators are then "flowed" in this dimension according to the classical diffusion equation

$$\frac{\partial\rho(t,x)}{\partial t} = \partial^2\rho(t,x), \tag{5.9}$$

where ρ is the flowed field and $\rho(0,x) \coloneqq \phi(x)$. We can use these fields to define a new two-point function at flow time t,

$$C_{\mu\nu}(t,q) = \frac{N}{g}a^3\sum_x e^{-iq\cdot x}\langle T^R_{\mu\nu}(x)\mathrm{Tr}\rho^2(t,0)\rangle, \tag{5.10}$$

where setting $t = 0$ recovers the previous two-point function. The benefit of doing this is that the contact term at $x = 0$ is regulated by the smearing effect on the $\mathrm{Tr}\rho^2$ term. Specifically, the contact term decays like $1/\sqrt{t}$ with a flow time of t [2]. This is in contrast to the operator mixing contribution, which affects $T_{\mu\nu}(x)$ across all values of x and is thus not smeared away. By measuring C across a range of different flow times, the contact term can be fit in terms of this scaling. This allows the remaining operator-mixing divergence to be isolated for each t and extrapolated in the $t \to 0$ limit.

Our operator-mixing renormalization term (eq. (5.8)) shifts C,

$$C_{\mu\nu}(t,q) = C^0_{\mu\nu}(t,q) - \frac{g}{a} c_3 \delta_{\mu\nu} C_2(t,q), \tag{5.11}$$

where

$$C^0(t,q) = \frac{N}{g} a^3 \sum_x e^{-iq\cdot x} \langle T^0_{\mu\nu}(x) \mathrm{Tr}\rho^2(t,0)\rangle, \tag{5.12}$$

$$C_2(t,q) = \frac{N}{g} a^3 \sum_x e^{-iq\cdot x} \langle \mathrm{Tr}\phi^2(x) \mathrm{Tr}\rho^2(t,0)\rangle. \tag{5.13}$$

The EMT and the two-point function $C_{\mu\nu}(t,q)$ are renormalized once c_3 is tuned such that eq. (5.6) is satisfied. This line of approach has been undertaken in our paper [2], where the parameter c_3 has been estimated for an $SU(2)$ scalar holographic dual theory.

5.4 The TT-correlator

The challenge of renormalizing the two-point function $C_{\mu\nu\rho\sigma} = \langle T_{\mu\nu} T_{\rho\sigma}\rangle$ remains however. Even using the renormalized EMT, as found through the Wilson Flow method, additional divergences persist. These divergences include a constant $1/a^3$-term and a new q^2/a divergence. These divergences can be seen in the one-loop correlator, which has the general form [108]

$$C^{1-loop}_{\mu\nu\rho\sigma}(q) = A_{\mu\nu\rho\sigma}\frac{1}{a^3} + B_{\mu\nu\rho\sigma}\frac{q^2}{a} + D_{\mu\nu\rho\sigma} q^3 + E_{\mu\nu\rho\sigma} a q^4. \tag{5.14}$$

The q^2 divergent piece decomposes into an $O(3)$-covariant piece and contributions that break the $O(3)$ symmetry. The terms $A_{\mu\nu\rho\sigma}$ and $D_{\mu\nu\rho\sigma}$ likewise receive contributions from $O(3)$-covariant and $O(3)$-breaking contributions. The aq^4 term need not concern us as it will vanish in the continuum limit. The $1/a^3$-divergence can be removed by the subtraction of $C_{\mu\nu\rho\sigma}(q = 0)$ from the correlator values at other momenta or by using Wilson Flow methods. However, the q^2 term presents a new challenge, requiring a change of strategy. In the following sections, we explore the viability of using Laplace transforms to remove this divergence.

5.5 The Discrete Fourier Transform

While we have used Fourier-transformed fields and correlators throughout this thesis, it is useful to review how the Fourier transform is defined on the lattice. This will hopefully provide inspiration for how to define a Laplace transform on the lattice.

Due to periodicity, all fields and all operators of the fields obey $O(x + Le_i) = O(x)$ for all dimensions i. When defining a Fourier transform, it is only necessary to sum over a single period of the lattice. With this in mind, the Fourier transform of lattice quantities can be defined simply by analogy to a continuous Fourier transform,

$$\mathcal{FT}[O](q) \coloneqq \left(\frac{a}{L}\right)^d \prod_{i=1}^{d} \left(\sum_{n_i=0}^{n_L-1} e^{-iaq_i n_i} \right) O((an_1, an_2, ..., an_d)), \tag{5.15}$$

where d is the number of dimensions of the lattice, n_i are integers, $n_L = L/a$, and we recognise that the i^{th} component of the position $x_i = an_i$. All sums act on everything to the right, while the product acts only on what is bracketed. We denote this by the shorthand

$$\tilde{O}(q) = \left(\frac{a}{L}\right)^d \sum_{x \in \Lambda} e^{-iq \cdot x} O(x), \tag{5.16}$$

where we understand the sum to only include one period in each dimension, and the tilde is used to represent that $\tilde{O}$ is a momentum space quantity. The momenta q have components q_i which take values in $2\pi\tilde{n}_i/L$, where $\tilde{n}_i$ are integers. By definition all Fourier transformed functions are periodic in momentum space, with periodicity of $2\pi/a$, meaning only values of $\tilde{n}_i$ in $\{0, 1, 2, ..., N_L - 1\}$ define momenta with unique operator values. We denote this periodic lattice of points in momentum space by $\tilde{\Lambda}$, where the lattice sites are separated by $\frac{2\pi}{L}$ in each dimension. This definition of the Fourier transform on the lattice has the bonus of having an easy inverse operation.

Proposition 5.1. *The inverse Fourier transform is given by*

$$\mathcal{FT}^{-1}\left[\tilde{O}\right](x) = \prod_{\tilde{i}=1}^{d} \left(\sum_{\tilde{n}_{\tilde{i}}=0}^{n_L-1} e^{+i\frac{2\pi\tilde{n}_{\tilde{i}} x_{\tilde{i}}}{L}} \right) \tilde{O}\left(\left(\frac{2\pi\tilde{n}_1}{L}, \frac{2\pi\tilde{n}_2}{L}, ..., \frac{2\pi\tilde{n}_d}{L} \right) \right). \tag{5.17}$$

Proof.

$$\begin{aligned}
\mathcal{FT}^{-1}\left[\mathcal{FT}\left[O\right]\right](y) &= \prod_{\tilde{i}=1}^{d} \left(\sum_{\tilde{n}_{\tilde{i}}=0}^{n_L-1} e^{+i\frac{2\pi\tilde{n}_{\tilde{i}} y_{\tilde{i}}}{L}} \right) \tilde{O}\left(\left(\frac{2\pi\tilde{n}_1}{L}, \frac{2\pi\tilde{n}_2}{L}, ..., \frac{2\pi\tilde{n}_d}{L} \right) \right) \qquad (5.18) \\
&= \prod_{\tilde{i}=1}^{d} \left(\sum_{\tilde{n}_{\tilde{i}}=0}^{n_L-1} e^{+i\frac{2\pi a\tilde{n}_{\tilde{i}} n'_{\tilde{i}}}{L}} \right) \left(\frac{a}{L}\right)^d \prod_{i=1}^{d} \left(\sum_{n_i=0}^{n_L-1} e^{-i\frac{2\pi a n_i \tilde{n}_i}{L}} \right) O((an_1, an_2, ..., an_d)) \\
&= \left(\frac{a}{L}\right)^d \prod_{i=1}^{d} \left(\sum_{\tilde{n}_i=0}^{n_L-1} e^{+i\frac{2\pi\tilde{n}_i n'_i}{n_L}} \sum_{n_i=0}^{n_L-1} e^{-i\frac{2\pi n_i \tilde{n}_i}{n_L}} \right) O((an_1, an_2, ..., an_d)) \\
&= \left(\frac{a}{L}\right)^d \prod_{i=1}^{d} \left(\sum_{n_i=0}^{n_L-1} \sum_{\tilde{n}_i=0}^{n_L-1} e^{+i\frac{2\pi\tilde{n}_i (n'_i - n_i)}{n_L}} \right) O((an_1, an_2, ..., an_d)),
\end{aligned}$$

where $n'_i = y_i/a$. In the third line, the products were combined. In the final line, we pulled the sums over n_i past the exponential factor involving n'_i and $\tilde{n}_i$ factors and then

swapped the order of the sums. Noting that the inner O only depends on n_i, we can perform the inner sums over $\tilde{n}_i$ giving

$$\mathcal{FT}^{-1}\left[\mathcal{FT}\left[O\right]\right](y) = \left(\frac{a}{L}\right)^d \prod_{i=1}^{d}\left(\sum_{n_i=0}^{n_L-1}\frac{L}{a}\delta_{n_i n_i'}\right) O((an_1, an_2, ..., an_d)) \tag{5.19}$$
$$= O((an_1', an_2', ..., an_d')) = O(y), \check{\mathrm{C}}$$

where $\delta_{nn'}$ is equal to 1 when $n = n'$ and 0 otherwise. □

We denote this inverse transform by the shorthand

$$O(x) = \sum_{q\in\tilde{\Lambda}} e^{+iq\cdot x}\tilde{O}(q). \tag{5.20}$$

The invertability of the transform also implies that Fourier transforms are one-to-one maps between the spaces of discrete functions on Λ and $\tilde{\Lambda}$. There is, however, an ambiguity in the definition of the Fourier transform that we have brushed under the carpet. One is free to add integer constants ρ_i (one for each dimension) to the sums in the Fourier transform,

$$\tilde{O}(q;\rho) = \left(\frac{a}{L}\right)^d \prod_{i=1}^{d}\left(\sum_{n_i=\rho_i}^{\rho_i+n_L-1} e^{-iaq_i n_i}\right) O((an_1, an_2, ..., an_d)), \tag{5.21}$$

where ρ_i form the elements of the vector ρ. This ambiguity is not of any concern, however, since it does not affect the resulting $\tilde{O}(q)$, which we will shortly prove. The reason this ambiguity is highlighted is to prepare ourselves for the definition of a Laplace transform, where we will face similar ambiguity.

Proposition 5.2. *The definition $\tilde{O}(q)$ is independent of the choice of ρ_i.*

Proof. Let $\tilde{O}(q;(\rho_1, \rho_2, ..., \rho_d))$ be the Fourier transform defined with a set of values ρ_i.

$$\tilde{O}(q;(\rho_1+1, \rho_2, ..., \rho_d)) = \tilde{O}(q;(\rho_1, \rho_2, ..., \rho_d)) + \left(\frac{a}{L}\right)^d \prod_{i=2}^{d}\left(\sum_{n_i=\rho_i}^{\rho_i+n_L-1} e^{-iaq_i n_i}\right) \tag{5.22}$$
$$\left[O((a\rho_1 + L, an_2, ..., an_d))e^{-iq_1(a\rho_1+L)} - O((a\rho_1, an_2, ..., an_d))e^{-iq_1 a\rho_1}\right]$$
$$= \tilde{O}(q;(\rho_1, \rho_2, ..., \rho_d)) + \left(\frac{a}{L}\right)^d \prod_{i=2}^{d}\left(\sum_{n_i=\rho_i}^{\rho_i+n_L-1} e^{-iaq_i n_i}\right) O((a\rho_1, an_2, ..., an_d))\left[e^{-iq_1 L} - 1\right]e^{-iq_1 a\rho_i}$$
$$= \tilde{O}(q;(\rho_1, \rho_2, ..., \rho_d)),$$

where in the second line we have used the periodicity of O, while in the final line, we used $q_1 = 2\pi n_1/L$ to tell us that $e^{-iq_1 L} = 1$. By induction, we conclude that

$$\tilde{O}(q;(\rho_1+\rho',\rho_2,...,\rho_d)) = \tilde{O}(q;(\rho_1,\rho_2,...,\rho_d)), \tag{5.23}$$

where ρ' is any integer. Since all dimensions are treated equivalently in the definition of the Fourier transform, this relation holds for the other dimensions (components of ρ) as well. □

Due to the invariance of the Fourier transform under the choice of ρ, we drop the ρ when denoting $\tilde{O}(q)$, and choose $\rho = 0$ when necessary.

5.6 The Continuum Laplace Transform

In the infinite-volume continuum we use a definition of the a one-sided Laplace transform,

$$\bar{O}(q) = \int_{x_1=0^-}^{\infty}\int_{x_2=0^-}^{\infty}...\int_{x_d=0^-}^{\infty} dx_1 dx_2 ... dx_d e^{-q\cdot x} O(x), \tag{5.24}$$

where d is the number of dimensions and 0^- indicates that $O(x=0)$ is fully included in the integration interval. This can be formalized by using $-\epsilon$ as the lower limit of the integral (where $\epsilon > 0$) and then taking the limit $\epsilon \to 0$. As a result of this, the one-sided Laplace transform of the Dirac delta function is

$$\bar{\delta}(q) = \int_{x_1=0^-}^{\infty}\int_{x_2=0^-}^{\infty}...\int_{x_d=0^-}^{\infty} dx_1 dx_2 ... dx_d e^{-qx}\delta^d(x) = 1. \tag{5.25}$$

We use a bar to distinguish $\bar{O}(q)$ from the Fourier transform result $\tilde{O}(q)$. Ignoring the limits on the integrals, the Laplace transform is an analytical continuation of the Fourier transform related by sending $q \to -iq$. As long as $\tilde{O}(q)$ is only non-zero in the interval $[0^-, +\infty]$ in position space $\bar{O}(q) = \tilde{O}(-iq)$. This motivates a method to cancel the q^2 divergence that appears in the $\langle TT\rangle$-correlator. A Laplace transform of the same position space function that leads to a q^2 divergence would also be a q^2 divergence of equal magnitude but with the opposite sign.

Proposition 5.3. *In the infinite-volume continuum limit, the Laplace transform of a d-dimensional Laplacian operator acting upon a d-dimensional δ function is equal to $-q^2$ while the Fourier transform of the same function is equal to q^2.*

Proof. The Laplacian operator is defined by

$$\Delta = \frac{d^2}{dx_1^2} + \frac{d^2}{dx_2^2} + ... + \frac{d^2}{dx_d^2} = \sum_{i=1}^{d}\frac{d^2}{dx_i^2} \tag{5.26}$$

Applying this function to the δ-function gives

$$
\begin{aligned}
\bar{D}(p) &= \prod_{j=1}^{d}\left(\int_{x_j=0^-}^{\infty} dx_i e^{-p_j x_j}\right)\Delta(\delta^d)(x), \qquad (5.27)\\
&= \prod_{j=1}^{d}\left(\int_{x_j=0^-}^{\infty} dx_j e^{-p_j x_j}\right)\sum_{i=1}^{d}\frac{d^2}{dx_i^2}\delta^d(x)\\
&= \sum_{i=1}^{d}\prod_{j\neq i}\left(\int_{x_j=0^-}^{\infty} dx_j e^{-p_j x_j}\right)\int_{x_i=0^-}^{\infty} dx_i e^{-p_i x_i}\frac{d^2}{dx_i^2}\delta(x)\\
&= \sum_{i=1}^{d}\prod_{j=1,j\neq i}^{d}\left(\int_{x_j=0^-}^{\infty} dx_j e^{-p_j x_j}\right)\Bigg(\Bigg[e^{-p_i x_i}\frac{d^2}{dx_i^2}\delta(x)\Bigg]_{0^-}^{\infty}\to 0,\\
&\qquad\qquad - (-p_i)\int_{x_i=0^-}^{\infty} dx_i e^{-p_i x_i}\frac{d}{dx_i}\delta(x)\Bigg)\\
&= \sum_{i=1}^{d}\prod_{j=1,j\neq i}^{d}\int_{x_j=0^-}^{\infty} dx_j e^{-p_j x_j} p_i\Bigg(\Bigg[e^{-p_i x_i}\frac{d}{dx_i}\delta(x)\Bigg]_{0^-}^{\infty}\to 0,\\
&\qquad\qquad - (-p_i)\int_{x_i=0^-}^{\infty} dx_i e^{-p_i x_i}\delta(x)\Bigg)\\
&= \sum_{i=1}^{d}\prod_{j=1}^{d}\left(\int_{x_j=0^-}^{\infty} dx_j e^{-p_j x_j}\right)p_i^2 e^{-p_j x_j}\left(\delta(x)\right),\\
&= \sum_{i=1}^{d} p_i^2\prod_{j=1}^{d}\left(\int_{x_j=0^-}^{\infty} dx_j e^{-p_j x_j}\right)\left(\delta(x)\right)\\
&= \sum_{i=1}^{d} p_i^2,\\
&= p^2,
\end{aligned}
$$

where the boundary terms evaluate to 0 because the δ-function is zero there. The proof can be repeated with the Fourier transform. The only difference is that each time we take a derivative of $e^{-ip_i x_i}$ we bring down a factor of $(-ip_i)$ instead of $(-p_i)$, and so end up with an $i^2 = -1$ difference between the two answers. □

Given that Fourier transforms are one-to-one, we see that $\Delta(\delta)(x)$ is the position space representation of a q^2 distribution. With proposition 5.6 we obtain the relation

$$\mathcal{LT}\left[\mathcal{FT}^{-1}\left[q^2\right]\right] = -q^2, \qquad (5.28)$$

where $\mathcal{LT}$ is a Laplace transform defined in eq. (5.24).

5.7 Laplace Transforms on the Lattice

To make use of equation (5.28), we need a definition of the Laplace transform of the lattice that reduces to (5.24) in the infinite-volume continuum limit. The rough form of a lattice Laplace transform must be

$$\tilde{O}(q) = \sum_x e^{-q\cdot x} O(x), \tag{5.29}$$

by analogy to the continuum case. However, unlike with the Fourier transform, the exponential term is not periodic in $x_i \to x_i + L$ since $e^{-q_0 L} \neq 1$ for $q_0 \neq 0$. Therefore, an equivalent of proposition 5.2 for Laplace transforms does not hold, and the choice of ρ_i does matter. Since the cubic lattice we have described is the same size in all directions, we consider definitions with $\rho_1 = \rho_2 = \rho_3 = ... = \rho_d = \bar{\rho}$. Including a point $x = 0$ in the lattice restricts us to $\bar{\rho} \in \{-(n_L - 1), -(n_L - 2), -(n_L - 3), ..., -2, -1, 0\}$. Our definition of the Laplace transform is then

$$\bar{O}(q) = \left(\frac{a}{L}\right)^d \prod_{i=1}^{d} \left(\sum_{n_i=\rho}^{\rho+n_L-1} e^{-aq_i n_i} \right) O((an_1, an_2, ..., an_d)), \tag{5.30}$$

where we have dropped the bar on ρ so that it represents an integer rather than a vector. To define the Laplace transform on a lattice we need to consider the 0^- lower bound of eq. (5.24). The $2n$-th discrete derivative of the δ function will be non-zero from $-na$ to na. Therefore, to be consistent with the 0^- limit in the continuum, we must take $\rho = -ma$ or smaller where $2m$ is the highest derivative of the δ function that makes a significant contribution to $T(x)$. When we later take the $a \to 0$ limit, this lower bound will approach the continuum lower bound. Since contributions from negative values of x contribute greatly to the Laplace transform (they are proportional to e^{-qx}), we wish to make ρ as close to zero as possible to avoid the risk of amplifying noise unnecessarily.

The Laplacian operator on the lattice can be defined in many ways that are consistent with the continuum result when the limit $a \to 0$ is taken. We use the definition

$$\Delta(f) = \sum_i \delta_i^+(\delta_i^-(f)), \tag{5.31}$$

where Δ is the Laplacian and δ_i^+ and δ_i^- are the forward and backward derivatives in the x_i axis, defined in eqns. (3.7) and (3.8) respectively. Expanding out the derivatives gives

$$\Delta(f)(x) = \sum_{i=1}^{d} \frac{f(x - e_i) - 2f(x) + f(x + e_i)}{a^2}, \tag{5.32}$$

where e_i is the unit vector in the i-direction. We define the δ function on our lattice Λ by

$$\delta(x) = \begin{cases} \left(\frac{L}{a}\right)^d, & x_i = nL \,\forall\, i \in \{1, 2, ..., d\}, n \in \mathbb{Z} \\ 0, & \text{otherwise.} \end{cases} \tag{5.33}$$

This function has very simple Fourier and Laplace transforms,

$$\tilde{\delta}(p) = \left(\frac{a}{L}\right)^d \sum_{x\in\Lambda} e^{-ip\cdot x}\delta(x) = 1, \tag{5.34}$$

$$\bar{\delta}(p) = \left(\frac{a}{L}\right)^d \sum_{x\in\Lambda} e^{-p\cdot x}\delta(x) = 1, \tag{5.35}$$

where we note our earlier choice to include $x = 0$ in Λ so the only non-zero contribution to the Laplace transform of $\delta(x)$ is at $x = 0$. We define the discrete function $D(x)$ on the lattice by applying the discrete Laplacian to the δ function $D(x) = \Delta(\delta)(x)$. This evaluates to

$$D(x) = \begin{cases} \frac{1}{a^2}\left(\frac{a}{L}\right)^d, & x_i/a = n_i n_L \pm 1, \quad x_j/a = n_j n_L \forall j \neq i, \\ -\frac{2}{a^2} d \left(\frac{a}{L}\right)^d, & x_j/a = n_j n_L \forall j, \\ 0, & \text{otherwise,} \end{cases} \tag{5.36}$$

where $n_i, n_j \in \mathbb{Z}$. We can perform a forward Fourier transform on this function,

$$\begin{aligned} \tilde{D}(p) &= \left(\frac{a}{L}\right)^d \sum_{x\in\Lambda} e^{-ip\cdot x} D(x) \\ &= \frac{1}{a^2}\sum_{i=1}^{d}\left(e^{-ip_i(L-a)} + e^{-ip_i a}\right) - \frac{2d}{a^2} \\ &= \frac{1}{a^2}\sum_{i=1}^{d}\left(e^{+ip_i a} + e^{-ip_i a}\right) - \frac{2d}{a^2} \\ &= \frac{1}{a^2}\sum_{i=1}^{d}\left(e^{-\frac{ip_i a}{2}} - e^{+\frac{ip_i a}{2}}\right)^2 \\ &= \frac{1}{a^2}\sum_{i=1}^{d}\left(-2i\sin\left(\frac{p_i a}{2}\right)\right)^2 \\ &= -\frac{4}{a^2}\sum_{i=1}^{d}\sin^2\left(\frac{p_i a}{2}\right) \\ &(= -\hat{p}^2), \end{aligned} \tag{5.37}$$

where $\hat{p} = (2/a)\sin(ap/2)$ is the lattice momentum. In the third line we have used $p_i = 2\pi m/L$, where m is an integer to give $e^{ip_i L} = 1$. For the Laplace transform we

assume that $-n_L + 2 \leq \bar{\rho} \leq -1$:

$$\begin{aligned}
\bar{D}(p;\rho) &= \left(\frac{a}{L}\right)^d \prod_{i'=1}^{d} \left(\sum_{n_i=\rho}^{\rho+n_L-1} e^{-ap_i n_i} \right) D((an_1, an_2, ..., an_d)) = 1 \qquad (5.38)\\
&= \frac{1}{a^2} \sum_{i=1}^{d} \left(e^{+p_i a} + e^{-p_i a} \right) - \frac{2d}{a^2} \\
&= \frac{1}{a^2} \sum_{i=1}^{d} \left(e^{+\frac{p_i a}{2}} - e^{-\frac{p_i a}{2}} \right)^2 \\
&= \frac{1}{a^2} \sum_{i=1}^{d} \left(2\sinh\left(\frac{p_i a}{2}\right) \right)^2 \\
&= \frac{4}{a^2} \sum_{i=1}^{d} \sinh^2\left(\frac{p_i a}{2}\right).
\end{aligned}$$

If $|p|a \ll 1$ then we can Taylor expand $\tilde{D}(p)$:

$$\begin{aligned}
\tilde{D}(p) &= -\frac{4}{a^2} \sum_{i=1}^{d} \left(\frac{p_i a}{2} - \frac{1}{3!}\left(\frac{p_i a}{2}\right)^3 + O((p_i a)^5) \right)^2 \\
&= -\frac{4}{a^2} \sum_{i=1}^{d} \left(\left(\frac{p_i a}{2}\right)^2 - \frac{2}{3!}\left(\frac{p_i a}{2}\right)^4 + O((p_i a)^6) \right) \\
&= -p^2 + \sum_{i=1}^{d} \frac{p_i^4 a^2}{12} + O((p_i a)^6).
\end{aligned}$$

We can similarly Taylor expand $\bar{D}(p;\rho)$, giving

$$\bar{D}(p;\rho) = +p^2 + \sum_{i=1}^{d} \frac{p_i^4 a^2}{12}(1 + O((p_i a)^2)). \qquad (5.39)$$

We nicely recover the continuum Fourier and Laplace transforms in the continuum ($a \to 0$) limit. However, for finite lattice spacing, these transforms do not exactly cancel - instead leaving a $p^4 a^2$ term,

$$\bar{D}(p;\rho) + \tilde{D}(p) = \sum_{i=1}^{d} \frac{p_i^4 a^2}{6}(1 + O((p_i a)^2). \qquad (5.40)$$

5.8 Cancelling the q^2 Divergence

We are interested in lattice quantities of the form $\tilde{C}_2(q)$, which are the Fourier transforms of two-point functions involving two position space operators. Specifically, we are

interested in two-point functions involving the energy-momentum tensor,

$$\tilde{C}_{2,\mu\nu\rho\sigma}(q) = \mathcal{FT}\left[\langle T_{\mu\nu}(x)T_{\rho\sigma}(0)\rangle\right] = \langle \mathcal{FT}\left[T_{\mu\nu}(x)T_{\rho\sigma}(0)\right]\rangle, \tag{5.41}$$

where we can change the order of the Fourier transform and the lattice average due to linearity. We can estimate these quantities efficiently on the lattice by taking

$$\begin{aligned}
\langle \mathcal{FT}\left[T_{\mu\nu}(x)T_{\rho\sigma}(0)\right]\rangle &= \left(\frac{a}{L}\right)^d \sum_{x\in\Lambda} e^{-iq\cdot x}\langle T_{\mu\nu}(x)T_{\rho\sigma}(0)\rangle \\
&= \left(\frac{a}{L}\right)^d \sum_{x\in\Lambda} e^{-iq\cdot x}\left(\frac{a}{L}\right)^d \sum_{y\in\Lambda}\langle T_{\mu\nu}(x+y)T_{\rho\sigma}(y)\rangle \\
&= \left(\frac{a}{L}\right)^{2d} \sum_{y\in\Lambda}\sum_{x\in\Lambda} e^{-iq\cdot x}\langle T_{\mu\nu}(x+y)T_{\rho\sigma}(y)\rangle \\
&= \left(\frac{a}{L}\right)^{2d} \sum_{y\in\Lambda}\sum_{x\in\Lambda} e^{-iq\cdot (x-y)}\langle T_{\mu\nu}(x)T_{\rho\sigma}(y)\rangle \\
&= \left\langle \left(\frac{a}{L}\right)^d \left(\sum_{y\in\Lambda} e^{+iq\cdot y}T_{\rho\sigma}(y)\right)\left(\left(\frac{a}{L}\right)^d \sum_{x\in\Lambda} e^{-iq\cdot x}T_{\mu\nu}(x)\right)\right\rangle \\
&= \langle \tilde{T}_{\mu\nu}(q)(\tilde{T}_{\rho\sigma}(q))^*\rangle, \\
&= \langle \tilde{T}_{\mu\nu}(q)\tilde{T}_{\rho\sigma}(-q)\rangle,
\end{aligned} \tag{5.42}$$

where $\langle\cdot\rangle$ represents taking an ensemble average. We have used the periodicity of the lattice and translational invariance of the two-point function and, in the fifth line, redefined the dummy variable x by sending $x \to x - y$. Once these momentum space correlators have been calculated, we subtract the disconnected contribution

$$\left\langle \left(\frac{a}{L}\right)^d \sum_{x\in\Lambda} T_{\mu\nu}(x)\right\rangle\left\langle \left(\frac{a}{L}\right)^d \sum_{x'\in\Lambda} T_{\rho\sigma}(x')\right\rangle = \langle \tilde{T}_{\mu\nu}(0)\rangle\langle \tilde{T}_{\rho\sigma}(0)\rangle. \tag{5.43}$$

Note that this disconnected contribution needs to be calculated first before finding the connected momentum-space two-point correlator

$$\tilde{C}^c_{2,\mu\nu\rho\sigma}(q) = \langle \tilde{T}_{\mu\nu}(q)\tilde{T}_{\rho\sigma}(-q)\rangle - \langle \tilde{T}_{\mu\nu}(0)\rangle\langle \tilde{T}_{\rho\sigma}(0)\rangle. \tag{5.44}$$

From here on, we will drop the superscript "c", but the correlator $\tilde{C}^c_{2,\mu\nu\rho\sigma}(q)$ is understood as connected. Given the lattice estimates of these correlators, we define the map

$$m\left[\tilde{O}\right](q) = \tilde{O}(q) + \mathcal{LT}\left[\mathcal{FT}^{-1}\left[\tilde{O}\right]\right](q), \tag{5.45}$$

and apply it to $\tilde{C}^c_{2,\mu\nu\rho\sigma}(q)$. For our forward Laplace transform, we choose $\rho = -1$ as we wish to encapsulate q^2 divergences but are not concerned with q^4 divergences. This map (in the continuum limit) completely removes the q^2 divergence from $\tilde{O}$. The inverse Fourier transform is defined exactly as before, however to reduce noise, we define

an extra parameter, n_{max} in our Laplace transform,

$$\bar{O}(q;\rho,n_{max}) = \left(\frac{a}{L}\right)^d \sum_{x\in\Lambda} e^{-q\cdot x} O(x) \tag{5.46}$$

$$= \left(\frac{a}{L}\right)^d \prod_{i'=1}^{d} \left(\sum_{n_i=\rho}^{\rho+n_L-1} e^{-aq_i n_i} H(n_i;n_{\max}) \right) O((an_1, an_2, ..., an_d)), \tag{5.47}$$

where

$$H(n;n_{\max}) = \begin{cases} 1, & |n| \le n_{max} \\ 0, & \text{otherwise.} \end{cases} \tag{5.48}$$

We choose $n_{\max} = 1$ in all results presented here. This value was chosen as it minimizes the statistical noise added by the Laplace transform.

5.9 Fitting to a Functional Form

We are interested in fitting our correlator data to linear functional forms

$$f(c)(q) = \sum_{i=1}^{r} c_i f_i(q), \tag{5.49}$$

where $c = (c_1, c_2, ..., c_r)$ are free fit parameters. However since our data has been processed by m (defined in eq. (5.45)) we are interested in fitting $m\left[\tilde{C}_{2,\mu\nu\rho\sigma}\right](q)$ to $m[f(c)](q)$. Due to the linearity of m,

$$m[f(c)](q) = \sum_{i=1}^{r} c_i m[f_i](q). \tag{5.50}$$

Because of the rotational invariance of the physical correlators, and for more stability during the fitting, we choose to fit data in a single momentum direction. We define our observed data

$$y(q) \coloneqq m\left[\tilde{C}_{2,\mu\nu\rho\sigma}\right](q), \tag{5.51}$$

$$h(q;c) \coloneqq m[f(c)](q) = \sum_{i=1}^{r} c_i m[f_i](q). \tag{5.52}$$

In practice will only fit over a subset of all q values, for example over a single dimension $q = (q', 0, 0)$. This is because the n_L^3 values available to us are far too large compared to the number of configurations to allow for accurate determination of the $L^3\times L^3$ covariance matrix of the data. We also do not include the $q = 0$ point in the fit as it is observed to have high statistical noise. Lastly, we apply a cut $|q| \le q_{max}$ so that the residual q^4 term left by m is minimized. The covariance matrix Cov of our input data y can be calculated empirically by using a statistical bootstrap [80]. We then fit the functional

form f by minimizing

$$\chi^2(c) = (\vec{y} - \vec{h}(c)) \cdot \mathrm{Cov}^{-1} \cdot (\vec{y} - \vec{h}(c)), \tag{5.53}$$

where "$\cdot$" represents an inner product over the superscript indices and

$$\vec{y} = (y(q^{(1)}), y(q^{(2)}), ..., y(q^{(n_q)})), \tag{5.54}$$
$$\vec{h}(c) = (h(q^{(1)}; c), h(q^{(2)}; c), ..., h(q^{(n_q)}; c)), \tag{5.55}$$

where $q^{(R)}$ is the R^{th} momenta included in the fit, of which there are n_q in total. The covariance matrix is given by

$$\mathrm{Cov}^{RS} = \langle y(q^{(R)}) y(q^{(S)}) \rangle - \langle y(q^{(R)}) \rangle \langle y(q^{(S)}) \rangle \tag{5.56}$$

where the angle brackets represent ensemble averages. We can decompose $\vec{h}(c)$ as

$$\vec{h}(c) = \sum_i c_i \vec{h}_i, \tag{5.57}$$

where

$$\vec{h}_i = (m[f_i](q^{(1)}), m[f_i](q^{(2)}), ..., m[f_i](q^{(n_q)})). \tag{5.58}$$

Expressing χ^2 in these terms and taking a derivative allows us to solve the minimization procedure exactly,

$$\chi^2 = \chi^2(c) = \left(\vec{y} - \sum_j c_j \vec{h}_j\right) \cdot \mathrm{Cov}^{-1} \cdot \left(\vec{y} - \sum_k c_k \vec{h}_k\right) \tag{5.59}$$
$$\implies \frac{\partial \chi^2}{\partial c_i} = 2\vec{h}_i \cdot \mathrm{Cov}^{-1} \cdot \left(\vec{y} - \sum_k c_k \vec{h}_k\right) = 0$$
$$\implies \vec{h}_i \cdot \mathrm{Cov}^{-1} \cdot \vec{y} = \sum_k c_k \vec{h}_i \cdot \mathrm{Cov}^{-1} \cdot \vec{h}_k.$$

Doing this for all c_i yields the matrix equation, $Ac = b$, where

$$A = \begin{pmatrix} A_{00} & A_{01} & \cdots & A_{0n} \\ A_{10} & A_{11} & \cdots & A_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n0} & A_{n1} & \cdots & A_{nn} \end{pmatrix}, \qquad c = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix}, \qquad b = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix}, \tag{5.60}$$

where there are n basis functions in the fit and

$$A_{ij} = \vec{h}_i \cdot \mathrm{Cov}^{-1} \cdot \vec{h}_j, \tag{5.61}$$
$$b_i = \vec{h}_i \cdot \mathrm{Cov}^{-1} \cdot \vec{y}.$$

We can then solve this for the coefficients c.

5.10 Testing on Synthetic Data

To gain a better understanding of this novel method we test it on synthetic data, which has properties similar to our actual lattice data. We create synthetic data on a three-dimensional lattice with $N_L = L/a$ points in each dimension. In our $\langle TT \rangle$-correlator data, the covariance matrix is observed to be close to diagonal, so in our synthetic data we add an independent noise contribution $\epsilon X(q)$ for each q, where

$$X(q) \sim N(0,1). \tag{5.62}$$

This form of noise is motivated by empirical covariance matrices of our lattice data where covariance between different momenta is relatively small. Following the perturbative anzatz (eq. (2.37)), we have

$$\tilde{f}_1(q) = N^2 \left(\alpha_1 (|\hat{q}|)^3 + \beta_1 g (|\hat{q}|)^2 \log\left(\frac{|\hat{q}|}{g}\right) + \gamma (|\hat{q}|)^2 \right) + \epsilon X(q) + \eta_1, \tag{5.63}$$

where α_1, β_1, γ, ϵ and η_1 are fixed values we choose. Here we are abusing notation somewhat and using g and $\hat{q}$ to represent the unitless values ag and $a|\hat{q}| = 2\sin(q/2)$ so the above expression is unitless.[1] The term η_1 accounts for the effect of divergences of the form $\sim 1/a^3$. Once we have transformed our synthetic data under the map m we wish to fit the data points $m(\tilde{f}_1)(q)$ to a noiseless signal with no q^2 contact term,

$$\tilde{f}_2(q) = N^2 \left(\alpha_2 (|\hat{q}|)^3 + \beta_2 g (|\hat{q}|)^2 \log\left(\frac{|\hat{q}|}{ag}\right) \right) + \eta_2. \tag{5.64}$$

For simplicity in what follows, we use

$$f_\alpha(q) = N^2 |\hat{q}|^3, \tag{5.65}$$

$$f_\beta(q) = N^2 |\hat{q}|^2 g \log\left(\frac{|\hat{q}|}{ag}\right), \tag{5.66}$$

$$f_\gamma(q) = N^2 |\hat{q}|^2 (ag), \tag{5.67}$$

$$f_\epsilon(q) = X(q), \tag{5.68}$$

$$f_\eta(q) = 1. \tag{5.69}$$

Expressed in terms of these functions, we have

$$\tilde{f}_1(q) = \alpha_1 f_\alpha(q) + \beta_1 f_\beta(q) + \gamma f_\gamma(q) + \epsilon f_\epsilon(q) + \eta_1 f_\eta(q), \tag{5.70}$$

$$\tilde{f}_2(q) = \alpha_2 f_\alpha(q) + \beta_2 f_\beta(q) + \eta_2 f_\eta(q). \tag{5.71}$$

[1] When performing the calculations presented later in the chapter this expression has been divided by $(ag)^2$. This does not affect the message of the results as it simply rescales them.

We can transform $\tilde{f}_2$ to $m[\tilde{f}_2]$, and determine the values of the parameters α_2, β_2 and η_2 that optimally fit the noisy $m[\tilde{f}_1]$ by performing a least squares fit. The χ^2 function (eq. (5.53)) is then given by

$$\chi^2 = \vec{z} \cdot \mathrm{Cov}^{-1} \cdot \vec{z}, \tag{5.72}$$

where

$$z^{(i)} = \Delta\alpha m[f_\alpha](q^{(i)}) + \Delta\beta m[f_\beta](q^{(i)}) + \Delta\eta m[f_\eta](q^{(i)}) + \epsilon m[f_\epsilon](q^{(i)}) + \gamma m[f_\gamma](q^{(i)}) \tag{5.73}$$

$$\implies \vec{z} = \Delta\alpha\vec{h}_\alpha + \Delta\beta\vec{h}_\beta + \Delta\eta\vec{h}_\eta + \epsilon'\vec{h}_\epsilon + \gamma\vec{h}_\gamma, \tag{5.74}$$

where $z^{(i)}$ are the components of $\vec{z}$ and $\Delta\alpha = \alpha_1 - \alpha_2$, $\Delta\beta = \beta_1 - \beta_2$ and $\Delta\eta = \eta_1 - \eta_2$. The value α_2 that minimizes χ^2, α_2^*, is given by

$$\alpha_2^* = \alpha_1 + F(\Delta\beta, \Delta\eta, \gamma, \epsilon), \tag{5.75}$$

where F is some unknown function. Equivalent relations hold for the optimum parameters $\beta_2 = \beta_2^*$ and $\eta_2 = \eta_2^*$. Given this, we choose $\alpha_1 = \beta_1 = \eta_1 = 0$ for simplicity in all testing since alternative choices only act to shift best-fit parameters by a constant amount.

5.11 Bias

There are two sources of error in the parameter estimates obtained with this method. One is the statistical error due to the noise in the data, while the other is the bias due to incomplete cancellation of the q^2 divergence. This incomplete cancellation leaves an aq^4 term (eq. (5.40)), which impacts the estimates of parameter values in a way that does not vanish in the large statistics limit.

We now investigate the dependence of the bias on γ, the coefficient of the q^2 divergence, in the large statistics limit. In this limit, we can take ϵ to be small such that the $\epsilon\vec{h}_\epsilon$ term can be ignored in the calculation of χ^2. The fit results are found by minimizing χ^2,

$$\frac{\partial\chi^2(\Delta\alpha, \Delta\beta, \Delta\eta, \gamma)}{\partial\Delta\alpha} = 2\vec{h}_\alpha \cdot \mathrm{Cov}^{-1} \cdot (\Delta\alpha\vec{h}_\alpha + \Delta\beta\vec{h}_\beta + \Delta\eta\vec{h}_\eta - \gamma\vec{h}_\gamma) = 0, \tag{5.76}$$

To simplify notation let $A_{XY} = \vec{h}_X \cdot \mathrm{Cov}^{-1} \cdot \vec{h}_Y$. Because of the linearity of the dot product,

$$\begin{aligned}\frac{\partial\chi^2(\Delta\alpha, \Delta\beta, \Delta\eta, \gamma)}{\partial\Delta\alpha} =& \Delta\alpha A_{\alpha\alpha} + \Delta\beta A_{\alpha\beta} + \Delta\eta A_{\alpha\eta} - \gamma A_{\alpha\gamma} = 0 \\ \implies& \Delta\alpha A_{\alpha\alpha} + \Delta\beta A_{\alpha\beta} + \Delta\eta A_{\alpha\eta} = \gamma A_{\alpha\gamma},\end{aligned} \tag{5.77}$$

and similarly for β and η. Putting these equations together, we get

$$\mathcal{A} \cdot d = \gamma w, \tag{5.78}$$

where

$$\mathcal{A} = \begin{pmatrix} A_{\alpha\alpha} & A_{\alpha\beta} & A_{\alpha\eta} \\ A_{\beta\alpha} & A_{\beta\beta} & A_{\beta\eta} \\ A_{\eta\alpha} & A_{\eta\beta} & A_{\eta\eta} \end{pmatrix}, \qquad d = \begin{pmatrix} \Delta\alpha \\ \Delta\beta \\ \Delta\eta \end{pmatrix}, \qquad w = \begin{pmatrix} A_{\alpha\gamma} \\ A_{\beta\gamma} \\ A_{\gamma\gamma} \end{pmatrix}. \tag{5.79}$$

Rearranging we get $d = \gamma\mathcal{A}^{-1}w$, so $d \propto \gamma$. Therefore, we have the key result that, in the large statistics limit,

$$\langle \Delta\alpha \rangle = \gamma K_\alpha, \tag{5.80}$$

where K_α is a constant. Equivalent linear relations hold for $\Delta\beta$ and $\Delta\eta$. These constants do however depend on the cut $q_{\max}$ we apply to the data. We, therefore, try varying this cut to see how it influences the bias and statistical error.

5.12 Synthetic Data - Results

To make synthetic data, samples of $\epsilon X(q)$ were drawn for a $L = 256$ cubic lattice and added to momentum space data made using the anzatz in eq. (5.63). This was done for a range of γ and $q_{\max}$ values, and the resultant fits are shown in figure 5.1. The statistical error was calculated by performing a bootstrap resampling, where the covariance matrix was calculated on each sample. Figure 5.1 demonstrates a few things. Firstly, we see that the statistical error decays rapidly as the cut-off momentum $q_{\max}$ is increased, due to the increased number of points in the fit. Secondly, we see increased deviation from the correct value of $\alpha = 1$ with increasing $q_{\max}$. This is because the residual contribution from the γq^2 term after processing with the map m (eq. (5.45)) is $\sim q^4$ at leading order. It is relatively small when $|q|$ is small, but if $q_{\max}$ is sufficiently large then it will become significant. The deviation is much more severe for larger values of γ as expected. There is an accuracy-precision trade-off that one would need to make with unseen data when choosing $q_{\max}$: as $q_{\max}$ increases, the precision increases (up to a certain point), but the accuracy decreases. The optimum point in this trade-off will depend on the value of γ, with smaller values of γ having an optimum trade-off point with larger $q_{\max}$. Unfortunately, in our simulation data, the value of γ is unknown, so it is unclear how one can a priori choose $q_{\max}$.

We may hope that the inaccurate fit results will be associated with small p-values since they fit the wrong functional form to the data (the fit function does not contain a q^2 piece). To explore this idea, the ratio of the bias over the statistical error $r = |\Delta\alpha|/\sigma_{\Delta\alpha}$

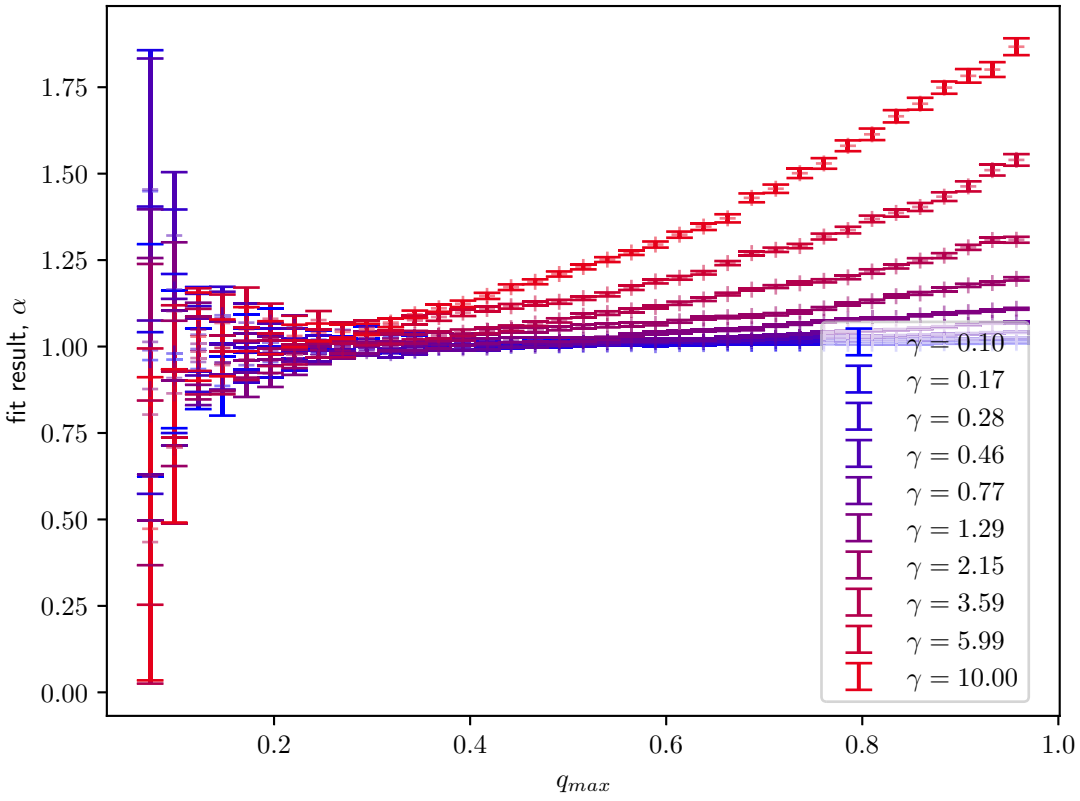

FIGURE 5.1: The fit result of the parameter α on synthetic data. Error bars are calculated using a statistical bootstrap. A range of different values of γ (the coefficient of the q^2-term) and $q_{\max}$ were used. Fitting was done on one dimension of momentum space data. The value of α used to create the data is 1.

is plotted against the p-value of the fit in figure 5.2. If there were no bias in the results, we would see distributions similar to $r \sim N(0, 1)$ for fits with acceptable p-values. Unfortunately, while the largest discrepancies are associated with unacceptable p-values, many fits with acceptable p-values have 4σ or higher inaccuracies. This appears to be true regardless of the value of γ.

Therefore, the method of combining Laplace and Fourier transforms to remove the q^2-dependence likely suffers from the risk of significant bias which is not easily detected. In the infinite-volume continuum, the Fourier transform of a Laplacian acting on a delta function at the origin has an exact q^2 form, while the Laplace transform of the same distribution also has a q^2 form but with a minus sign. In the continuum there is absolute cancellation when adding these transforms together, however, for lattice data, this is no longer true. The cancellation is incomplete with a aq^4 term remainder, which is a source of bias when we fit the transformed data. We therefore seek an alternative strategy that does not suffer from such biases. Such a method must remove the q^2 term completely. One promising method is the windowing method described in the next section.

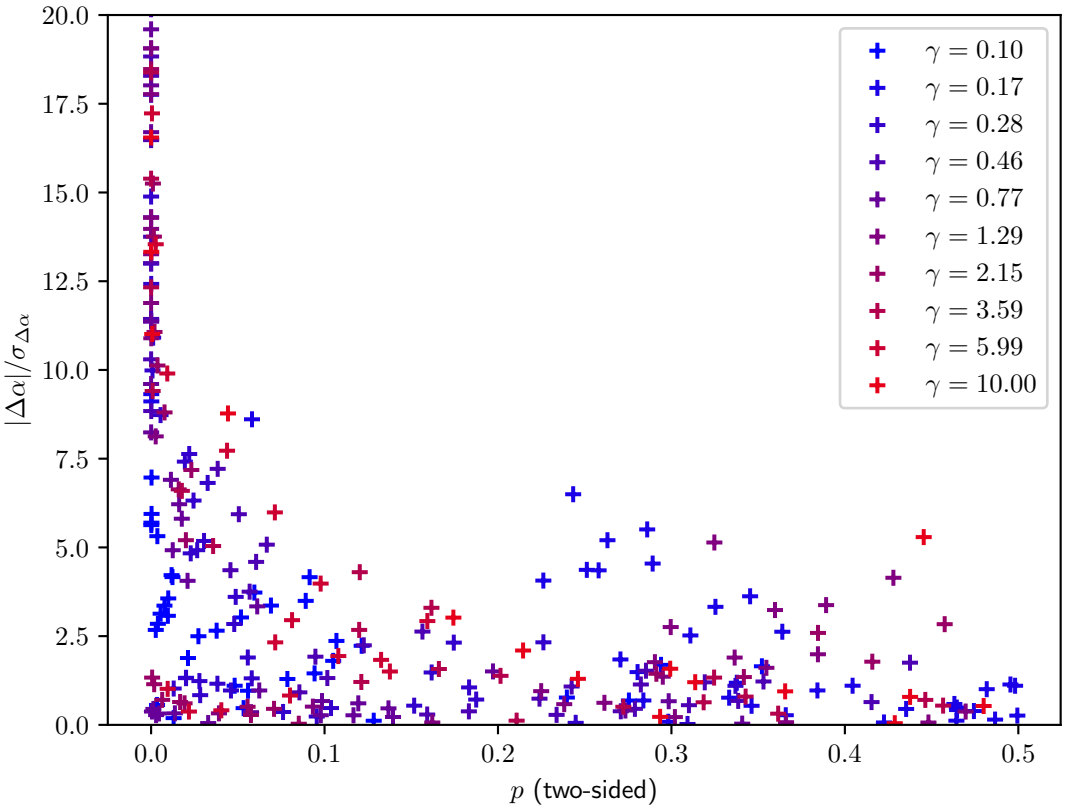

FIGURE 5.2: Relationship between the bias (in units of the statistical error) against the p-value of the fit for synthetic data fits. A range of different values of γ (the coefficient of the q^2-term) and $q_{\max}$ were used. Fitting was done on one dimension of the momentum space data.

5.13 The Windowing Method

The idea behind windowing is similar to the Fourier-Laplace method in that it uses a linear transformation correlator data to try to remove the q^2 divergence. Unlike the Fourier-Laplace method, which uses the cancellation between Fourier and Laplace transforms, the windowing method takes advantage of the localization of the q^2 divergence near the origin in position space. Given eq. (5.36), the function $f(x) = \mathcal{FT}^{-1}[r](x)$ (where $r(q) = q^2$) is zero for all sites at a distance greater than a from the origin. By multiplying our data by a function that is zero in a region around the origin, the q^2 divergence can be completely removed. We define the map

$$\tilde{m}[\tilde{f}] \coloneqq \mathcal{FT}[\mathcal{FT}^{-1}[\tilde{f}] \cdot \Gamma], \tag{5.81}$$

where $\cdot$ represents simple multiplication of two real-space functions and $\Gamma(x)$ is a function which is zero in the region $|x| \leq a$. The details of this method are discussed in detail in [108]. The function Γ is defined to be 0 for all values of x which are closer than a

certain distance r_0 to the origin, 1 for values of x further than $r_0 + \epsilon$ from the origin and smooth and monotonic in between. The specific choice of window affects the errors and convergence of the method. See [129] for a discussion of the Fourier properties of different windows. This technique has been shown to work in synthetic data studies, however, when applied to our lattice data the resulting signal is hard to detect amongst the noise [108].

To make further progress, we either require larger statistics (at the cost of more computer time) or methods to overcome or mitigate noise in our data. In the following chapter, we explore the feasibility of using a noise-reduction algorithm called multilevel for our holographic cosmology project.

Chapter 6

The Multilevel Algorithm

6.1 Introduction

In the CMB, lower multipoles result from quantum fluctuations at earlier times in the cosmic inflation process. It is in these low-multipole modes that we see a tension between the predictions of ΛCDM and observations of the CMB (see section (2.9)). In the dual theory, these low-multipole modes correspond to two-point functions in momentum space with a small value of the unitless ratio q/g, where g is the coupling constant, and q is the momentum.

In lattice simulations with a lattice spacing of a, the momenta take discrete values in the set $2\pi\tilde{n}a/L$, where $\tilde{n} \in \{0, 1, ..., L-1\}$. The ratio q/g is equal to $(aq)/(ag)$. Therefore, for a fixed ag, we are interested in two-point functions with small values of $a\tilde{n}/L$. These low momentum modes on the lattice are sensitive to long-distance correlations. This sensitivity poses a problem for accurate lattice estimation of these low-momentum modes as the two-point function of long-ranged correlators suffers from the exponential *signal-to-noise problem*. This is a well-known issue in lattice QFT (see for example [130]), which occurs when the noise in a lattice measurement is constant, but the signal decays like e^{-xm_0} at long distances x, where $m_0 = 1/\xi_0$ is the ground state mass of the system with associated correlation length ξ_0. The signal-to-noise ratio therefore decays like e^{-xm_0} at long distances.

Multilevel algorithms have proved to be successful at overcoming the signal-to-noise problem. The name derives from the multiple layers a lattice simulation is separated into. The lattice sites in layer n form boundaries around regions of the lattice simulated in layer $(n+1)$. The first multilevel algorithm was proposed in [131], and then extended in [132], where the term "multilevel" was coined.

A commonly used multilevel setup is one with two layers: ∂B, which we will henceforth refer to as the boundary, and Λ_r, the sub-lattice layers. Here $r \in \{1, 2, ..., n_s\}$ with n_s

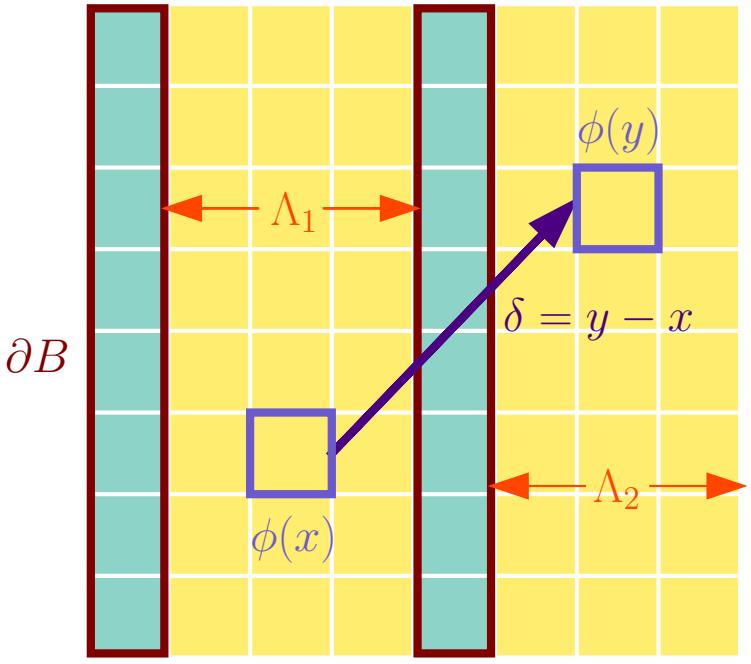

FIGURE 6.1: Two-level multilevel setup with sub-lattices Λ_1 and Λ_2 shown in yellow separated by a boundary layer ∂B shown in green [3].

being the number of sub-lattices. A system with $n_s = 2$ is shown in Figure 6.1 [3]. The multilevel algorithm is compatible with many boundary conditions commonly used in lattice simulations. In this project, we use periodic ('Pac-Man') boundaries unless otherwise stated, as this matches the implementation used in our simulation code. If non-periodic boundaries are used, then only one boundary layer is needed to separate the sub-lattices instead of two.

Consider the decomposition of the action,

$$S[\Lambda] = S_a[\partial B] + \sum_{r=1}^{n_s} S_{b_r}[\Lambda_r] + \sum_{r=1}^{n_s} S_{c_r}[\Lambda_r, \partial B] + \sum_{r=1}^{n_s} \sum_{s \neq r} S_{d_{rs}}[\Lambda_r, \Lambda_s] + S_{\text{other}}[\Lambda], \quad (6.1)$$

where S_a is the contribution to the action purely localized within the boundary, S_{b_r} is the contribution from the r^{th} sub-lattice, S_{c_r} captures the cross terms between the boundary and the r^{th} sub-lattice, $S_{d_{rs}}$ captures the contribution to the action from cross-terms between the r^{th} and s^{th} sub-lattices and $S_{\text{other}}[\Lambda]$ captures any cross-terms between three or more different sites on the lattice. In a model with a local action that involves terms between lattice sites at most r away and boundary layers with thickness $t \geq r$, interactions between different sub-lattices are not present, so $S_{d_{rs}}[\Lambda_r, \Lambda_s] = 0 \forall r, s$ and $S_{\text{other}}[\Lambda] = 0$. We assume from here on that we are always discussing a theory that is local in this way. The remaining terms in the action facilitate a useful decomposition

∂B, *boundary*	The lattice sites in the first layer of the multilevel algorithm
Λ_r, *sub-lattice(s)*	The sites in the second layer of the multilevel algorithm
r	Labels the sub-lattices, in the set $\{1, 2, ..., n_s\}$
n_s	The number of sub-lattices
configuration, ∂B_i	A given Monte-Carlo sample of the boundary
sub-configuration, $(\Lambda_r)_{ij_r}$	A given Monte-Carlo sample of the sub-lattice Λ_r
i	Indexes the configurations, in the set $\{1, 2, ..., n_0\}$
j_r	Indexes the sub-configurations of Λ_r, in the set $\{1, 2, ..., n_1\}$
n_0	Number of boundary configurations
n_1	Number of configurations of each sub-lattice
$\mathbf{x}$	Position on a two-dimensional lattice, (x, y)
$\phi_i(\mathbf{x})$, $\phi_{ij_r}(\mathbf{x})$	The value of a field variable at position $\mathbf{x}$ on the lattice. Indices depend on if $\mathbf{x}$ is in the boundary or in a sub-lattice
$\Phi_i(x), \Phi_{ij_1j_2}(x)$	Slice-coordinate field, equal to $\frac{1}{L}\sum_y \phi((x, y))$ with ϕ indexed appropriately.

TABLE 6.1: Notation used in this thesis for topics related to multilevel.

of the action,

$$\int_{x\in\Lambda} \mathcal{D}\phi(x) e^{-S[\Lambda]} = \int_{x\in\partial B} \mathcal{D}\phi(x) e^{-S_a[\partial B]} \prod_{r=1}^{n_s} \int_{x_r\in\Lambda_r} \mathcal{D}\phi(x_r) e^{-S_{f_r}[\Lambda_r, \partial B]}, \tag{6.2}$$

where $S_{f_r}[\Lambda_r, \partial B] = S_{b_r}[\Lambda_r] + S_{c_r}[\Lambda_r, \partial B]$. This decomposition demonstrates that the multilevel algorithm is an unbiased method. The path integral, eq. (6.2), is sampled in a two-step process by first sampling the distribution $e^{-S_a[\partial B]}$. For each of these samples, we sample the conditional distribution $e^{-S_{f_r}[\Lambda_r, \partial B]}$. To keep track of the notation used in this chapter the table (6.1) has been provided.

6.2 Background: Multilevel in Lattice Gauge Theories

The historical origins of the multilevel algorithm lie in lattice gauge theories. We note that the lattice gauge action (eq. (3.81)) is completely local in the link variables, as interaction terms are only present between link variables that share plaquettes. This locality was exploited in the 1980s [131] when considering Polyakov loops,

$$P(x, y, t) = Tr\left[\prod_{i=0}^{L_z-1} U_{\mu_3}(x, y, z + ia_z, t)\right], \tag{6.3}$$

where μ_3 denotes the z-axis, and L_z and a_z are the lattice size and lattice spacing in the z-direction respectively. Due to the periodicity of the lattice, these Polyakov loops are gauge invariant. If we consider the two-point correlator of two such loops separated by $(0, 0, t)$,

$$C_2(t; x, y) = \langle P(x, y, t_0) P(x, y, t + t_0)\rangle, \tag{6.4}$$

then assuming that $t \geq 2a_t$ (where a_t is the lattice spacing in the t-direction), the two entries in this expectation are statistical variables that depend only on the values of the other link variables in the lattice, and not directly on each other. If we call the set of these other link variables C and define a new average, where we keep C fixed,

$$C_2(t;x,y) = \langle \bar{P}(x,y,t_0)\bar{P}(x,y,t+t_0)\rangle, \tag{6.5}$$

where $\bar{X}$ is a lattice average of the quantity X under fixed C. The outer $\langle$ and $\rangle$ remain to indicate that we still take a Monte-Carlo average over different configurations of C. Because the Polyakov loops are straight lines, contiguous elements of the loop are also independent of each other, depending only on the other elements of their respective plaquettes that are in C. Therefore

$$\bar{P}(x,y,t_0) = Tr\left[\prod_{i=0}^{L_z-1} \bar{U}_{\mu_3}(x,y,z+ia_z,t)\right], \tag{6.6}$$

leading the authors of [131] to take heatbath averages of each of the links in the Polyakov loops individually by updating the Polyakov links multiple times while keeping C fixed. This technique is referred to as a *multihit* method.

The ideas in this founding paper were expanded on in [132]. Here the authors advocated splitting the lattice into time-slices of various thicknesses. The gauge links comprising a given time-slice from time t_0 to t_1 involve the gauge links with $\mu \in \{1,2,3\}$ (space-like) with $t \in \{t_0+a, t_0+2a, ..., t_1-a\}$, and the positive time-directed gauge links ($\mu = 0$) with $t \in \{t_0, t_0+a, ..., t_1-a\}$. We can define sub-lattice averages over time-slices, which we denote by

$$[X]_{t_0}^{t_1}. \tag{6.7}$$

Because of the locality of the action, this average only depends on the space-like gauge links at $t = t_0$ and $t = t_1$. For simulations, the lattice can be split into multiple time-slices, only keeping the space-like gauge links at the boundaries of these time-slices constant. This decomposition allows numerical averages over all time-slices to be found simultaneously. The problem considered in [132] involves products over two-link quantities

$$\mathbb{T}(t,r)_{\alpha\beta\gamma\delta} = ((U_{\mu_0})_{\alpha\beta}(t,0,0,0))^*(U_{\mu_0})_{\gamma\delta}(t,r,0,0), \tag{6.8}$$

where Greek indices are used to represent colors. The two-point correlator between Polyakov loops separated by a distance r is given by a product over $\mathbb{T}$ variables,

$$\begin{aligned} P(0,0,0)^*P(r,0,0) &= \langle(\mathbb{T}(0,r)\mathbb{T}(a,r)\cdots\mathbb{T}(T-a,r))_{\alpha\alpha\gamma\gamma}\rangle \\ &= \langle([\mathbb{T}(0,r)\mathbb{T}(a,r)]_0^{2a}[\mathbb{T}(2a,r)\mathbb{T}(3a,r)]_{2a}^{4a}\cdots[\mathbb{T}(T-2a,r)\mathbb{T}(T-a,r)]_{T-2a}^{T})_{\alpha\alpha\gamma\gamma}\rangle, \end{aligned} \tag{6.9}$$

where T is the time extent of the lattice. In the second line, the locality of the action has been used to decompose the product in a multilevel manner. This could be done using any number of layers of averaging, hence the name multilevel. The authors suggest tuning the number of sub-lattice updates so that the signal-to-noise ratio is approximately unity ($n_1 \propto e^{2m_0 r}$, given that each of the sub-lattice bracketed terms should be scale like $e^{-m_0 r}$, while the noise is roughly constant). Given that there are $T/2a_t$ of these brackets, we get a statistical gain proportional to $e^{-m_0 rT/2}$ for a cost of only $e^{2m_0 r}$. This leads to the central claim of [132] that this technique can achieve exponential error reduction and thus has the potential to solve the signal-to-noise problem.

This exponential error reduction, however, relies on the fact that these Polyakov loop correlators decompose into a product of many variables. In the case of a simple two-point function, there is only one product over variables that can be decomposed, so the gains are not exponential. This two-point case will be discussed more later on in the chapter.

Since these foundational papers, the multilevel algorithm has been used in a wide range of lattice calculations, with both [131] and [132] cited by several hundred papers. Predominantly these papers are in the field of lattice gauge theory, where the aforementioned exponential error reduction strategy is well understood. In [133], the authors outline how an exponential error reduction scheme could be achieved for more two-point functions by utilizing symmetries of the action. There is also ongoing research looking at multilevel schemes in QCD with fermions, where the action is only approximately local, so thicker boundaries and approximate strategies must be used [134]. Other research has investigated the potential of multilevel in overcoming the critical slowing down problem [135]. More recently, multilevel methods have been used in studies of the muon magnetic moment anomaly [136].

6.3 A Statistical View of Multilevel

In the following, we will assume that each configuration of the boundary, ∂B_i, is independent and identically distributed (*i. i. d.*). Similarly, for a *given* boundary configuration ∂B_{i_0}, we assume that sub-configurations of the sub-lattice Λ_r with $i = i_0$ are likewise *i. i. d.*, with a distribution that is dependent on ∂B_{i_0}. These assumptions are valid as long as our Monte-Carlo samples are separated by a Monte-Carlo time, $\Delta T >> \tau_{\text{int}}$, where τ_{int} is the relevant integrated autocorrelation time. To derive the scaling (with n_0 and n_1) of two-point function errors in a multilevel algorithm, we consider two field insertions $\phi_{ij_1}(\mathbf{x})$ and $\phi_{ij_2}(\mathbf{y})$, where $\mathbf{x} \in \Lambda_1$ and $\mathbf{y} \in \Lambda_2$. These variables are samples of

the distributions $D_{\mathbf{x},\partial B_i}$ and $D_{\mathbf{y},\partial B_i}$ respectively. We now consider the variables

$$X_i = \frac{1}{n_1}\sum_{j_1=1}^{n_1}\phi_{ij_1}(\mathbf{x}), \qquad Y_i = \frac{1}{n_1}\sum_{j_2=1}^{n_1}\phi_{ij_2}(\mathbf{y}). \tag{6.10}$$

By applying the central limit theorem, in the limit $n_1 \to \infty$ we obtain the distribution of these variables as

$$X_i \sim N\left(\mu_{\mathbf{x}}(\partial B_i), \frac{\sigma^2_{\mathbf{x}}(\partial B_i)}{n_1}\right), \qquad Y_i \sim N\left(\mu_{\mathbf{y}}(\partial B_i), \frac{\sigma^2_{\mathbf{y}}(\partial B_i)}{n_1}\right), \tag{6.11}$$

where $\mu_{\mathbf{x}}(\partial B_i)$ and $\sigma^2_{\mathbf{x}}(\partial B_i)$ are the means and standard deviations respectively of the distributions $D_{\mathbf{x},\partial B_i}$. We introduce the two-point function variable $Z_i = X_i Y_i$. Since X_i and Y_i depend only on the boundary configuration, there is no residual correlation between these variables, so they can be considered statistically independent. This is not to say that the lattice variables $\phi(\mathbf{x})$ and $\phi(\mathbf{y})$ are not correlated, just that all of the correlation have already been accounted for in the values of $\mu_{\mathbf{x}}(\partial B_i)$ and $\mu_{\mathbf{y}}(\partial B_i)$. If we take repeat samples of Z_i with fixed ∂B_i, the samples have a mean

$$\mu_{Z_i}(\partial B_i) = \mu_{Y_i}(\partial B_i)\mu_{X_i}(\partial B_i), \tag{6.12}$$

where $\mu_{X_i} = \mu_{\mathbf{x}}(\partial B_i)$ and $\mu_{Y_i} = \mu_{\mathbf{y}}(\partial B_i)$. To find the variance of Z_i takes a little more work,

$$\begin{aligned}\sigma^2_{Z_i}(\partial B_i) &= \langle [Z_i - \langle Z_i\rangle]^2\rangle \\ &= \langle [X_iY_i - \langle X_i\rangle\langle Y_i\rangle]^2\rangle.\end{aligned} \tag{6.13}$$

Reorganizing the terms we have,

$$\begin{aligned}\sigma^2_{Z_i}(\partial B_i) &= \langle [(X_i - \langle X_i\rangle)(Y_i - \langle Y_i\rangle) - 2\langle X_i\rangle\langle Y_i\rangle + X_i\langle Y_i\rangle + Y_i\langle X_i\rangle]^2\rangle \\ &= \langle [(X_i - \langle X_i\rangle)(Y_i - \langle Y_i\rangle) + \langle Y_i\rangle(X_i - \langle X_i\rangle) + \langle X_i\rangle(Y_i - \langle Y_i\rangle)]^2\rangle.\end{aligned} \tag{6.14}$$

Expanding the bracket, and noting that because of statistical independence of expectations of products, $X_i - \langle X_i\rangle$ and $Y_i - \langle Y_i\rangle$ separate into products of their individual expectations (which are zero), we get

$$\begin{aligned}\sigma^2_{Z_i}(\partial B_i) &= \langle (X_i - \langle X_i\rangle)^2(Y_i - \langle Y_i\rangle)^2\rangle + \langle \langle Y_i\rangle^2(X_i - \langle X_i\rangle)^2\rangle + \langle \langle X_i\rangle^2(Y_i - \langle Y_i\rangle)^2\rangle \\ &\quad + 2\langle \langle Y_i\rangle(X_i - \langle X_i\rangle)^2(Y_i - \langle Y_i\rangle)\rangle + 2\langle \langle X_i\rangle(X_i - \langle X_i\rangle)(Y_i - \langle Y_i\rangle)^2\rangle \\ &\quad + 2\langle \langle X_i\rangle\langle Y_i\rangle(X_i - \langle X_i\rangle)(Y_i - \langle Y_i\rangle)\rangle \\ &= \langle (X_i - \langle X_i\rangle)^2(Y_i - \langle Y_i\rangle)^2\rangle + \langle \langle Y_i\rangle^2(X_i - \langle X_i\rangle)^2\rangle + \langle \langle X_i\rangle^2(Y_i - \langle Y_i\rangle)^2\rangle \\ &= \frac{1}{n_1^2}\sigma^2_{\mathbf{x}}(\partial B_i)\sigma^2_{\mathbf{y}}(\partial B_i) + \frac{1}{n_1}[\mu^2_{\mathbf{x}}(\partial B_i)\sigma^2_{\mathbf{y}}(\partial B_i) + \mu^2_{\mathbf{y}}(\partial B_i)\sigma^2_{\mathbf{x}}(\partial B_i)].\end{aligned} \tag{6.15}$$

The total variance of Z_i gains another contribution when accounting for resampling of ∂B_i since this affects the mean of Z_i. The total variance of a variable conditional on another variable is given by the law of total variance. In this case tells us that

$$\mathrm{Var}(Z) = \mathrm{Var}(\mu_Z) + \mathrm{E}(\sigma_Z^2), \tag{6.16}$$

where E and Var correspond to the expectation value and variance as we change ∂B. Substituting in our previous results gives

$$\mathrm{Var}(Z) = \mathrm{Var}(\mu_{\mathrm{x}}\mu_{\mathrm{y}}) + \mathrm{E}\left(\frac{1}{n_1^2}\sigma_{\mathrm{x}}^2\sigma_{\mathrm{y}}^2 + \frac{1}{n_1}(\mu_{\mathrm{x}}^2\sigma_{\mathrm{y}}^2 + \mu_{\mathrm{y}}^2\sigma_{\mathrm{x}}^2)\right), \tag{6.17}$$

where for convenience we have dropped the reference ∂B in the notation for μ and σ. Finally we define the two-point function variable $\tilde{Z} = (1/n_0)\sum_i Z_i$ and get (see also [137])

$$\mathrm{Var}(\tilde{Z}) = \frac{1}{n_0}A + \frac{1}{n_0 n_1}B + \frac{1}{n_0 n_1^2}C, \tag{6.18}$$

where

$$\begin{aligned} A &= \mathrm{Var}(\mu_{\mathrm{x}}\mu_{\mathrm{y}}), \\ B &= \mathrm{E}(\mu_{\mathrm{x}}^2\sigma_{\mathrm{y}}^2 + \mu_{\mathrm{y}}^2\sigma_{\mathrm{x}}^2), \\ C &= \mathrm{E}(\sigma_{\mathrm{x}}^2\sigma_{\mathrm{y}}^2). \end{aligned} \tag{6.19}$$

The relative size of the terms in this sum will determine if multilevel offers an improvement over a computationally equivalent single-level algorithm. When we run a multilevel simulation, we simulate the sub-lattices simultaneously, so the number of Monte-Carlo steps is proportional to $n_0 \times n_1$. In a single-level simulation with $n_0 \times n_1$ configurations, we would find that the variance of our two-point correlators scales like $\frac{1}{n_0 n_1}$. In the best case scenario, the $\frac{1}{n_0 n_1^2}C$ term is dominant, and multilevel gives us a n_1 reduction in the variance of our correlators. However, if the $\frac{1}{n_0}A$ term dominates the error, then two-point correlators calculated with the multilevel algorithm will have a variance n_1 times larger than an equivalent single-level algorithm. The B term reveals that the ideal $1/n_0 n_1^2$ scaling can only be achieved when the quantity of interest has an expectation value close to zero. This can usually be achieved for two-point functions by subtracting the disconnected part of the correlator.

6.4 The Ising Model

In our holographic model, the dual quantum field theory is simulated at the massless (critical) point. Multilevel is usually applied in systems with a finite correlation length, such that one can create correlators between points sufficiently far into the sub-lattice

that they are not dominated by the boundaries. As we move towards a critical regime, this no longer holds. To investigate whether the multilevel algorithm could still improve the statistics of our two-point correlators in this regime we have conducted a study in the Ising model [138]. The two-dimensional Ising Model is one of the simplest systems that exhibits a second-order phase transition between a symmetric phase and a broken phase. The system has the path integral

$$Z = \int \mathcal{D}\phi \exp\left(-\beta\left(J \sum_{(i,j)_{\mathrm{NN}}} \phi_i \phi_j + B \sum_i \phi_i\right)\right), \tag{6.20}$$

where B is due to an external magnetic field β is the inverse of the temperature T and NN means nearest neighbors. In the case of a system with two dimensions and $B = 0$, the critical point is at $\beta_c = \frac{1}{2}\log(1+\sqrt{2})$ ([139]). By varying the parameter β in the approach to the critical point, the performance of the multilevel algorithm across a range of correlation lengths was investigated. This was done in Python code, organized in a Jupyter-Notebook [8]. This code also reproduces all the results presented in this chapter including plots. The purpose of this project was to investigate the performance of the multilevel algorithm in the approach to criticality, and not to write a competitive simulation code for the Ising Model. The code presented could be improved by using clustering techniques such as those in [140], through parallelization and by transcribing into a compiled language.

6.5 Slice-Fields

To reduce the numerical and analytical complexity of the system, we will use *slice-fields* from now on. These are fields that have been averaged over all but one direction of the lattice,

$$\Phi(x) = \frac{1}{L^{d-1}} \sum_{y_1,\ldots,y_{d-1}} \phi((y_1,\ldots,y_{d-1},x)), \tag{6.21}$$

where d is the number of dimensions of the system. These fields are also often referred to as zero-momentum projected fields. We define the slices to be parallel to the boundaries, so a given slice is wholly in the boundary layers or wholly in a sub-lattice. We can relate two-point functions of slice-fields, $c_2(x)$ represented by lower case 'c' to two-point functions of unaveraged fields, $C_2(\mathbf{x})$ represented by upper case 'C' by

$$c_2(x) = \sum_p e^{+ipx} \tilde{C}_2(0,0,\ldots,0,p) = C_2(0,0,\ldots,0,x), \tag{6.22}$$

where the $\tilde{C}_2$ is the momentum space two-point correlator, while $C_2(0,0,\ldots,0,x)$ a mixed correlator with the first $d-1$ entries in momentum space and the final entry in position space. This result is derived in the appendix (D). In section (2.1) of [67], the leading

behavior of these functions is shown to decay at leading order like $e^{-m|x|}$ assuming that $C_2(\mathbf{x}) \sim e^{-m|\mathbf{x}|}$ at leading order. From now on, we consider slice-fields in two dimensions,

$$\Phi_i(x) = \frac{1}{L}\sum_y \phi_i((x,y)), \tag{6.23}$$

$$\Phi_{ij_r}(x) = \frac{1}{L}\sum_y \phi_{ij_r}((x,y)), \tag{6.24}$$

where the indices depend on if the fields are in the boundary or the sub-lattices.

The scaling of the multilevel algorithm, eq. (6.19) holds for slice coordinates. In this system, we can calculate the two-point correlator for a given slice

$$C_2(\delta;x) = \frac{1}{n_0}\sum_{i=1}^{n_0}\sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_1}\Phi_{ij_1j_2}(x+\delta)\Phi_{ij_1j_2}(x), \tag{6.25}$$

where

$$\Phi_{ij_1j_2}(x) = \begin{cases} \Phi_i(x) & x \in \partial B, \\ \Phi_{ij_1}(x) & x \in \Lambda_1, \\ \Phi_{ij_2}(x) & x \in \Lambda_2. \end{cases} \tag{6.26}$$

In the theoretical model of multilevel performance, it will be important to treat the slice-fields as normally distributed. Given that they are an average over binary distributed variables, this is clearly an assumption, however figure 6.2, which shows the histograms of slice-fields at various values of β, demonstrates that this approximation is reasonable in the symmetric phase as long as β is not too close to critical.

6.6 The Covariance Matrix of Correlators

In this project, the performance of multilevel with ensembles of n_0 boundary configurations and n_1 sub-lattice configurations was compared to single-level performance with n_0n_1 configurations. This comparison was chosen as these two ensembles demand equal amounts of computational cost to produce. The two-point correlators of the single-level and multilevel data sets are defined by

$$C_2^s(\delta;x) = \frac{1}{n_0n_1}\sum_{i=1}^{n_0n_1}\Phi_i(x)\Phi_i(x+\delta) \tag{6.27}$$

and

$$C_2^m(\delta;x) = \frac{1}{n_0n_1^2}\sum_{i=1}^{n_0}\sum_{j_1=1}^{n_1}\sum_{j_2=1}^{n_1}\Phi_{ij_1j_2}(x)\Phi_{ij_1j_2}(x+\delta) \tag{6.28}$$

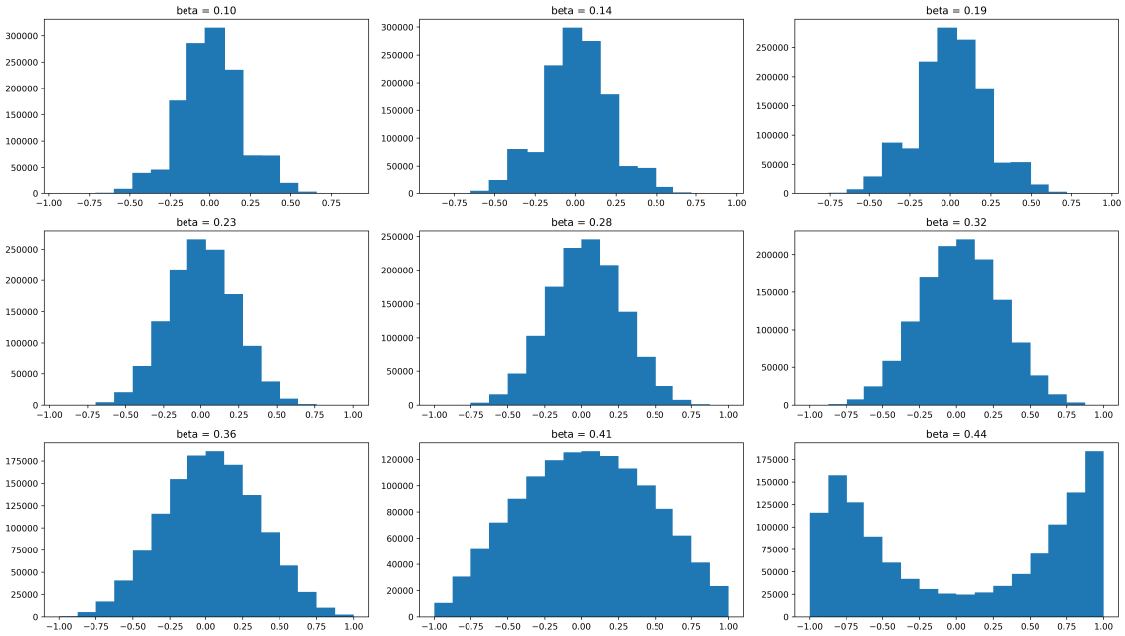

FIGURE 6.2: Histogram of slice-field values over a range of β in a single-level simulation of the 2D-Ising Model with lattice size $L = 32$ and 25000 configurations.

respectively, where we apply periodic boundary conditions: $\Phi_i(x+\delta) = \Phi_i(x+\delta-L)$, if $x+\delta \geq L$. Assuming that n_0 is sufficiently large, we apply the central limit theorem to give

$$C_2^s(\delta;x) \sim N\left(C_2(\delta), \sigma_s^2(x)\right), \tag{6.29}$$

$$C_2^m(\delta;x) \sim N\left(C_2(\delta), \sigma_m^2(x)\right), \tag{6.30}$$

where the expectation value of the multilevel correlator is the same as for the single-level correlator, because the 2D-Ising model only has interactions between neighboring lattice sites, so the decomposition of the path integral is exact. We make a final estimate of the two-point correlator by taking a weighted average of the $C_2(\delta,x)$ over x

$$C_2^s(\delta) = \sum_x W_x^s C_2^s(\delta;x), \tag{6.31}$$

$$C_2^m(\delta) = \sum_x W_x^m C_2^m(\delta;x),$$

where $\sum_x W_x^s = \sum_x W_x^m = 1$. We define weight vectors

$$\mathbf{W}^m = (W_1^m, W_2^m, ..., W_L^m), \tag{6.32}$$

$$\mathbf{W}^s = (W_1^s, W_2^s, ..., W_L^s). \tag{6.33}$$

Using the fact that the individual contributions of the weighted average are normal, the overall estimates of $C_2(\delta)$ are also normally distributed, and have variances of $\mathbf{W}^s \cdot \mathrm{Cov}^s(\delta) \cdot \mathbf{W}^s$ and $\mathbf{W}^m \cdot \mathrm{Cov}^m(\delta) \cdot \mathbf{W}^m$ respectively. Here we have defined $L \times L$ Covariance

matrices with components

$$\begin{aligned} \mathrm{Cov}^m(\delta)_{x_1x_2} &= \langle C_2^m(\delta;x_1)C_2^m(\delta;x_2)\rangle - \langle C_2^m(\delta;x_1)\rangle\langle C_2^m(\delta;x_2)\rangle, \\ \mathrm{Cov}^s(\delta)_{x_1x_2} &= \langle C_2^s(\delta;x_1)C_2^s(\delta;x_2)\rangle - \langle C_2^s(\delta;x_1)\rangle\langle C_2^s(\delta;x_2)\rangle. \end{aligned} \tag{6.34}$$

An example covariance matrix for a multilevel correlator is shown in figure (6.3). There are several important features to point out in this covariance matrix. Firstly, since the lattice uses periodic boundaries, we have $C_2^m(\delta;x_1) = C_2^m(\delta;x_1+L/2)$ for $x_1 < L/2$, where L is the lattice size. The covariance matrix is therefore singular, and we must be careful in our minimization procedure. To overcome this, we solve a reduced minimization problem where we minimize $W^{m\prime} \cdot \mathrm{Cov}^m(\delta) \cdot W^{m\prime}$, where $W^{m\prime}(x) = W^m(x)$ for $x < L/2$ and 0 otherwise. However, this is only necessary for the $\delta = L/2$ correlator.

The second thing worth noting about this correlator is that the covariance at $(0,0)$ is very high. This is to be expected since this is the boundary-to-boundary correlator. This correlator is fixed across all sub-lattice updates, so only receives a single contribution for each boundary configuration. The covariance of correlators near the boundaries is also higher, due to their correlation to the boundary. This leads to a contribution to the covariance similar to the A/n_0 term in (6.19). As we move away from the boundary the covariance of the correlators decreases, and so the correlators in the center of the sub-lattices will get the highest weights.

Theorem 6.1. *For a general $L \times L$ covariance matrix, C, and an L-weight vector $\mathbf{w}$ subject to $\sum_x w_x = 1$, the quadratic form $R = \mathbf{w} \cdot C \cdot \mathbf{w}$, is minimized by*

$$w_x = \frac{\sum_y C^{-1}_{xy}}{\sum_{xy} C^{-1}_{xy}}. \tag{6.35}$$

Proof. Use the method of Lagrange multipliers, minimize $\mathcal{L}$ with respect to w_x and λ.

$$\begin{aligned} \mathcal{L} &= \mathbf{w} \cdot C \cdot \mathbf{w} - \lambda\left(\sum_x w_x - 1\right) \\ \frac{\partial \mathcal{L}}{\partial \mathbf{w}} = 0 &\implies C \cdot \mathbf{w} + \lambda \mathbf{1} = 0, \end{aligned} \tag{6.36}$$

where $\mathbf{1} = 1 \forall x$. Multiplying through by C^{-1} to the left gives

$$\begin{aligned} \mathbf{w} &= -\lambda C^{-1}\mathbf{1}, \\ \implies w_x &= -\lambda \sum_y C^{-1}_{xy}. \end{aligned} \tag{6.37}$$

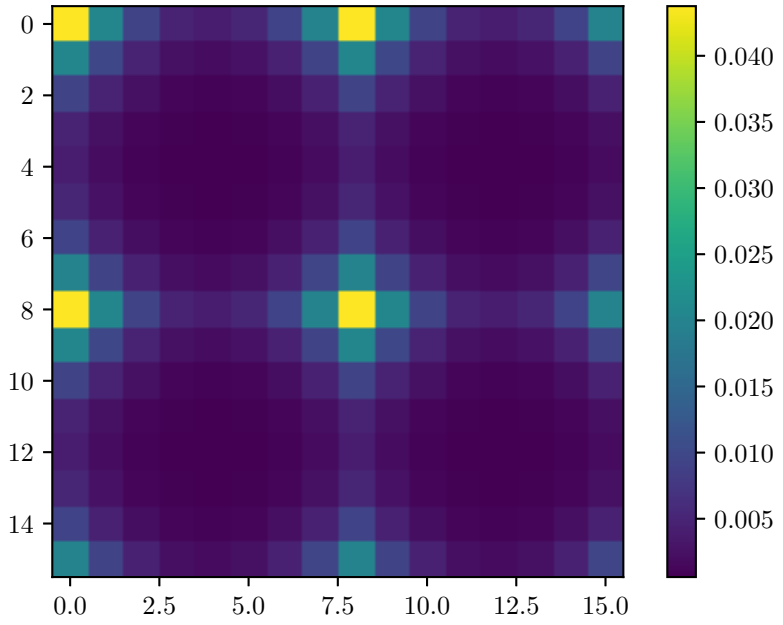

FIGURE 6.3: Covariance matrix of correlators of the two-point function of slice-fields for a 16×16 multilevel setup with two sub-lattices, 50 configurations, and 50 sub-lattice configurations per configuration. The correlator is of length $\delta = 8$ and the inverse temperature of the system was $\beta = 0.35$.

Taking the derivative $\partial\mathcal{L}/\partial\lambda = 0$ gives $\sum_x w_X = 1$, so

$$\lambda = \frac{1}{\sum_{xy} C^{-1}_{xy}}. \tag{6.38}$$

Substituting back into (6.37) gives

$$w_x = \frac{\sum_y C^{-1}_{xy}}{\sum_{xy} C^{-1}_{xy}}, \tag{6.39}$$

as required. □

Corollary 6.2. *The optimum weights we obtain are the same as if we performed a χ^2-minimization to a fit of a constant, μ, to the full correlator data.*

Proof. The χ^2 function is given by

$$\chi^2 = \sum_{xy} (Y_x - \mu) C^{-1}_{xy} (Y_y - \mu), \tag{6.40}$$

where $Y_x = C^m(\delta; x)$. Minimizing with respect to μ and using the symmetry of the covariance matrix gives

$$\frac{d\chi^2}{d\mu} = 0 \implies \sum_{xy} C_{xy}^{-1}(Y_y - \mu) = 0 \tag{6.41}$$

$$\implies \mu = \frac{\sum_{xy} C_{xy}^{-1} Y_y}{\sum_{xy} C_{xy}^{-1}} = \sum_x w_y Y_y, \tag{6.42}$$

where

$$w_y = \frac{\sum_x C_{xy}^{-1}}{\sum_{xy} C_{xy}^{-1}}. \tag{6.43}$$

Noting that the covariance matrix is symmetric $C_{xy} = C_{yx}$, and redefining $x \leftrightarrow y$ we get

$$w_x = \frac{\sum_y C_{xy}^{-1}}{\sum_{xy} C_{xy}^{-1}}. \tag{6.44}$$

□

6.7 Alternative Weighting Schemes

To demonstrate the need for an optimal weighting scheme, we have also calculated numerical results for alternative weighting schemes. The first and most basic scheme equally weights all contributions to the correlator. This scheme is appropriate for the single-level algorithm as all correlator contributions have the same expected variance. The second scheme we propose as a comparison is a basic weighting scheme which is as follows:

- Correlators between two boundary slice-fields have a weight of 1,
- Correlators between a boundary and a non-boundary slice-field are weighted at M,
- Correlators between two slice-fields in the same sub-lattice have a weight of M,
- Correlators between slice-fields in two different sub-lattices have a weight of M^2,
- These weights are finally normalized with $\sum_x W_x^m = 1$.

The motivation for this scheme is to weight correlators by the number of unique samples that contribute to that correlator per configuration. This seems like a logical thing to do, however, it ignores the fact that slice-fields in the sub-lattices are correlated to the boundary.

6.8 Numerical Results

A 32×32 Ising Model was simulated across a range of temperatures in the symmetric phase of the theory using a single-level and a multilevel algorithm. For the multilevel algorithm, $n_0 = 500$ configurations were generated, with $n_1 = 500$ sub-configurations per configuration, while for the single-level algorithm $25,000$ configurations were generated. This was done for 32 values of the inverse temperature β so that a range of correlation lengths of the system were explored. A fit was done on the single-level ensemble to determine the correlation length of the system, using the anzatz

$$C_2^s(\delta) = \frac{1}{A}\left[\exp\left(-\frac{\delta}{\xi}\right) + \exp\left(-\frac{(L-\delta)}{\xi}\right)\right], \tag{6.45}$$

where $A = 1+\exp(-L/\xi)$. It would be equally valid to fit the multilevel ensemble data as multilevel is not systematically biased compared to single-level (as discussed previously). To demonstrate that this is the case, the ratio

$$R = \frac{C_2^m(\delta) - C_2^s(\delta)}{\sigma_r(\delta)} \tag{6.46}$$

was calculated, where $\sigma_r(\delta) = \sqrt{((\sigma^s(\delta))^2 + (\sigma^m(\delta))^2)}$ with σ^m and σ^s being the standard deviation of the multilevel and single-level two-point functions respectively. The standard deviation in both cases was calculated under a bootstrap resampling of the configurations. For multilevel, this means resampling the boundary configurations, not the sub-lattice configurations. The values of R across all δ and β studied for $L = 32$ are shown in figure 6.4. If the two methods agree, we expect a distribution of R that follows $N(0,1)$. Note that, for the same value of β, the two-point function for different separations δ are correlated, so these do not represent independent samples of $N(0,1)$.

As an example, the two-point function for $\beta = 0.32$ is shown in figure 6.5. This graph demonstrates both the accuracy of the multilevel technique as well as its benefit in reducing the statistical uncertainty of long-range correlators of the system.

The performance improvement of the multilevel algorithm compared to the single-level one was plotted against the values of ξ from this fit for the three weighting schemes thus far discussed. This performance improvement was measured by the ratio σ_m/σ_s. The results are shown in figure 6.6.

Points to note are:

- For $\delta = 0$, it is impossible for multilevel to beat the performance of single-level. This is because there are no correlators that mix the two sub-lattices, and so there are no $1/NM^2$ scaling terms in the variance.

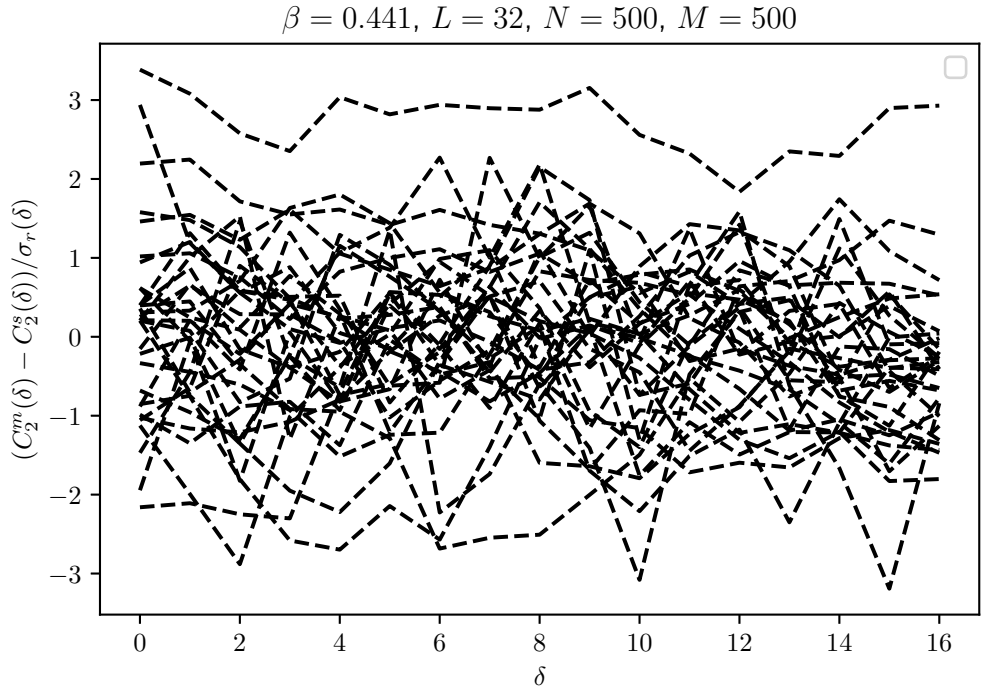

FIGURE 6.4: Plot of the ratio between the deviation of the multilevel and single-level results and their combined standard deviations for a range of β from 0.1 to $\log((1+\sqrt{2})/2)$.

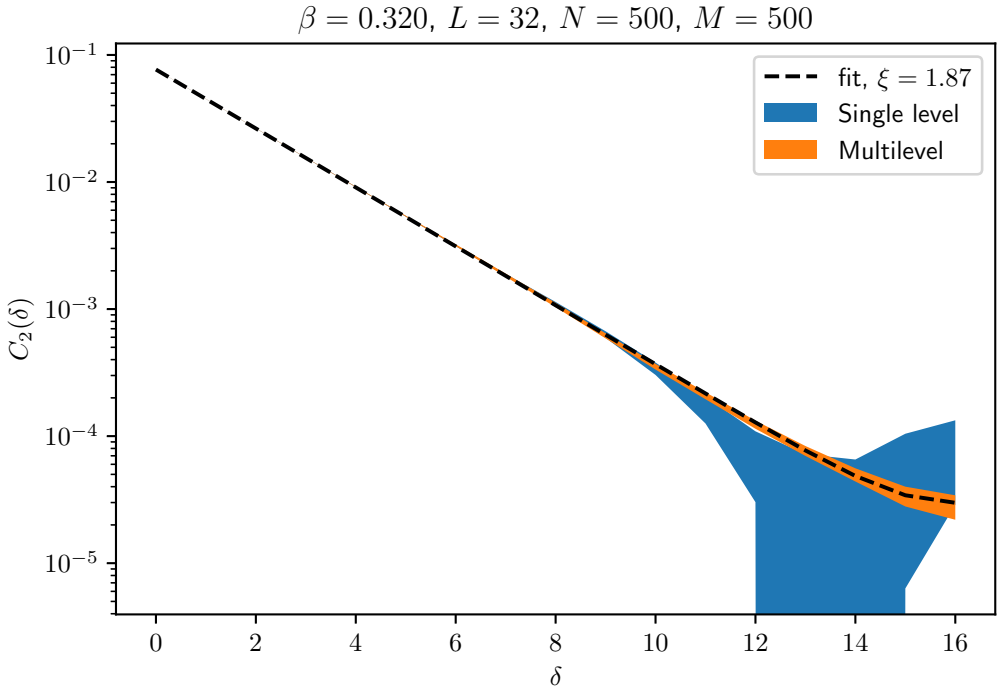

FIGURE 6.5: Two-point correlator for a 32×32 Ising model under computation-time equivalent multilevel and single-level algorithms at an inverse temperature of $\beta = 0.32$. Also shown is the fit result used to calculate the correlation length of the system, ξ.

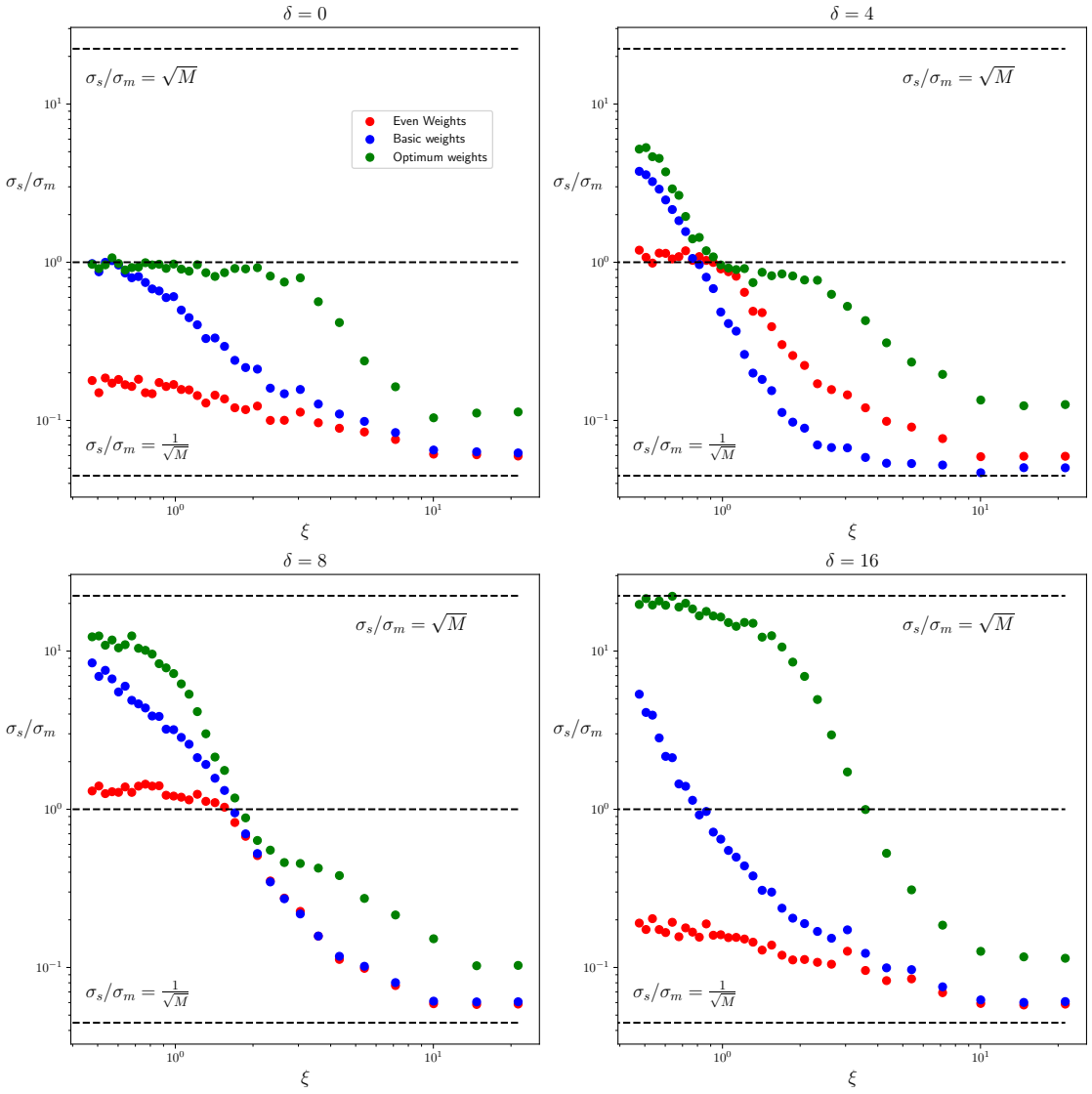

FIGURE 6.6: Ratio of the standard deviation of two-point correlator estimates for a multilevel scheme compared to a computationally equivalent single-level scheme. This was done on a 32×32 lattice with $N = 500$, $M = 500$ with two sub-lattices for the multilevel ensemble, and $N = 25000$ for the single-level one.

- As the correlation length is increased, the performance of the multilevel algorithm as compared to the single-level one decreases. This is a result of the sub-lattice sites becoming increasingly determined by the boundary, and therefore increasing the coefficient of the $1/N$ scaling term in the variance.

- The optimum weighting scheme always performs better or as well as the simpler alternative schemes.

- In the limit $\xi \to 0$, the performance of the basic weighting scheme approaches that of the optimum weighting scheme. This is due to the flaw in the basic weighting scheme (the failure to recognize that slice-fields in the sub-lattices are correlated to the boundary) disappears in this limit.

- The performance of the multilevel algorithm as compared to the single-level one improves as the correlator length, δ, increases. The first reason is that the longer the correlator the more contributions there are connecting slice fields in different sub-lattices. Specifically, there are $2(\delta - 1)$ such contributions for $\delta \geq 1$. Other factors aside, this means the best possible performance of multilevel in the limit $M \to \infty$ is reduced by a factor of $\sqrt{2(\delta - 1)/L}$. The second factor which inhibits the performance of multilevel for shorter correlators is that for correlators mixing the two sub-regions, the average distance of the slice-fields from the boundary is $\delta/2$. Therefore the $1/NM^2$ contributions present for these shorter correlators suffer from larger contributions from terms with poorer scaling.

6.9 Scaling

All the arguments used in the previous section to explain the numerical performance of multilevel only rely on the unitless ratios $r_1 = \delta/L$ and $r_2 = \xi/L$. We can therefore hypothesize that if we change L but keep these unitless ratios the same, we will see no change in the performance of multilevel. This is exactly what figure 6.7 shows. One could alternatively consider different unitless ratios of ξ, δ and L, and the conclusion would remain the same. Often in other lattice studies $L \gg \xi$, so the ratios δ/ξ and δ/L are a more natural way to study algorithm performance.

6.10 Theoretical Model

Given the remarkable scaling property of multilevel performance demonstrated in the previous section, it seems hopeful that these curves could be explained by a theoretical framework. To do this the theoretical framework needs to be able to estimate the covariance matrices $\mathrm{Cov}^s(\delta)$ and $\mathrm{Cov}^m(\delta)$ defined earlier. Underlying this model is the assumption that the slice-fields are roughly normally distributed with a mean of 0. This assumption was discussed in section 6.5.

6.10.1 Single Level Model

We will use the shorthand notation $\phi_i^R = \phi_i(x = R)$. We wish to consider the covariance between two correlators, so we choose $S' = S+\delta$ and $T' = T+\delta$ and consider the correlators

$$C_2^{SS'} = \frac{1}{N} \sum_{i=1}^{N_{\text{single}}} \phi_i^S \phi_i^{S'}, \qquad C_2^{TT'} = \frac{1}{N} \sum_{i=1}^{N_{\text{single}}} \phi_i^T \phi_i^{T'}. \tag{6.47}$$

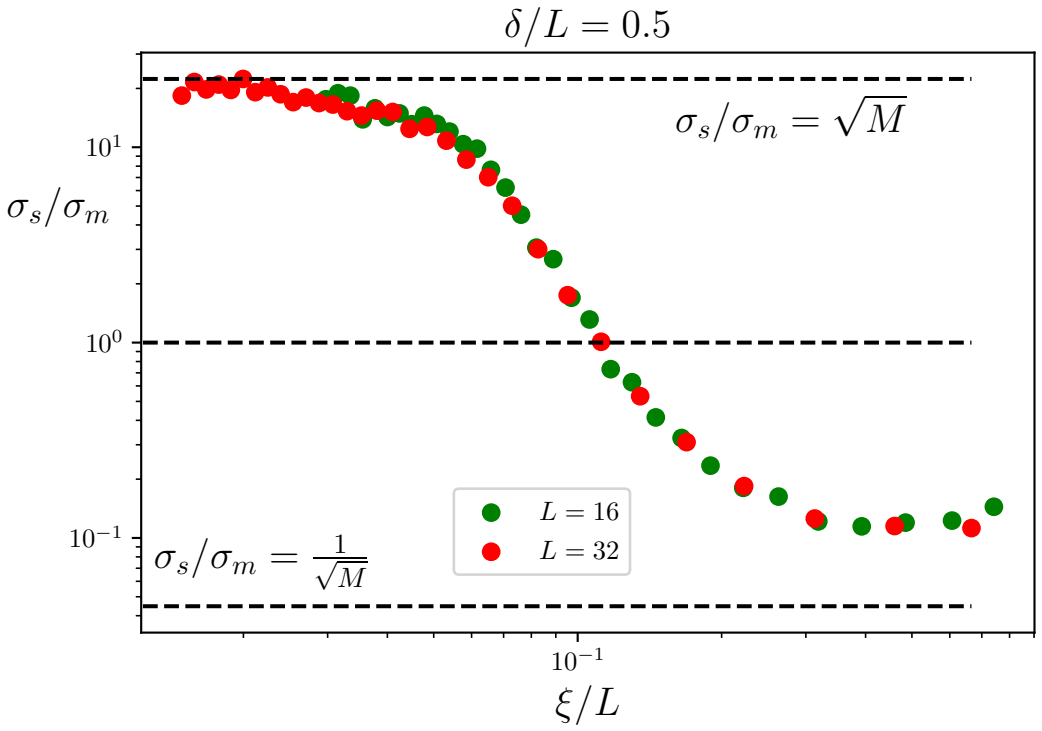

FIGURE 6.7: Performance of the multilevel algorithm across a range of r_2 and constant r_1. Ensembles generated with $N = 500$ and $M = 500$ for the multilevel system and $N = 25000$ for the single-level one. Lattices of size $L = 16$ and $L = 32$ were used, with two sub-lattices for the multilevel ensemble.

These two correlators are related by the covariance

$$\mathrm{Cov}(C_2^{SS'}, C_2^{TT'}) = \langle C_2^{SS'} C_2^{TT'} \rangle - \langle C_2^{SS'} \rangle \langle C_2^{TT'} \rangle = \frac{1}{N} \langle \phi^S \phi^{S'} \phi^T \phi^{T'} \rangle - C_2(\delta)^2. \tag{6.48}$$

We choose now to normalize the fields for convenience, defining $\varphi^R = \phi^R/\sigma_\phi$, where σ_ϕ is the standard deviation of the distribution from which the field ϕ is sampled. When we later take the ratio σ_m/σ_s, this normalization factor will cancel. We recall the equation for the correlation of the two-point correlator in the symmetric phase,

$$\alpha := \exp\left(-\frac{|S-T|}{\xi}\right) = C_2(|S-T|). \tag{6.49}$$

This leads us to assert that $\varphi^T = \alpha\varphi^S + f(\alpha)\epsilon^T$, where $\epsilon^T \sim N(0,1)$. Given that $\varphi^S = \epsilon^S \sim N(0,1)$ we have that $\sigma^2_{\varphi^T} = \alpha^2 + (f(\alpha))^2$. Asserting that $\sigma^2_{\varphi^T} = 1$ gives us that $f(\alpha) = \sqrt{(1-\alpha^2)}$. We now have random variables which model two of the slice-fields on the lattice in such a way that the expected two-point correlation function holds. We need to add-in further fields in a similar way, taking an account of all correlation functions. In general if we already have n fields decomposed, $\{\varphi^{R_1}, \varphi^{R_2}, ..., \varphi^{R_n}\}$ when we add the $(n+1)^{th}$ field we assert that $\varphi^{R_{n+1}} = \sum_{i=1}^{n} a_i \varphi^{R_i} + a_{n+1}\epsilon^{n+1}$, where $\epsilon^i \sim N(0,1)$. To find the coefficients a_i up to $i = n$, we use the two-point correlation functions, which give us

the matrix equation

$$\begin{bmatrix} \langle\varphi^{R_1}\varphi^{R_1}\rangle & \langle\varphi^{R_1}\varphi^{R_2}\rangle & \langle\varphi^{R_1}\varphi^{R_3}\rangle & \dots & \langle\varphi^{R_1}\varphi^{R_n}\rangle \\ \langle\varphi^{R_2}\varphi^{R_1}\rangle & \langle\varphi^{R_2}\varphi^{R_2}\rangle & \langle\varphi^{R_2}\varphi^{R_3}\rangle & \dots & \langle\varphi^{R_2}\varphi^{R_n}\rangle \\ \langle\varphi^{R_3}\varphi^{R_1}\rangle & \langle\varphi^{R_3}\varphi^{R_2}\rangle & \langle\varphi^{R_3}\varphi^{R_3}\rangle & \dots & \langle\varphi^{R_3}\varphi^{R_n}\rangle \\ \dots & \dots & \dots & \dots & \dots \\ \langle\varphi^{R_n}\varphi^{R_1}\rangle & \langle\varphi^{R_n}\varphi^{R_2}\rangle & \langle\varphi^{R_n}\varphi^{R_3}\rangle & \dots & \langle\varphi^{R_n}\varphi^{R_n}\rangle \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \\ a_n \end{bmatrix} = \begin{bmatrix} \langle\varphi^{R_1}\varphi^{R_{n+1}}\rangle \\ \langle\varphi^{R_2}\varphi^{R_{n+1}}\rangle \\ \langle\varphi^{R_3}\varphi^{R_{n+1}}\rangle \\ \dots \\ \langle\varphi^{R_n}\varphi^{R_{n+1}}\rangle \end{bmatrix}. \tag{6.50}$$

We can then obtain a_{n+1} by decomposing $\varphi^{R_{n+1}}$ into its constituent random variables ϵ_i and then imposing $\sigma_\varphi = 1$. This process has been automated in Python code [8]. We can now write all slice-fields according to the decomposition

$$\varphi^R = \sum_{x\in\Lambda} c_x^R \epsilon^x. \tag{6.51}$$

It is then straightforward to evaluate the four-point function $\langle\phi^S\phi^{S'}\phi^T\phi^{T'}\rangle$ by noting that $\langle\epsilon^x\epsilon^x\epsilon^x\epsilon^x\rangle = 3$, $\langle\epsilon^x\epsilon^x\epsilon^y\epsilon^y\rangle = 1$ (where $x \neq y$ and all other contributions are 0).

6.10.2 Multilevel Model

The multilevel model is much the same as the single-level one, with two key exceptions. Firstly, fields in one sub-region are only correlated indirectly to fields in a different sub-region via boundary interactions. In a two sub-lattice set-up there are two boundary fields ϕ^{B_1} and ϕ^{B_2}, N_1 fields in sub-lattice Λ_1, and N_2 in sub-lattice Λ_2. We use i to index the boundary configurations, j_1 to index configurations of Λ_1 and j_2 to index configurations of Λ_2. We perform the decomposition procedure first to the two boundaries, and then add fields from the sub-lattices. This gives us the decomposition

$$\varphi^S_{ij_1j_2} = \sum_{x_1\in\Lambda_1} c^S_{x_1}\epsilon^{x_1}_{ij_1} + \sum_{x_2\in\Lambda_2} d^S_{x_2}\epsilon^{x_2}_{ij_2} + \sum_{x_B\in\partial B} f^S_{x_B}\epsilon^{x_B}_i, \tag{6.52}$$

where $c^S_{x_1}$, $d^S_{x_2}$ and $f^S_{x_B}$ are constants and $\epsilon^{x_1}_{ij_1}$, $\epsilon^{x_2}_{ij_2}$ and $\epsilon^{x_B}_i$ are independent random variables with a distributed like $N(0,1)$. If $S \in \Lambda_1$ then by construction $d_{x_2} = 0\,\forall\, x_2$, and similarly if $S \in \Lambda_2$, then $c_k = 0\,\forall\, k$; if $S \in \partial B$ then $c_k = 0\,\forall\, k$ and $d_{x_2} = 0\,\forall\, x_2$. The two-point function is then given by

$$\begin{aligned} C_{RS} &= \left\langle \frac{1}{N}\frac{1}{M^2}\sum_{i=1}^{N}\sum_{j_1=1}^{M}\sum_{j_2=1}^{M}\varphi^R_{ij_1j_2}\varphi^S_{ij_1j_2}\right\rangle \\ &= \frac{1}{N}\frac{1}{M^2}\sum_{i=1}^{N}\sum_{j_1=1}^{M}\sum_{j_2=1}^{M}\left\langle\left(\sum_{x_1\in\Lambda_1} c^R_{x_1}\epsilon^{x_1}_{ij_1} + \sum_{x_2\in\Lambda_2} d^R_{x_2}\epsilon^{x_2}_{ij_2} + \sum_{x_B\in\partial B} f^R_{x_B}\epsilon^{x_B}_i\right)\right. \\ &\qquad \left.\times\left(\sum_{x'_1\in\Lambda_1} c^S_{x'_1}\epsilon^{x'_1}_{ij_1} + \sum_{x'_2\in\Lambda_2} d^S_{x'_2}\epsilon^{x'_2}_{ij_2} + \sum_{x'_B\in\partial B} f^S_{x'_B}\epsilon^{x'_B}_i\right)\right\rangle. \end{aligned} \tag{6.53}$$

Since the ϵ terms are by construction uncorrelated then $\langle \epsilon^x \epsilon^y \rangle = 0$ if $x \neq y$, $\forall\,(x, y)$, so we can drop all cross terms,

$$
\begin{aligned}
C_{RS} &= \frac{1}{N}\frac{1}{M^2}\sum_{i=1}^{N}\sum_{j_1=1}^{M}\sum_{j_2=1}^{M}\Bigg(\sum_{x_1\in\Lambda_1}\sum_{x_1'\in\Lambda_1} c^R_{x_1}c^S_{x_1'}\langle \epsilon^{x_1'}_{ij_1}\epsilon^{x_1}_{ij_1}\rangle + \sum_{x_2\in\Lambda_2}\sum_{x_2'\in\Lambda_2} d^R_{x_2}d^S_{x_2'}\langle \epsilon^{x_2'}_{ij_2}\epsilon^{x_2}_{ij_2}\rangle \\
&\quad + \sum_{x_B\in\partial B}\sum_{x_B'\in\partial B} f^R_{x_B}f^S_{x_B'}\langle \epsilon_i^{x_B'}\epsilon_i^{x_B}\rangle\Bigg) \\
&= \frac{1}{N}\frac{1}{M^2}\sum_{i=1}^{N}\sum_{j_1=1}^{M}\sum_{j_2=1}^{M}\Bigg(\sum_{x_1\in\Lambda_1}\sum_{x_1'\in\Lambda_1} c^R_{x_1}c^S_{x_1'}\delta_{x_1x_1'} + \sum_{x_2\in\Lambda_2}\sum_{x_2'\in\Lambda_2} d^R_{x_2}d^S_{x_2'}\delta_{x_2x_2'} \\
&\qquad\qquad + \sum_{x_B\in\partial B}\sum_{x_B'\in\partial B} f^R_{x_B}f^S_{x_B'}\delta_{x_Bx_B'}\Bigg) \\
&= \sum_{x_1\in\Lambda_1} c^R_{x_1}c^S_{x_1} + \sum_{x_2\in\Lambda_2} d^R_{x_2}d^S_{x_2} + \sum_{x_B\in\partial B} f^R_{x_B}f^S_{x_B}.
\end{aligned}
\tag{6.54}
$$

Note that by construction of the coefficients c, d and f, $C_{RS} = C_2(|R - S|)$ as in the single level model. However, unlike the single level model, we haven't directly correlated points in Λ_1 with those in Λ_2, only doing this indirectly via the boundaries. However, due to the decomposition property $e^{x-z}e^{z-y} = e^{x-y}$, this indirect correlation gives an accurate estimate of the true correlation function, as evidenced by figure (6.8). When the correlation length, ξ is large the pure-mode correlation function $C_2(\delta) = (e^{-\delta/\xi} + e^{-(L-\delta)/\xi})/(1 + e^{-L/\xi})$ is no longer valid as we are near criticality. Therefore, it is no surprise, our model (which assumes this correlation function) becomes inaccurate. By using a more accurate anzatz, the model's accuracy would likely be improved.

The covariance between two correlators under variation of the boundary configuration (e.g. changing of the index, i) is given by

$$
\mathrm{Cov}(C_{PQ}, C_{RS}) = \left\langle \frac{1}{M^2}\sum_{j_1=1}^{M}\sum_{j_2=1}^{M}\varphi^P_{ij_1j_2}\varphi^Q_{ij_1j_2}\frac{1}{M^2}\sum_{j_1'=1}^{M}\sum_{j_2'=1}^{M}\varphi^R_{ij_1'j_2'}\varphi^S_{ij_1'j_2'}\right\rangle - C_{PQ}C_{RS}. \tag{6.55}
$$

In the appendix C it is shown that the value of this covariance is given by

$$
\mathrm{Cov}(C_{PQ}, C_{RS}) = A + \frac{B}{M} + \frac{C}{M^2}, \tag{6.56}
$$

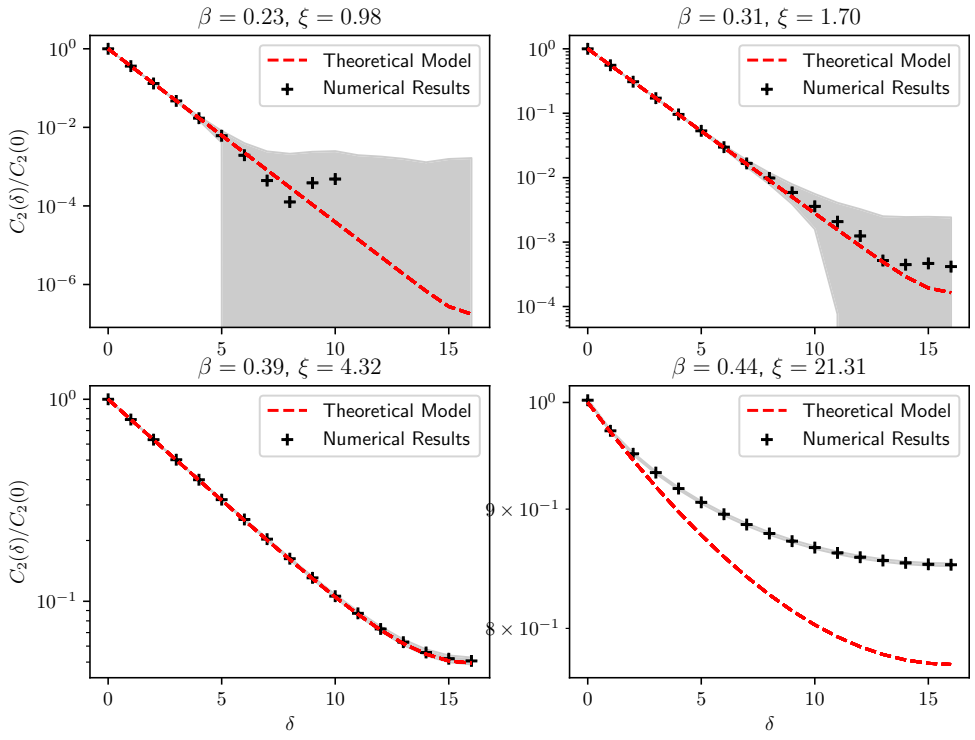

FIGURE 6.8: Comparison between the two-point correlator as obtained by numerical results and from calculations of the theoretical model. Results were obtained from $L = 32$ lattice with a single-level simulation with $N = 25000$.

where

$$A = \sum_{x,x'} \left[f_x^P f_{x'}^Q f_x^R f_{x'}^S + f_x^P f_{x'}^Q f_{x'}^R f_x^S \right], \tag{6.57}$$

$$\begin{aligned} B = \sum_{x,x'} \Big[& c_x^P c_{x'}^Q c_x^R c_{x'}^S + c_x^P c_{x'}^Q c_{x'}^R c_x^S + d_x^P d_{x'}^Q d_x^R d_{x'}^S + d_x^P d_{x'}^Q d_{x'}^R d_x^S \\ & + (c_x^P f_{x'}^Q + f_{x'}^P c_x^Q)(c_x^R f_{x'}^S + f_{x'}^R c_x^S) + (d_x^P f_{x'}^Q + f_{x'}^P d_x^Q)(d_x^R f_{x'}^S + f_{x'}^R d_x^S), \\ C = \sum_{x,x'} & (c_x^P d_{x'}^Q + d_{x'}^P c_x^Q)(c_x^R d_{x'}^S + d_{x'}^R c_x^S). \end{aligned}$$

Intuitively these expressions make sense. The contribution which does not scale at all with M is the one resulting from contributions to the fields directly resulting from the boundaries (f-contributions). On the opposite end of things, we see that the only terms contributing to the ideal $1/M^2$ scaling are those involving contributions mixing the two sub-lattices. Specifically, we have to have c and d contributions from each correlator, e.g. both correlators individually span the two sub-regions.

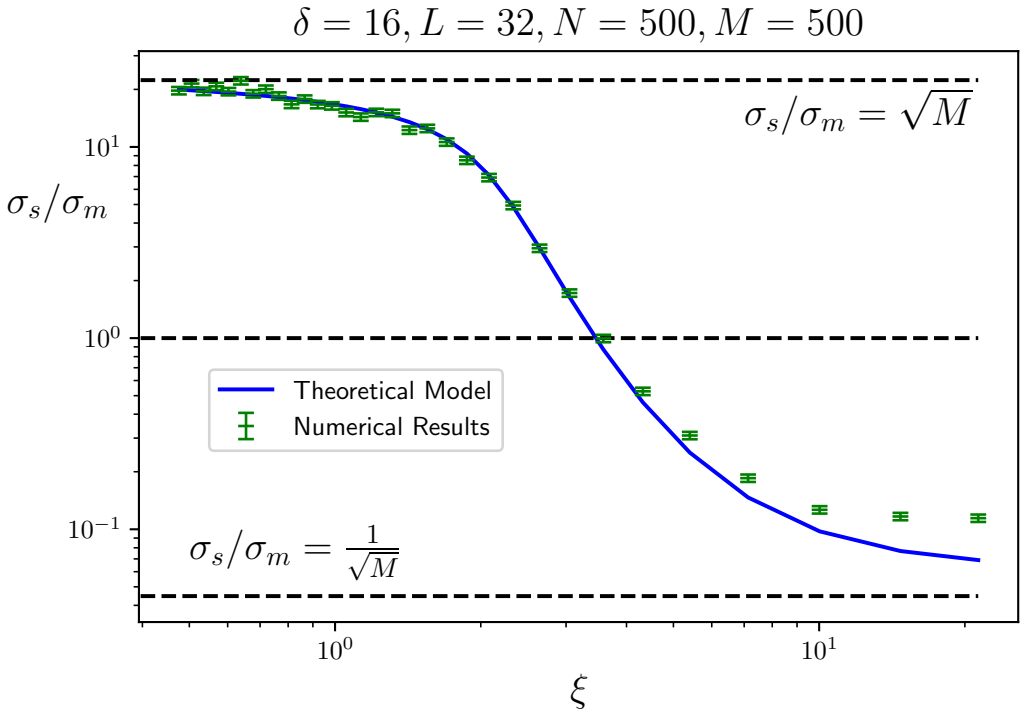

FIGURE 6.9: Theoretical vs. observed performance gain of the multilevel algorithm as compared to the single-level one, as measured by the ratio of the standard deviations of the longest correlator ($\delta = L/2$) in the system. Numerical results use an $L = 32$ lattice with $N = 500$ and $M = 500$ for multilevel and $N = 250000$ for single-level. Error band on numerical results calculated by taking the error in quadrature from the single level and multilevel contributions.

6.11 Theoretical Model Performance

Given that we can estimate the covariance matrix of the correlators, we can use the theoretical model to calculate an optimum weighting of correlator contributions. This allows us to estimate the overall variance of the correlators in both the single-level and multilevel algorithms. The model proves to be very accurate to the numerical results, failing only as ξ becomes large. As discussed previously, this is a regime where the functional form for the two-point correlator becomes inaccurate, and therefore any attempt to model the system using this is likely to be unsuccessful.

To relate this study back to the challenge of testing holographic cosmology through the simulation of a lattice-regulated dual theory, we arrive at an unfortunate conclusion. Our theory needs to be simulated at the critical-mass point so that it corresponds to a continuum theory with generalized conformal invariance. However, we have now seen that the critical point, where the correlation length of the system diverges, is exactly the point at which multilevel techniques fail to give any performance improvement (in fact they perform more poorly than a single-level method). Therefore, we conclude that

it is unlikely that these techniques will be useful in testing the holographic cosmology conjecture.

Chapter 7

Conclusion

Throughout the centuries, breakthroughs in physics have allowed us to look back further into the past of our universe. This has taught us that the universe is and always has been expanding. The Cosmic Microwave Background Radiation (CMB) is a window into the very earliest moments of our universe. Its incredible uniformity in the night sky, along with other observations, motivates new physics. The most popular model of this new physics is cosmic inflation, where a short, violent expansion of the universe effectively extends conformal time allowing disparate regions of the night sky to become causally linked. This model is lacking in several aspects, however. It does not resolve the initial singularity problem of the Big Bang and suffers from self-consistency issues. It is not a UV-complete theory - it is an effective theory, incapable of explaining physics at energies above the Planck scale, which are the energy scales at the start of the universe. To describe physics at these energy scales would require a quantum description of gravity.

One very prominent paradigm of quantum gravity uses the holographic principle. Here there is a duality between $(d+1)$-dimensional theories with gravity and d-dimensional quantum field theories (QFTs) with a large-N limit. The most famous of these dualities is the AdS/CFT correspondence. A duality between a four-dimensional gravity theory with the geometry of our universe and a dual Euclidean QFT with a generalized conformal structure has been proposed. This dual theory can contain scalars, gauge fields, and fermions in the adjoint of $SU(N)$. These models are collectively referred to as holographic cosmology models, and they directly connect fluctuations in the CMB to two-point functions of the energy-momentum tensor of the dual QFT. This offers an opportunity to test the holographic principle in our universe. If shown to be correct it would be one of the great leaps in humanity's understanding of the natural world. Time, for example, would be viewed not simply as a dimension with a minus sign[1] in the metric, but as a flow in energy scale of the dual theory.

[1] Or plus sign depending on your tastes

Experimental tests of holographic cosmology using perturbation theory and CMB data have found this new paradigm of cosmology to be competitive with the ΛCDM model, the "standard model of cosmology". However, some of the largest discrepancies of ΛCDM with observational evidence lie in the low-multipole range. Here the holographic theory is non-perturbative, so a change of strategies is required. The LatCos collaboration, of which the author is a member, is tasked with using lattice QFT methods to make predictions of holographic cosmology, which are valid non-perturbatively. Lattice regularization regulates both the infrared and the ultraviolet of a theory by placing the theory on a discrete lattice of finite size. The theory in this form is then amenable to study using computer simulation with Monte Carlo Markov chains.

Before predictions can be made using lattice QFT, to test against the CMB, three major problems need to be overcome. The first problem is that the action of the dual QFT is super-renormalizable. This means that the theory is expected to be asymptotically free in the ultraviolet limit, but also to suffer from divergences in the infrared. These divergences would reveal the theory to be ill-defined. Thankfully, a potential mechanism has been proposed that implies this class of theory is non-perturbatively infrared finite. In chapter 4, we explored the infrared finiteness of the theory in great detail. For simplicity, a pure scalar model has been investigated as the infrared properties of such a model are representative of holographic dual theory models. In this model, a bare-mass parameter m^2 is needed to cancel radiative corrections which contribute to a ϕ^2 term. Such a cancellation requires this mass parameter to be at its critical value $m^2 = m_c^2$. As the bare mass is varied across this critical value in lattice simulations, a phase transition is observed, allowing the critical mass to be pinpointed. In this study, the Binder Cumulant was used to detect the phase transition (and therefore the critical mass) across a range of lattice sizes L and coupling constants g. The behavior of the critical mass in the infrared limit ($gL \to \infty$) is exemplary of the infrared behavior of the theory and formed the focus of this study. By comparing the performance of infrared-finite and infrared-infinite anzätze on critical mass data, we have demonstrated that the theory is infrared-finite. This was done through statistical analysis, including frequentist and Bayesian approaches, answering the question of which anzatz is preferred. On top of this success, the study also yielded estimates of finite-size scaling parameters such as the critical exponent ν. These parameters provide a formula for the critical mass as a function of the coupling and the lattice size, allowing us to better choose bare-mass parameters in future studies.

While we have successfully allayed our fears of infrared-finiteness issues in our theory, the challenge of the ultraviolet remains - the second major problem. When defining a lattice QFT, the lattice spacing a provides an ultraviolet cut-off with maximum momentum $2\pi/a$. To make physical predictions of our holographic theory, a continuum $a \to 0$ limit of correlators of interest (in this case the two-point momentum of the EMT) needs to be taken. This limit will only be finite if all divergences present in the correlators

are removed. This is a two-step problem: firstly, the EMT operator itself needs to be renormalized, meaning any contributions to the operator which are divergent in the ultraviolet must be removed. Secondly, additional divergences when taking the two-point function of the renormalized EMT operator must be subtracted. The first step has been done with a Wilson Flow method, while the second step is the focus of chapter 5. In this chapter, we considered how the q^2-divergence appearing in this two-point function can be removed from our lattice data. The method pursued utilizes properties of the Fourier and Laplace transforms, namely that, depending on the definitions, the Laplace transform of a function is the analytic continuation $q \rightarrow -iq$ of the Fourier transform of that same function. By the addition of Fourier and Laplace transforms, the q^2-divergence can be canceled exactly in the continuum. Defining a Laplace transform on a lattice is challenging, however, with various ambiguities appearing in the definition. Nevertheless, in chapter 5, we have provided a definition of the Laplace transform which in the continuum limit $a \rightarrow 0$ gives the expected cancellation with the Fourier transform. For a finite lattice spacing, this cancellation is not exact, however, and a q^4-term remains. This remaining contribution is proportional to the size of the initial q^2-divergence and causes a bias in the estimated parameters proportionally. Given the size of the q^2-divergence in our data is unknown, the bias in this method leads us to be wary of this technique. Therefore, going forward, we use an alternative method (windowing), which removes the q^2-divergence entirely with no remainder.

In principle, having demonstrated infrared-finiteness and renormalized the ultraviolet, our collaboration is almost ready to make non-perturbative predictions of holographic cosmology based on our three-dimensional dual QFT. However, throughout the investigation of holographic dual theories on the lattice, issues of statistical noise have plagued the precision of results - the third problem. One method to resolve this is through computing power, however, given that ensemble generation is already in the millions of CPU hours, a significant increase in statistics through this method is not viable [2]. Therefore techniques that could improve the signal-to-noise ratio of our predictions are being investigated. One method which we explore in detail in chapter 6 is the multi-level algorithm. This technique divides a lattice into sub-lattices separated by boundary layers. By first sampling the boundary, and then sampling the sub-lattices, the effective number of contributions to a two-point function can be increased. Specifically, if N boundary configurations and M sub-lattice configurations are produced, a multilevel algorithm may be able to achieve $1/\sqrt{NM^2}$ scaling in two-point correlator estimates - superior to the $1/\sqrt{NM}$ scaling of a computer-time equivalent single level algorithm. Such a method relies on the contributing points in the two-point function being sufficiently far from the boundary such that they are statistically (close to) independent from it. The required distance scales like the correlation length ξ of the system, and in chapter 6, we

[2] It should be noted here that efforts have been made by the LatCos collaboration to port our code onto the GPU using CUDA, which have resulted in performance improvements. Whether this will fully alleviate statistical noise issues remains the subject of ongoing study.

demonstrate (both with a numerical study done in the 2D-Ising model and a theoretical framework) that when the correlation length is larger than around a tenth of the lattice size, the multilevel algorithm fails to offer performance improvements over a single level method. For the holographic duality to hold, the dual theory needs to have the property of generalized conformal invariance, requiring the mass to be tuned to the critical mass value. Here the correlation length diverges, and we expect the multilevel algorithm will fail to offer performance improvements. Therefore, this algorithm is not expected to be used to improve the signal-to-noise ratio in our future simulations.

Looking forward, exploring holographic models of cosmology by performing simulations of lattice-regularized dual QFTs remains an exciting avenue of research. Infrared-finiteness has been demonstrated, and a procedure for renormalization has been developed. A code for generating ensembles with scalar and gauge fields using the GRID library has now been completed, allowing the LatCos collaboration to explore a greater area of holographic cosmology theory space. This is especially exciting as the majority of well-explored holographic dualities contain gauge fields in the boundary theory - so much so that these dualities are often referred to as gauge-gravity dualities. Moreover, this code (along with the pure scalar code) has been ported for use with NVidia GPUs. Already simulations with such a setup are being performed on the Tursa supercomputer in Edinburgh with Nvidia A100 GPU nodes. This technology offers the opportunity to produce larger ensembles more efficiently and could reduce statistical noise in estimates of EMT two-point functions. In this thesis, ensembles with scalars in the adjoint of $SU(N)$ with $N = 2, 3, 4$ and 5 have been used. Since holographic dualities are valid in the large-N limit, a future avenue of research would extrapolate our results into the large-N limit. This process will likely involve simulations at higher values of N.

Appendix A

Binder Cumulant Scaling

To predict the value of ζ (see eq. (4.37)) at $B = \bar{B}$, we use two expressions from [90]. The first expression is the scaling-law for the Binder Cumulant at $m^2_{c,\infty}(g)$ (where $\zeta = 0$),

$$B = B_c + \sum_{k=1}^{\infty} \tilde{c}_k (ag)(gL)^{-\omega k} + \sum_j \tilde{e}_j (ag)(gL)^{-\epsilon_j}. \tag{A.1}$$

This expression states that the Binder Cumulant at the infinite-volume critical point is equal to a constant critical value B_c, plus an adjustment due to correction-to-scaling terms. The next expression we use is for the gradient of the Binder Cumulant at a particular value B_0,

$$\begin{aligned} \left.\frac{dB}{d\zeta}\right|_{\zeta=\tilde{f}(B_0)} &= a(B_0)(gL)^{\frac{1}{\nu}} \left(1 + \sum_{k=1}^{\infty} d_k(g, B_0)(gL)^{-\omega k} + \sum_{j=1}^{2} e_j(g, B_0)(gL)^{-\epsilon_j}\right) \\ &\qquad + \sum_{k=1}^{\infty} c_k(g, B_0)(gL)^{-\omega k}. \\ &= f'(\tilde{f}(B_0)), \end{aligned} \tag{A.2}$$

where the apostrophes are used to indicate the derivative of a function with respect to its argument and we recount the definitions of the inverse functions f and $\tilde{f}$,

$$B = f(\zeta), \tag{A.3}$$

$$\zeta = \tilde{f}(B). \tag{A.4}$$

We can predict $\zeta(\bar{B})$ by integrating along the path $f(0) \to \bar{B}$:

$$\begin{aligned}\zeta(\bar{B}) &= 0 + \int_{f(0)}^{\bar{B}} \left.\frac{d\zeta}{dB}\right|_{B=B'} dB' \\ &= \int_{f(0)}^{\bar{B}} \left(\left.\frac{dB}{d\zeta}\right|_{\zeta=\tilde{f}(B')}\right)^{-1} dB' \\ &= \int_{f(0)}^{\bar{B}} \frac{1}{f'(\tilde{f}(B'))} dB',\end{aligned} \tag{A.5}$$

where we've used $(dy/dx)|_{x=x_0} = ((dx/dy)|_{y=y(x_0)})^{-1}$. This integral can be split into two parts, $\int_{f(0)}^{B} = I_1 + I_2$, where

$$I_1 = \int_{B_c}^{\bar{B}} \frac{1}{f'(\tilde{f}(B))} dB, \tag{A.6}$$

$$I_2 = \int_{f_0}^{B_c} \frac{1}{f'(\tilde{f}(B))} dB, \tag{A.7}$$

where B_c is the critical Binder Cumulant value defined in eq. (A.1).

We first consider I_1,

$$\begin{aligned}I_1 &= \int_{B_c}^{\bar{B}} dB \left(a(B)(gL)^{\frac{1}{\nu}} \left(1 + \sum_{k=1}^{\infty} d_k(g,B)(gL)^{-\omega k} + \sum_{j=1}^{2} e_j(g,B)(gL)^{-\epsilon_j}\right) + \sum_{k=1}^{\infty} c_k(g,B)(gL)^{-\omega k}\right)^{-1} \\ &= \int_{B_c}^{\bar{B}} dB (gL)^{-\frac{1}{\nu}} a(B)^{-1} \left(1 + \sum_{k=1}^{\infty} d_k(g,B)(gL)^{-\omega k} + \sum_{j=1}^{2} e_j(g,B)(gL)^{-\epsilon_j} + \sum_{k=1}^{\infty} \frac{c_k(g,B)}{a(B)}(gL)^{-\omega k - \frac{1}{\nu}}\right)^{-1} \\ &= \int_{B_c}^{\bar{B}} dB (gL)^{-\frac{1}{\nu}} a(B)^{-1} \Bigg(1 - \sum_{k=1}^{2} d_k(g,B)(gL)^{-\omega k} + d_1(g,B)^2 (gL)^{-2\omega} - \sum_{j=1}^{2} e_j(g,B)(gL)^{-\epsilon_j} \\ &\qquad + O\left((gL)^{-2\epsilon}\right) + O\left((gL)^{-3\omega}\right)\Bigg).\end{aligned}$$

The scaling terms can be pulled outside the integral, and since $a(B)$ has no zeros, the integrals of the parameters over B can be performed, giving

$$\begin{aligned}I_1 &= A(B)(gL)^{-\frac{1}{\nu}} + \sum_{k=1}^{2} D_k(g,B)(gL)^{-\frac{1}{\nu}+k\omega} + \tilde{D}(g,B)(gL)^{-\frac{1}{\nu}+2\omega} \\ &\quad + \sum_{j=1}^{2} E_j(g,B)(gL)^{-\frac{1}{\nu}+\epsilon_j} + O\left((gL)^{-2\epsilon-\frac{1}{\nu}}\right) + O\left((gL)^{-3\omega-\frac{1}{\nu}}\right) + O\left((gL)^{-2/\nu-\omega}\right),\end{aligned} \tag{A.8}$$

where

$$A(\bar{B}) = \int_{B_c}^{\bar{B}} dB \frac{1}{a(B)}, \tag{A.9}$$
$$D_k(g,\bar{B}) = \int_{B_c}^{\bar{B}} dB \frac{d_k(B)}{a(B)},$$
$$\tilde{D}(g,\bar{B}) = \int_{B_c}^{\bar{B}} dB \frac{d_1^2(B)}{a(B)},$$
$$E_k(g,\bar{B}) = \int_{B_c}^{\bar{B}} dB \frac{e_k(g,B)}{a(B)}.$$

We evaluate I_2, keeping terms up to the same order as for I_1:

$$I_2 = \int_{f_0}^{B_c} \left(\frac{(gL)^{-\frac{1}{\nu}}}{a(B)} - \frac{d(g,B)(gL)^{-\frac{1}{\nu}-\omega}}{a(B)} \right) dB + O\left((gL)^{\frac{1}{\nu}+3\omega}\right) + O\left((gL)^{-\frac{1}{\nu}-2\epsilon}\right) \tag{A.10}$$
$$+ O\left((gL)^{-\frac{1}{\nu}-(\epsilon+\omega)}\right).$$

Since the Binder Cumulant is strictly monotonic, $a(B')$ has no zeros and $1/a(B)$ and $d(g,B)/a(B)$ are analytic, so we can Taylor expand around B_c with coefficients a_n and d'_n respecively,

$$I_2 = \sum_{n=0}^{\infty} \left[a_n (gL)^{-\frac{1}{\nu}} \int_{B_c}^{f(0)} (B-B_c)^n dB + d'_n (gL)^{-\frac{1}{\nu}-\omega} \int_{B_c}^{f(0)} (B-B_c)^n dB \right] \tag{A.11}$$
$$+ O\left((gL)^{-\frac{1}{\nu}-3\omega}\right) + O\left((gL)^{-\frac{1}{\nu}-2\epsilon}\right) + O\left((gL)^{-\frac{1}{\nu}-(\epsilon+\omega)}\right)$$
$$= (gL)^{-\frac{1}{\nu}} \sum_{n=0}^{\infty} \frac{a_n}{n+1} [f(0)-B_c]^{n+1} + (gL)^{-\frac{1}{\nu}-\omega} \sum_{n=0}^{\infty} \frac{d'_n}{n+1} [f(0)-B_c]^{n+1}$$
$$+ O\left((gL)^{\frac{1}{\nu}+3\omega}\right) + O\left((gL)^{-\frac{1}{\nu}-2\epsilon}\right) + O\left((gL)^{-\frac{1}{\nu}-(\epsilon+\omega)}\right)$$
$$= (gL)^{-\frac{1}{\nu}} \sum_{n=0}^{\infty} \left(\frac{a_n}{n+1} + (gL)^{-\omega} \frac{d'_n}{n+1} \right) \left[\sum_{k=1}^{\infty} \tilde{c}_k(g)(gL)^{-\omega k} + \sum_j \tilde{e}_j(g)(gL)^{-\epsilon_j} \right]^{n+1}$$
$$+ O\left((gL)^{\frac{1}{\nu}+3\omega}\right) + O\left((gL)^{-\frac{1}{\nu}-2\epsilon}\right) + O\left((gL)^{-\frac{1}{\nu}-(\epsilon+\omega)}\right),$$

where in the second line we have used eq. (A.1). Keeping terms with scaling coefficients up to $1/\nu+\epsilon_j$ and $1/\nu+2\omega$, combining with I_1 and redefining scaling coefficients gives

$$\zeta(\bar{B}) = A(\bar{B})(gL)^{-\frac{1}{\nu}} + \sum_{k=1}^{2} D_k(g,\bar{B})(gL)^{-\frac{1}{\nu}+\omega} + D_2(g,\bar{B})(gL)^{-\frac{1}{\nu}+2\omega} \tag{A.12}$$
$$+ \sum_{j=1}^{2} E_j(g,\bar{B})(gL)^{-\frac{1}{\nu}+\epsilon_j} + O\left((gL)^{-2\epsilon}\right) + O\left((gL)^{-3\omega}\right),$$

where we have redefined

$$\begin{aligned} D_1(g,B) &= D_1(g,B) + a_0\tilde{c}_1(g), \\ D_2(g,B) &= D_2(g,B) + \tilde{D}(g,B) + a_0\tilde{c}_2(g) + \frac{a_1}{2}(\tilde{c}_1(g))^2 + d_0\tilde{c}_1(g), \\ E_k(g,B) &= E_k(g,B) + a_1\tilde{e}_k(g). \end{aligned} \tag{A.13}$$

Note that we can take the leading order of eq. (A.12) and rearrange for $\bar{B}$. We can do this since the function $A(\bar{B})$ is strictly monotonic and non-zero. This gives us

$$\bar{B} = A^{-1}\left(\frac{m^2 - m^2_{c,\infty}(g)}{g^2}(gL)^{1/\nu}\right), \tag{A.14}$$

where we have used the definition of ζ. This is exactly the formula quoted in [1], taking $A^{-1} = f$. As a further aside we note here that taking up to the linear order of the Taylor expansion of f is not strictly necessary if one is considering only two values of $\bar{B}$ in a given fit. This is because taking the linear order of the Taylor series gives two parameters ($f(0)$ and $f'(0)$), which is exactly the number of parameters one has if one does not make this approximation - namely $A(B_1)$ and $A(B_2)$. The two pairs of parameters are linked by

$$\begin{aligned} A(B_1) &= \frac{B_1 - f(0)}{f'(0)}, \\ A(B_2) &= \frac{B_2 - f(0)}{f'(0)}. \end{aligned} \tag{A.15}$$

Appendix B

Binder Cumulant Plots

In this appendix the plots of the Binder Cumulant for models A, B, C and D for $N \in \{2, 3, 4, 5\}$ are shown.

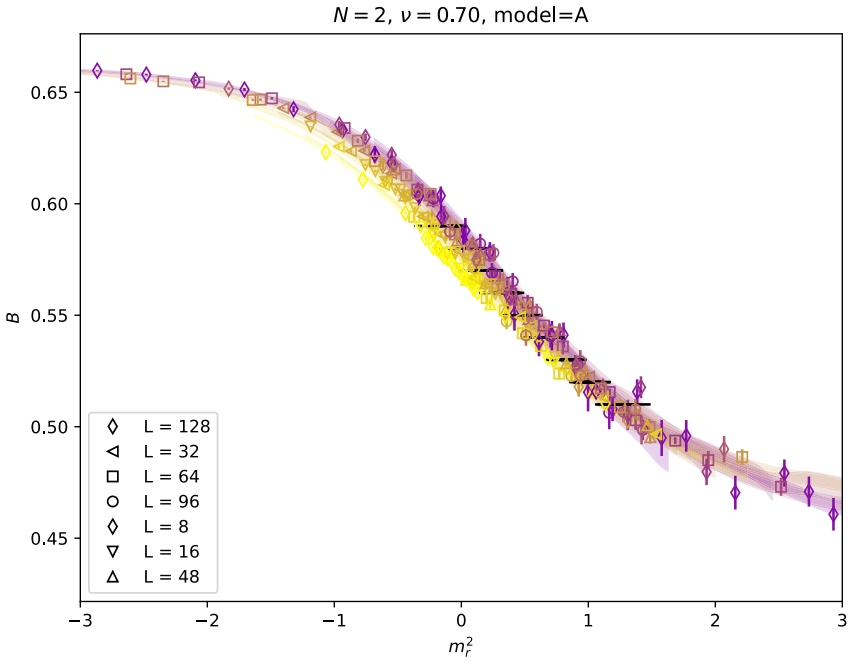

FIGURE B.1: The Binder Cumulant for $N = 2$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model A_1. All data down to and including $gL_{\min} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

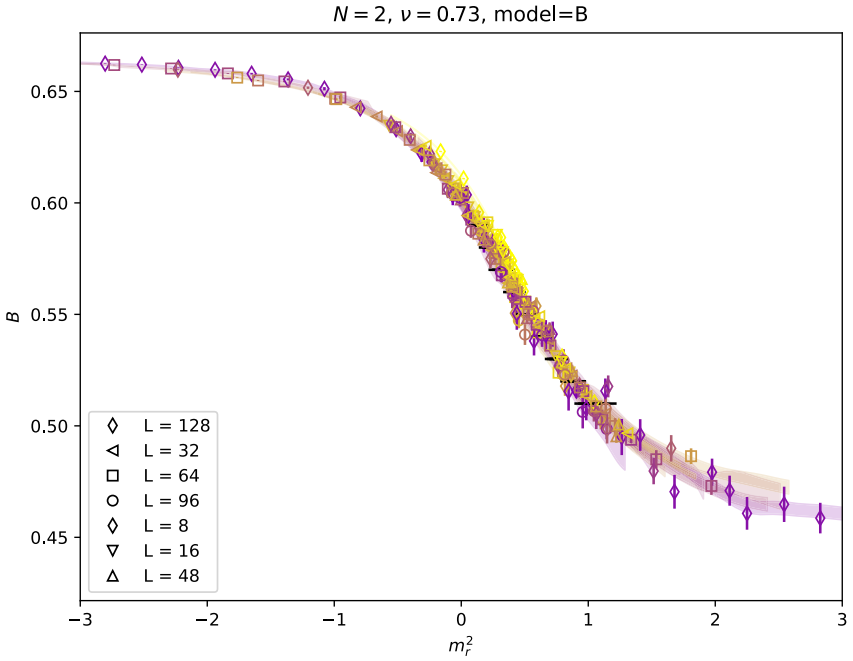

FIGURE B.2: The Binder Cumulant for $N = 2$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model B_1. All data down to and including $gL_{\min} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

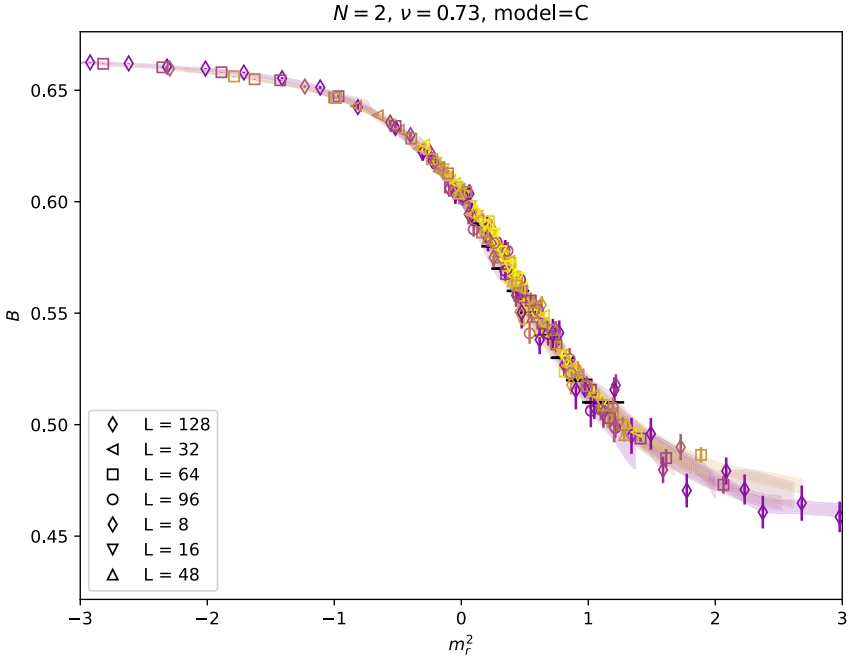

FIGURE B.3: The Binder Cumulant for $N = 2$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model C_1. All data down to and including $gL_{\min} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

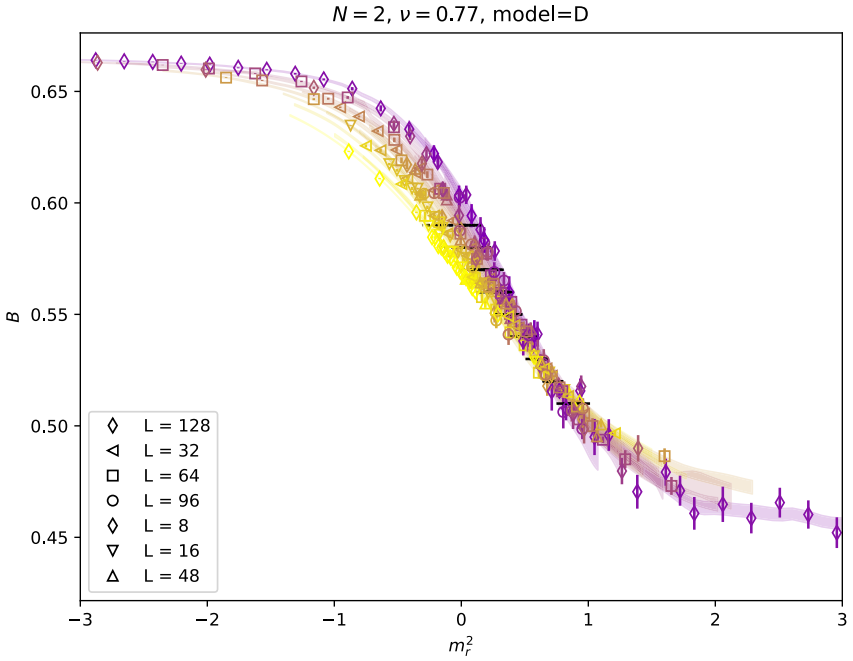

FIGURE B.4: The Binder Cumulant for $N = 2$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model D_1. All data down to and including $gL_{\text{min}} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

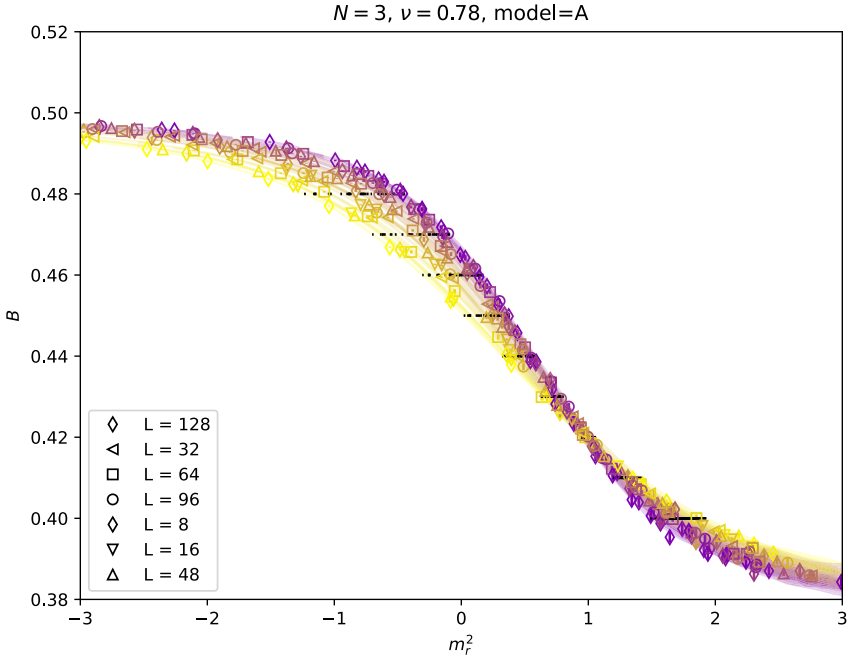

FIGURE B.5: The Binder Cumulant for $N = 3$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model A_1. All data down to and including $gL_{\text{min}} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

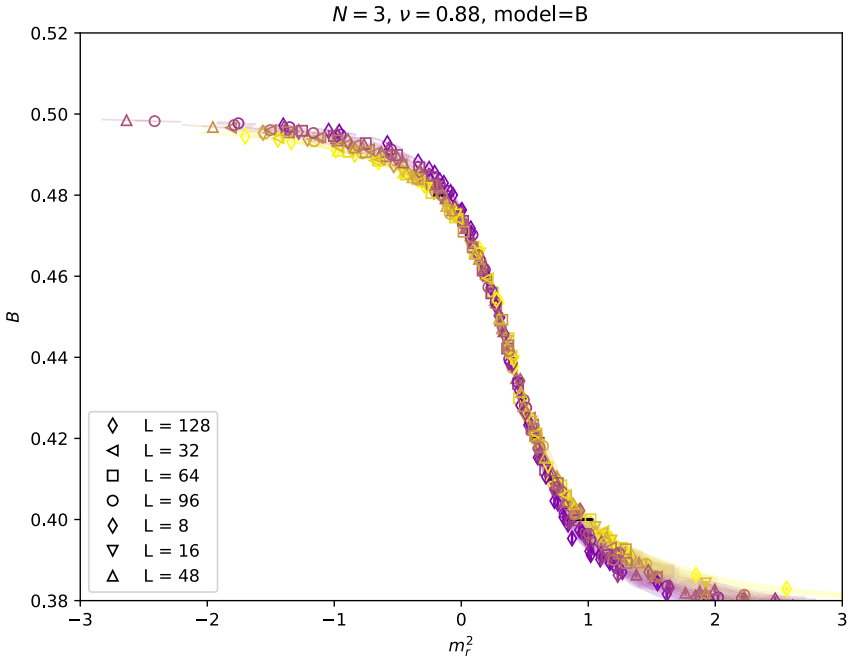

FIGURE B.6: The Binder Cumulant for $N = 3$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model B_1. All data down to and including $gL_{\min} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

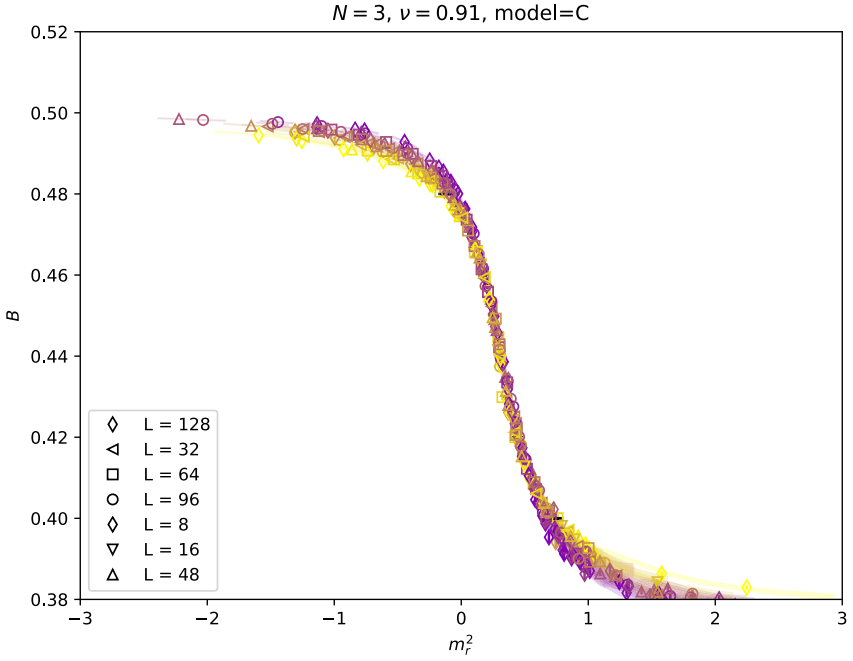

FIGURE B.7: The Binder Cumulant for $N = 3$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model C_1. All data down to and including $gL_{\min} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

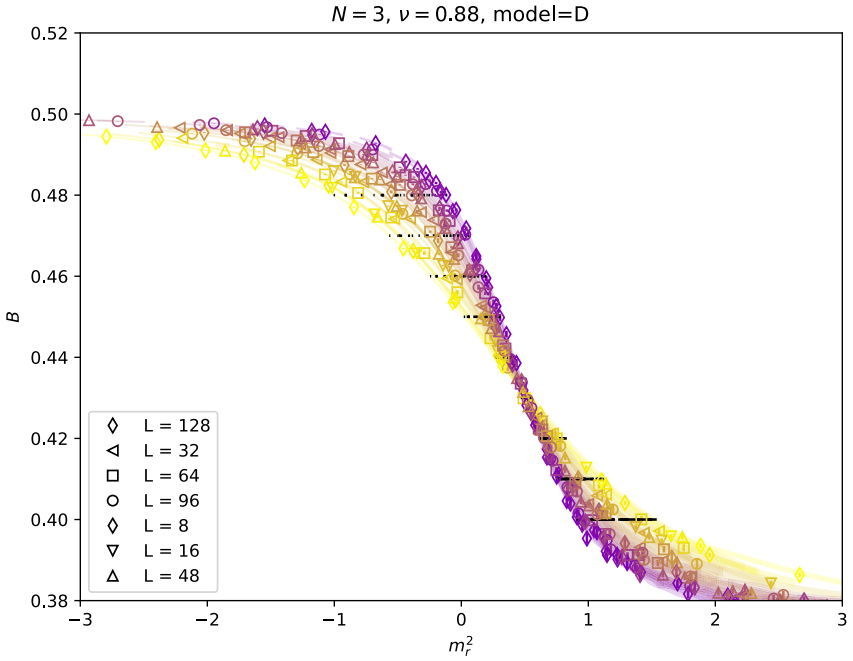

FIGURE B.8: The Binder Cumulant for $N = 3$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model D_1. All data down to and including $gL_{\text{min}} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

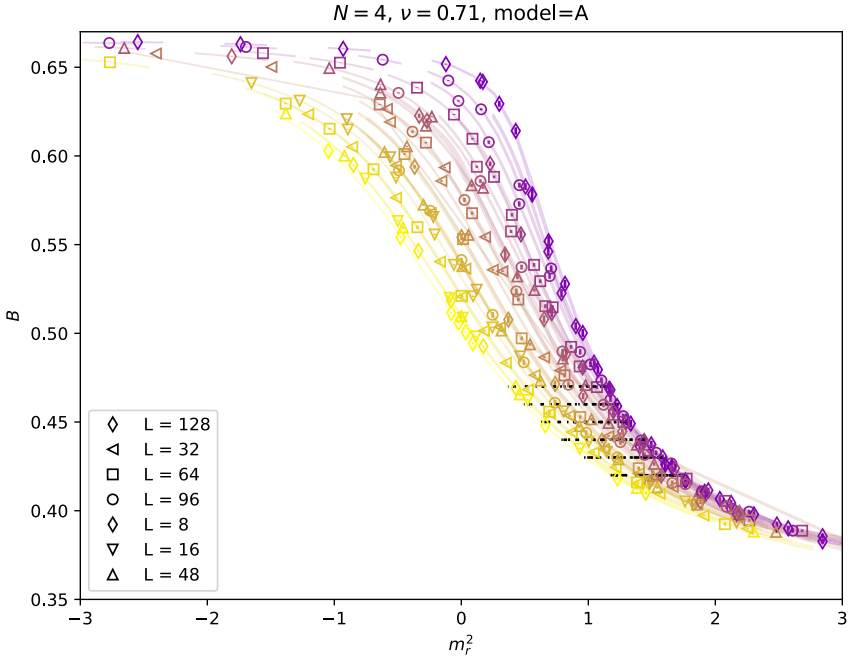

FIGURE B.9: The Binder Cumulant for $N = 4$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model A_1. All data down to and including $gL_{\text{min}} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

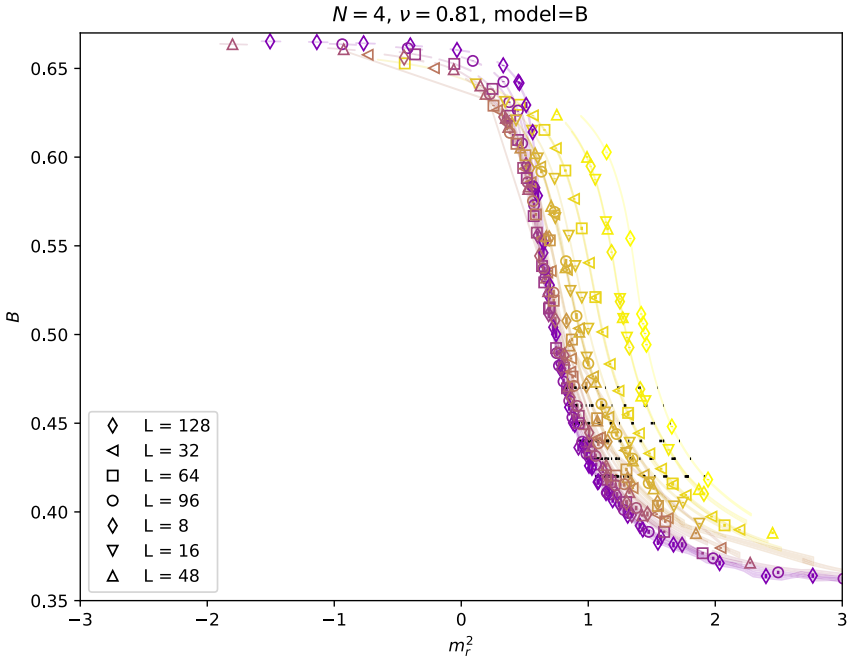

FIGURE B.10: The Binder Cumulant for $N = 4$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model B_1. All data down to and including $gL_{\min} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

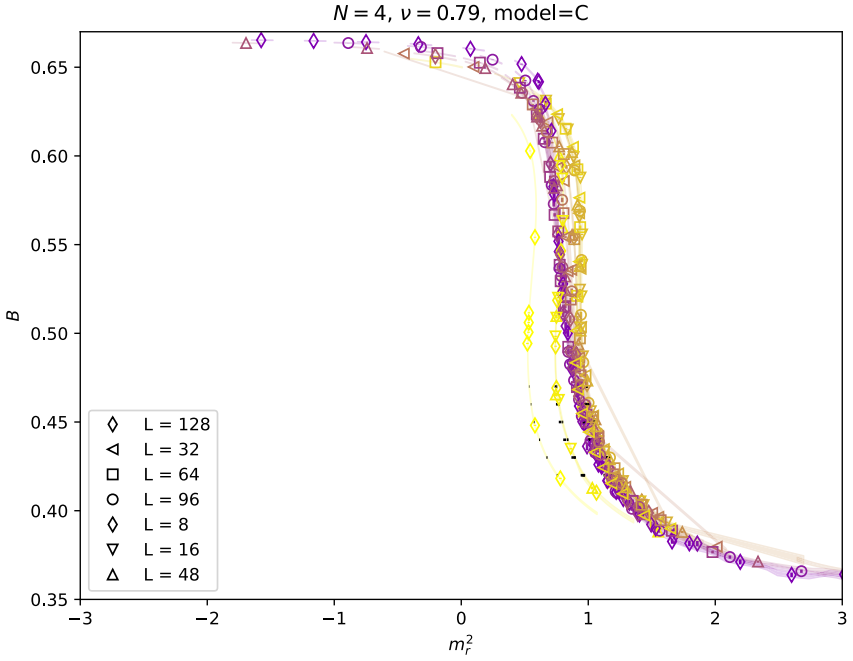

FIGURE B.11: The Binder Cumulant for $N = 4$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model C_1. All data down to and including $gL_{\min} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

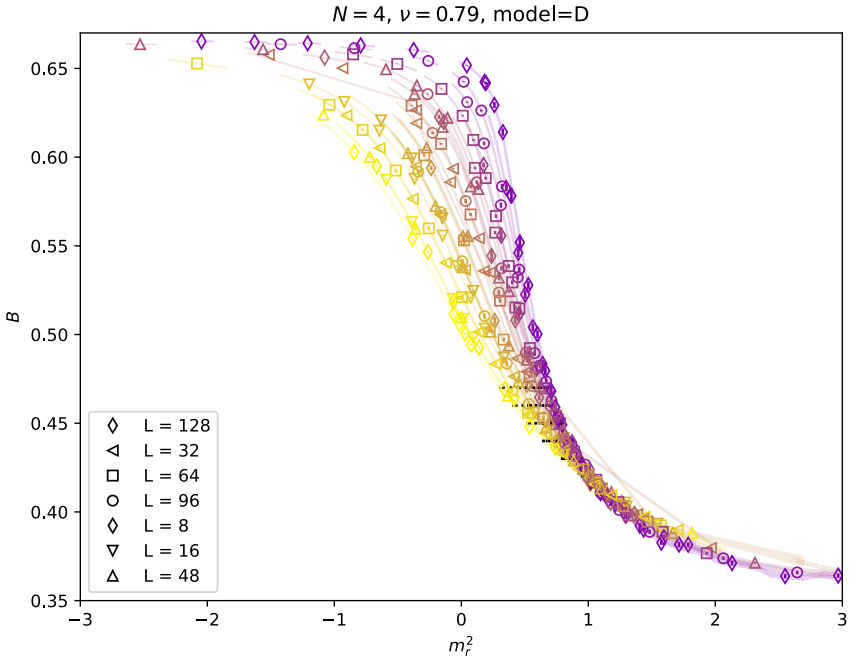

FIGURE B.12: The Binder Cumulant for $N = 4$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model D_1. All data down to and including $gL_{\min} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

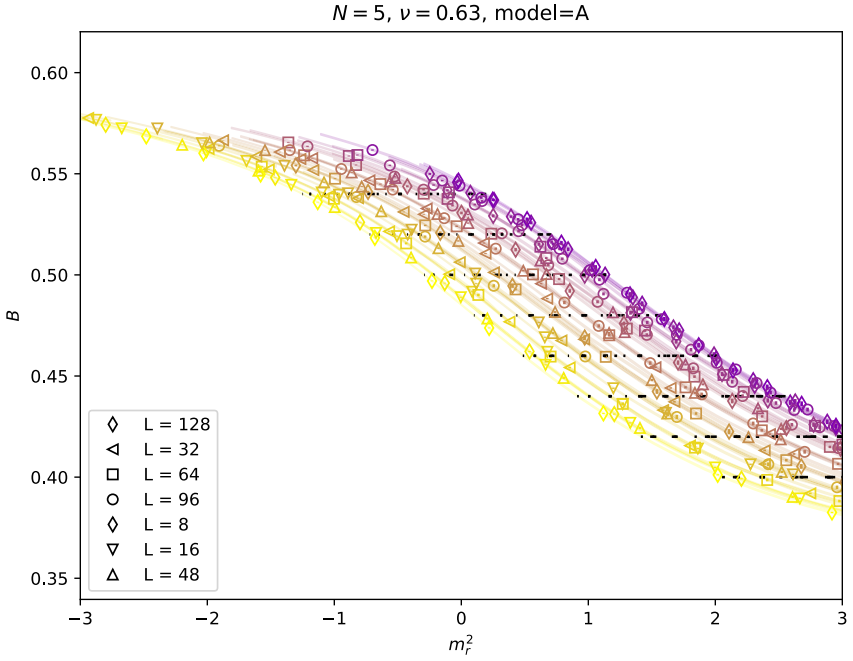

FIGURE B.13: The Binder Cumulant for $N = 5$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model A_1. All data down to and including $gL_{\min} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

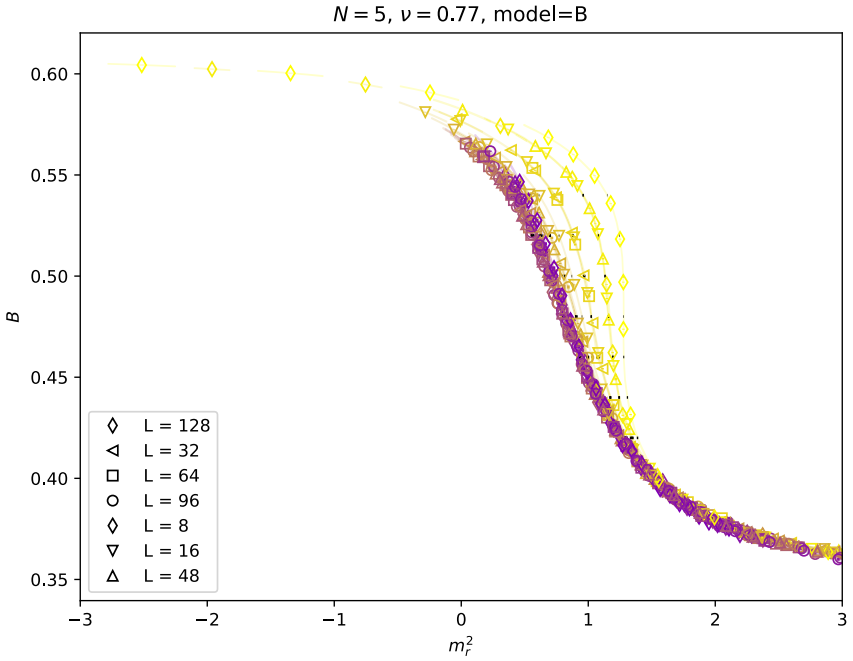

FIGURE B.14: The Binder Cumulant for $N = 5$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model B_1. All data down to and including $gL_{\min} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

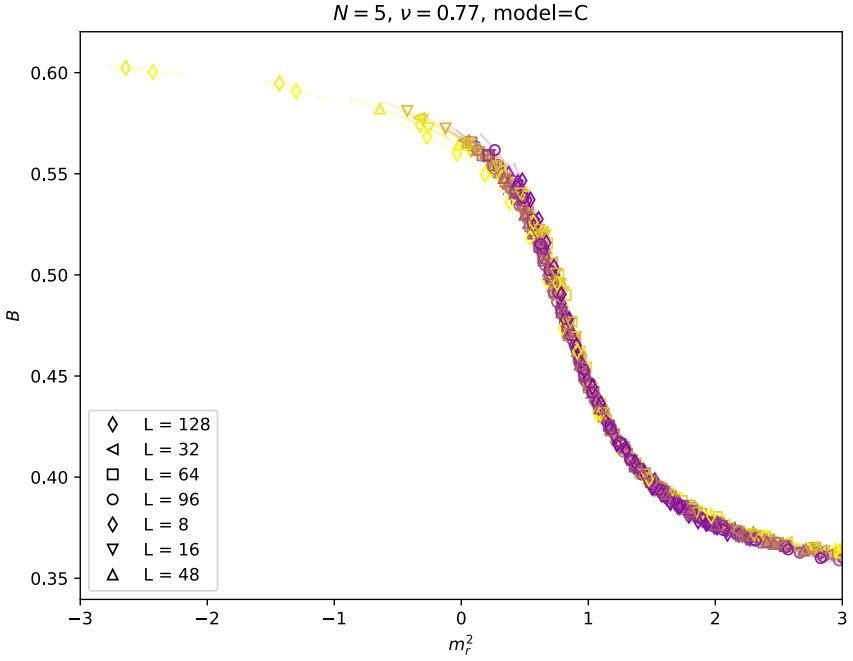

FIGURE B.15: The Binder Cumulant for $N = 5$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model C_1. All data down to and including $gL_{\min} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

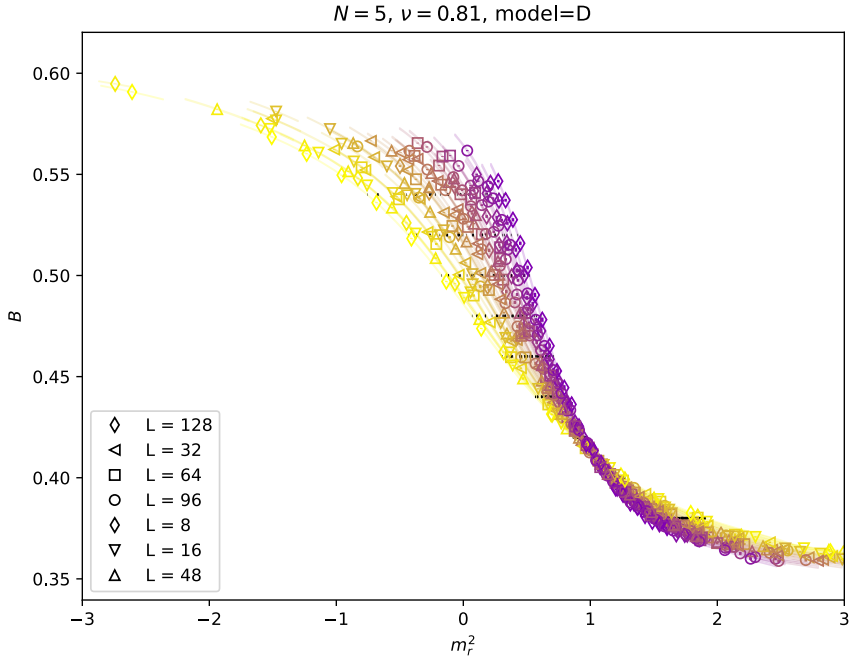

FIGURE B.16: The Binder Cumulant for $N = 5$ data against the rescaled mass defined in eq. (4.44). Data has been fit using parameter estimates using model D_1. All data down to and including $gL_{\min} = 4$ are included, with the color gradient shifting from yellow to purple as gL increases.

Appendix C

Multilevel Covariance

To derive equation (6.56) we need to evaluate the multilevel expression,

$$Z = \left\langle \frac{1}{M^4} \sum_{j_1=1}^{M} \sum_{j_2=1}^{M} \sum_{j_1'=1}^{M} \sum_{j_2'=1}^{M} \varphi^P_{ij_1j_2} \varphi^Q_{ij_1j_2} \varphi^R_{ij_1'j_2'} \varphi^S_{ij_1'j_2'} \right\rangle \tag{C.1}$$

The four-point function is similar to the above expression except all fields use the same j_1 and j_2 indices. Therefore, while the four-point function is a physical quantity, (C.1) is not. Therefore, this expression is not the same as a single level one and must be evaluated explicitly. This expression evaluates to the same value for all i, so from here on we drop this index to reduce clutter. Substituting in the formula for the φ (eq. (6.52)) we get the expression

$$\begin{aligned} Z = \frac{1}{M^4} \sum_{j_1,j_2,j_1',j_2'=1}^{M} \Bigg\langle & \left(\sum_{x_1 \in \Lambda_1} c^P_{x_1} \epsilon^{x_1}_{j_1} + \sum_{x_2 \in \Lambda_2} d^P_{x_2} \epsilon^{x_2}_{j_2} + \sum_{x_B \in \partial B} f^P_{x_B} \epsilon^{x_B} \right) \\ & \left(\sum_{x_1' \in \Lambda_1} c^Q_{x_1'} \epsilon^{x_1'}_{j_1} + \sum_{x_2' \in \Lambda_2} d^Q_{x_2'} \epsilon^{x_2'}_{j_2} + \sum_{x_B' \in \partial B} f^Q_{x_B'} \epsilon^{x_B'} \right) \left(\sum_{x_1'' \in \Lambda_1} c^R_{x_1''} \epsilon^{x_1''}_{j_1'} + \sum_{x_2'' \in \Lambda_2} d^R_{x_2''} \epsilon^{x_2''}_{j_2'} + \sum_{x_B'' \in \partial B} f^R_{x_B} \epsilon^{x_B''} \right) \\ & \left(\sum_{x_1''' \in \Lambda_1} c^S_{x_1'''} \epsilon^{x_1'''}_{j_1'} + \sum_{x_2''' \in \Lambda_2} d^S_{x_2'''} \epsilon^{x_2'''}_{j_2'} + \sum_{x_B''' \in \partial B} f^S_{x_B'''} \epsilon^{x_B'''} \right) \Bigg\rangle. \end{aligned} \tag{C.2}$$

The only non-zero combinations of ϵ values are of the form $\langle \epsilon\epsilon\epsilon'\epsilon' \rangle = \langle \epsilon\epsilon \rangle \langle \epsilon'\epsilon' \rangle = 1$ and $\langle \epsilon\epsilon\epsilon\epsilon \rangle = 3$. Two ϵ random variables are only the same if they have the same x coordinate and j indices. For example:

$$\begin{aligned} \langle \epsilon^{x_1}_{j_1} \epsilon^{x_2}_{j_1} \rangle &= \delta^{x_1 x_2}, \\ \langle \epsilon^{x_1}_{j_1} \epsilon^{x_2}_{j_1'} \rangle &= \delta^{x_1 x_2} \delta_{j_1 j_1'}. \end{aligned} \tag{C.3}$$

There are two distinct terms that are non-zero. The first involves four coefficients of the same type, e.g. four c's, four d's or four f's. We call this contribution G. The second

non-zero contribution involves pairs of coefficients, e.g. two c's and two d's for example: call this contribution H.

$$Z = G + H \tag{C.4}$$

Considering first G, we have

$$\begin{aligned} G = \frac{1}{M^4} \sum_{j_1,j_2,j_1',j_2'=1}^{M} \Bigg(& \sum_{x_1,x_1',x_1'',x_1'''\in\Lambda_1} c^P_{x_1} c^Q_{x_1'} c^R_{x_1''} c^S_{x_1'''} \langle \epsilon^{x_1}_{j_1} \epsilon^{x_1'}_{j_1'} \epsilon^{x_1''}_{j_1'} \epsilon^{x_1'''}_{j_1'} \rangle \\ & + \sum_{x_2,x_2',x_2'',x_2'''\in\Lambda_2} d^P_{x_2} d^Q_{x_2'} d^R_{x_2''} d^S_{x_2'''} \langle \epsilon^{x_2}_{j_2} \epsilon^{x_2'}_{j_2'} \epsilon^{x_2''}_{j_2'} \epsilon^{x_2'''}_{j_2'} \rangle \\ & + \sum_{x_B,x_B',x_B'',x_B'''\in\partial B} f^P_{x_B} f^Q_{x_B'} f^R_{x_B''} f^S_{x_B'''} \langle \epsilon^{x_B} \epsilon^{x_B'} \epsilon^{x_B''} \epsilon^{x_B'''} \rangle \Bigg) \end{aligned} \tag{C.5}$$

$$\begin{aligned} = \sum_j \Bigg(& \sum_{\Lambda_1} c^P_{x_1} c^Q_{x_1'} c^R_{x_1''} c^S_{x_1'''} \langle \epsilon^{x_1}_{j_1} \epsilon^{x_1'}_{j_1'} \epsilon^{x_1''}_{j_1'} \epsilon^{x_1'''}_{j_1'} \rangle + \sum_{\Lambda_2} d^P_{x_2} d^Q_{x_2'} d^R_{x_2''} d^S_{x_2'''} \langle \epsilon^{x_2}_{j_2} \epsilon^{x_2'}_{j_2'} \epsilon^{x_2''}_{j_2'} \epsilon^{x_2'''}_{j_2'} \rangle \\ & + \sum_{\partial_B} f^P_{x_B} f^Q_{x_B'} f^R_{x_B''} f^S_{x_B'''} \langle \epsilon^{x_B} \epsilon^{x_B'} \epsilon^{x_B''} \epsilon^{x_B'''} \rangle \Bigg), \end{aligned}$$

where we have used the abbreviations

$$\begin{aligned} \sum_j &= \frac{1}{M^4} \sum_{j_1,j_2,j_1',j_2'=1}^{M}, \\ \sum_{\Lambda_1} &= \sum_{x_1,x_1',x_1'',x_1'''\in\Lambda_1}, \\ \sum_{\Lambda_2} &= \sum_{x_2,x_2',x_2'',x_2'''\in\Lambda_2}, \\ \sum_{\partial_B} &= \sum_{x_B,x_B',x_B'',x_B'''\in\partial B}. \end{aligned} \tag{C.6}$$

The simplest term in the above is the one involving boundary contributions. We evaluate the expectation value,

$$\begin{aligned} \langle \epsilon^{x_B} \epsilon^{x_B'} \epsilon^{x_B''} \epsilon^{x_B'''} \rangle &= 3\delta^{x_B x_B'} \delta^{x_B x_B''} \delta^{x_B x_B'''} + \delta^{x_B x_B'} \delta^{x_B'' x_B'''} (1 - \delta^{x_B x_B''}) \\ &\quad + \delta^{x_B x_B''} \delta^{x_B' x_B'''} (1 - \delta^{x_B x_B'}) + \delta^{x_B x_B'''} \delta^{x_B' x_B''} (1 - \delta^{x_B x_B''}), \\ &= \delta^{x_B x_B'} \delta^{x_B'' x_B'''} + \delta^{x_B x_B''} \delta^{x_B' x_B'''} + \delta^{x_B x_B'''} \delta^{x_B' x_B''} \end{aligned} \tag{C.7}$$

giving

$$\sum_j \sum_{\partial B} f^P_{x_B} f^Q_{x'_B} f^R_{x''_B} f^S_{x'''_B} \langle \epsilon^{x_B} \epsilon^{x'_B} \epsilon^{x''_B} \epsilon^{x'''_B} \rangle = \sum_j \sum_{\partial B} f^P_{x_B} f^Q_{x'_B} f^R_{x''_B} f^S_{x'''_B} (\tag{C.8}$$

$$\delta^{x_B x'_B} \delta^{x''_B x'''_B} + \delta^{x_B x''_B} \delta^{x'_B x'''_B} + \delta^{x_B x'''_B} \delta^{x'_B x''_B}),$$

$$= \sum_{x_B, x''_B \in \partial B} (f^P_{x_B} f^Q_{x_B} f^R_{x'_B} f^S_{x'_B} + f^P_{x_B} f^Q_{x'_B} f^R_{x_B} f^S_{x'_B} + f^P_{x_B} f^Q_{x'_B} f^R_{x'_B} f^S_{x_B}) \tag{C.9}$$

where in the last line we relabelled x''_B to x'_B and performed the sum over j, which gave a simple factor of 1, since there are no j indices in the expression. The other terms are a little more complicated since we have to consider the impact of the j-indices. This causes us to pick up factors of $\delta_{j_1 j'_1}$ and $(1-\delta_{j_1 j'_1})$. When all combinations are considered we get

$$\begin{aligned} \langle \epsilon^{x_1}_{j_1} \epsilon^{x'_1}_{j_1} \epsilon^{x''_1}_{j'_1} \epsilon^{x'''_1}_{j'_1} \rangle &= 3\delta^{x_1 x'_1} \delta^{x_1 x''_1} \delta^{x_1 x'''_1} \delta_{j_1 j'_1} + \delta^{x_1 x'_1} \delta^{x_1 x''_1} \delta^{x_1 x'''_1} (1 - \delta_{j_1 j'_1}) \\ &\quad + \delta^{x_1 x'_1} \delta^{x''_1 x'''_1} (1 - \delta^{x_1 x''_1}) + (\delta^{x_1 x''_1} \delta^{x'_1 x'''_1} + \delta^{x_1 x'''_1} \delta^{x'_1 x''_1})(1 - \delta^{x_1 x'_1}) \delta_{j_1 j'_1}, \\ &= \delta^{x_1 x'_1} \delta^{x''_1 x'''_1} + (\delta^{x_1 x''_1} \delta^{x'_1 x'''_1} + \delta^{x_1 x'''_1} \delta^{x'_1 x''_1}) \delta_{j_1 j'_1}, \end{aligned} \tag{C.10}$$

giving

$$\sum_j \sum_{\Lambda_1} c^P_{x_1} c^Q_{x'_1} c^R_{x''_1} c^S_{x'''_1} \langle \epsilon^{x_1}_{j_1} \epsilon^{x'_1}_{j_1} \epsilon^{x''_1}_{j'_1} \epsilon^{x'''_1}_{j'_1} \rangle = \sum_j \sum_{\Lambda_1} c^P_{x_1} c^Q_{x'_1} c^R_{x''_1} c^S_{x'''_1} [\tag{C.11}$$

$$\delta^{x_1 x'_1} \delta^{x''_1 x'''_1} + (\delta^{x_1 x''_1} \delta^{x'_1 x'''_1} + \delta^{x_1 x'''_1} \delta^{x'_1 x''_1}) \delta_{j_1 j'_1}].$$

Noting that $\sum_j 1 = 1$ and $\sum_j \delta_{j_1 j'_1} = 1/M$, we can evaluate the j sum,

$$\sum_j \sum_{\Lambda_1} c^P_{x_1} c^Q_{x'_1} c^R_{x''_1} c^S_{x'''_1} \langle \epsilon^{x_1}_{j_1} \epsilon^{x'_1}_{j_1} \epsilon^{x''_1}_{j'_1} \epsilon^{x'''_1}_{j'_1} \rangle = \left(c^P_{x_1} c^Q_{x_1} c^R_{x'_1} c^S_{x'_1} + \frac{1}{M} (c^P_{x_1} c^Q_{x'_1} c^R_{x_1} c^S_{x'_1} + c^P_{x_1} c^Q_{x'_1} c^R_{x'_1} c^S_{x_1}) \right), \tag{C.12}$$

where we relabelled the x's to group them under the same sum. We can extend the definitions of c, d and f as follows:

$$c_x = \begin{cases} c_{x_1} & x \in \Lambda_1 \\ 0 & otherwise \end{cases}, \tag{C.13}$$

$$d_x = \begin{cases} d_{x_2} & x \in \Lambda_2 \\ 0 & otherwise \end{cases},$$

$$f_x = \begin{cases} f_{x_B} & x \in \partial B \\ 0 & otherwise \end{cases}.$$

Evaluate the terms involving sites in Λ_2 similarly to those involving Λ_1 gives the overall expression

$$\begin{aligned} G = \sum_{x,x'\in\Lambda} \Big(& f^P_x f^Q_x f^R_{x'} f^S_{x'} + f^P_x f^Q_{x'} f^R_x f^S_{x'} + f^P_x f^Q_{x'} f^R_{x'} f^S_x) \\ & + c^P_x c^Q_x c^R_{x'} c^S_{x'} + \frac{1}{M}(c^P_x c^Q_{x'} c^R_x c^S_{x'} + c^P_x c^Q_{x'} c^R_{x'} c^S_x) \\ & + d^P_x d^Q_x d^R_{x'} d^S_{x'} + \frac{1}{M}(d^P_x d^Q_{x'} d^R_x d^S_{x'} + d^P_x d^Q_{x'} d^R_{x'} d^S_x) \Big) \end{aligned} \tag{C.14}$$

The remaining contributions involve all combinations of c, d and f pairs (six combinations) with all variations of x parameters (three combinations). There are therefore 18 terms which have been expressed in full below:

$$\begin{aligned} H = & \sum_j \sum_{x_1,x_1'\in\Lambda_1} c^P_{x_1} c^Q_{x_1'} \langle \epsilon^{x_1}_{j_1} \epsilon^{x_1'}_{j_1} \rangle \Big(\sum_{x_2'',x_2'''\in\Lambda_2} d^R_{x_2''} d^S_{x_2'''} \langle \epsilon^{x_2''}_{j_2'} \epsilon^{x_2'''}_{j_2'} \rangle + \sum_{x_B'',x_B'''\in\partial B} f^R_{x_B''} f^S_{x_B'''} \langle \epsilon^{x_B''} \epsilon^{x_B'''} \rangle \Big) \\ & + \sum_j \sum_{x_1,x_1''\in\Lambda_1} c^P_{x_1} c^R_{x_1''} \langle \epsilon^{x_1}_{j_1} \epsilon^{x_1''}_{j_1'} \rangle \Big(\sum_{x_2',x_2'''\in\Lambda_2} d^Q_{x_2'} d^S_{x_2'''} \langle \epsilon^{x_2'}_{j_2} \epsilon^{x_2'''}_{j_2'} \rangle + \sum_{x_B',x_B'''\in\partial B} f^Q_{x_B'} f^S_{x_B'''} \langle \epsilon^{x_B'} \epsilon^{x_B'''} \rangle \Big) \\ & + \sum_j \sum_{x_1,x_1'\in\Lambda_1} c^P_{x_1} c^S_{x_1'''} \langle \epsilon^{x_1}_{j_1} \epsilon^{x_1'''}_{j_1'} \rangle \Big(\sum_{x_2'',x_2'\in\Lambda_2} d^R_{x_2''} d^Q_{x_2'} \langle \epsilon^{x_2''}_{j_2'} \epsilon^{x_2'}_{j_2} \rangle + \sum_{x_B'',x_B'\in\partial B} f^R_{x_B''} f^Q_{x_B'} \langle \epsilon^{x_B''} \epsilon^{x_B'} \rangle \Big) \\ & + \sum_j \sum_{x_2,x_2'\in\Lambda_2} d^P_{x_2} d^Q_{x_2'} \langle \epsilon^{x_2}_{j_2} \epsilon^{x_2'}_{j_2} \rangle \Big(\sum_{x_1'',x_1'''\in\Lambda_1} c^R_{x_1''} c^S_{x_1'''} \langle \epsilon^{x_1''}_{j_1'} \epsilon^{x_1'''}_{j_1'} \rangle + \sum_{x_B'',x_B'''\in\partial B} f^R_{x_B''} f^S_{x_B'''} \langle \epsilon^{x_B''} \epsilon^{x_B'''} \rangle \Big) \\ & + \sum_j \sum_{x_2,x_2''\in\Lambda_2} d^P_{x_2} d^R_{x_2''} \langle \epsilon^{x_2}_{j_2} \epsilon^{x_2''}_{j_2'} \rangle \Big(\sum_{x_1',x_1'''\in\Lambda_1} c^Q_{x_1'} c^S_{x_1'''} \langle \epsilon^{x_1'}_{j_1} \epsilon^{x_1'''}_{j_1'} \rangle + \sum_{x_B',x_B'''\in\partial B} f^Q_{x_B'} f^S_{x_B'''} \langle \epsilon^{x_B'} \epsilon^{x_B'''} \rangle \Big) \\ & + \sum_j \sum_{x_2,x_2'''\in\Lambda_2} d^P_{x_2} d^S_{x_2'''} \langle \epsilon^{x_2}_{j_2} \epsilon^{x_2'''}_{j_2'} \rangle \Big(\sum_{x_1'',x_1'''\in\Lambda_1} c^R_{x_1''} c^Q_{x_1'} \langle \epsilon^{x_1''}_{j_1'} \epsilon^{x_1'}_{j_1} \rangle + \sum_{x_B'',x_B'\in\partial B} f^R_{x_B''} f^Q_{x_B'} \langle \epsilon^{x_B''} \epsilon^{x_B'} \rangle \Big) \\ & + \sum_j \sum_{x_B,x_B'\in\partial B} f^P_{x_B} f^Q_{x_B'} \langle \epsilon^{x_B} \epsilon^{x_B'} \rangle \Big(\sum_{x_1'',x_1'''\in\Lambda_1} c^R_{x_1''} c^S_{x_1'''} \langle \epsilon^{x_1''}_{j_1'} \epsilon^{x_1'''}_{j_1'} \rangle + \sum_{x_2'',x_2'''\in\Lambda_2} d^R_{x_2''} d^S_{x_2'''} \langle \epsilon^{x_2''}_{j_2'} \epsilon^{x_2'''}_{j_2'} \rangle \Big) \\ & + \sum_j \sum_{x_B,x_B'\in\partial B} f^P_{x_B} f^R_{x_B''} \langle \epsilon^{x_B} \epsilon^{x_B''} \rangle \Big(\sum_{x_1',x_1'''\in\Lambda_1} c^Q_{x_1'} c^S_{x_1'''} \langle \epsilon^{x_1'}_{j_1} \epsilon^{x_1'''}_{j_1'} \rangle + \sum_{x_2',x_2'''\in\Lambda_2} d^Q_{x_2'} d^S_{x_2'''} \langle \epsilon^{x_2'}_{j_2} \epsilon^{x_2'''}_{j_2'} \rangle \Big) \\ & + \sum_j \sum_{x_B,x_B''\in\partial B} f^P_{x_B} f^S_{x_B'''} \langle \epsilon^{x_B} \epsilon^{x_B'''} \rangle \Big(\sum_{x_1'',x_1'''\in\Lambda_1} c^R_{x_1''} c^Q_{x_1'} \langle \epsilon^{x_1''}_{j_1'} \epsilon^{x_1'}_{j_1} \rangle + \sum_{x_2'',x_2'\in\Lambda_2} d^R_{x_2''} d^Q_{x_2'} \langle \epsilon^{x_2''}_{j_2'} \epsilon^{x_2'}_{j_2} \rangle \Big). \end{aligned}$$

Evaluating the expectation values gives

$$
\begin{aligned}
H = &\sum_j \sum_{x_1,x_1'\in\Lambda_1} c^P_{x_1} c^Q_{x_1'} \delta^{x_1 x_1'} \Bigg(\sum_{x_2'',x_2'''\in\Lambda_2} d^R_{x_2''} d^S_{x_2'''} \delta^{x_2'' x_2'''} + \sum_{x_B'',x_B'''\in\partial B} f^R_{x_B''} f^S_{x_B'''} \delta^{x_B'' x_B'''} \Bigg) \\
&+ \sum_j \sum_{x_1,x_1''\in\Lambda_1} c^P_{x_1} c^R_{x_1''} \delta^{x_1 x_1''} \delta_{j_1 j_1'} \Bigg(\sum_{x_2',x_2'''\in\Lambda_2} d^Q_{x_2'} d^S_{x_2'''} \delta^{x_2' x_2'''} \delta_{j_2 j_2'} + \sum_{x_B',x_B'''\in\partial B} f^Q_{x_B'} f^S_{x_B'''} \delta^{x_B' x_B'''} \Bigg) \\
&+ \sum_j \sum_{x_1,x_1'\in\Lambda_1} c^P_{x_1} c^S_{x_1'''} \delta^{x_1 x_1'''} \delta_{j_1 j_1'} \Bigg(\sum_{x_2'',x_2'\in\Lambda_2} d^R_{x_2''} d^Q_{x_2'} \delta^{x_2' x_2''} \delta_{j_2 j_2'} + \sum_{x_B'',x_B'\in\partial B} f^R_{x_B''} f^Q_{x_B'} \delta^{x_B' x_B''} \Bigg) \\
&+ \sum_j \sum_{x_2,x_2'\in\Lambda_2} d^P_{x_2} d^Q_{x_2'} \delta^{x_2 x_2'} \Bigg(\sum_{x_1'',x_1'''\in\Lambda_1} c^R_{x_1''} c^S_{x_1'''} \delta^{x_1'' x_1'''} + \sum_{x_B'',x_B'''\in\partial B} f^R_{x_B''} f^S_{x_B'''} \delta^{x_B'' x_B'''} \Bigg) \\
&+ \sum_j \sum_{x_2,x_2''\in\Lambda_2} d^P_{x_2} d^R_{x_2''} \delta^{x_2 x_2''} \delta_{j_2 j_2'} \Bigg(\sum_{x_1',x_1'''\in\Lambda_1} c^Q_{x_1'} c^S_{x_1'''} \delta^{x_1' x_1'''} \delta_{j_1 j_1'} + \sum_{x_B',x_B'''\in\partial B} f^Q_{x_B'} f^S_{x_B'''} \delta^{x_B' x_B'''} \Bigg) \\
&+ \sum_j \sum_{x_2,x_2'''\in\Lambda_2} d^P_{x_2} d^S_{x_2'''} \delta^{x_2 x_2'''} \delta_{j_2 j_2'} \Bigg(\sum_{x_1'',x_1'''\in\Lambda_1} c^R_{x_1''} c^Q_{x_1'} \delta^{x_1' x_1''} \delta_{j_1 j_1'} + \sum_{x_B'',x_B'\in\partial B} f^R_{x_B''} f^Q_{x_B'} \delta^{x_B' x_B''} \Bigg) \\
&+ \sum_j \sum_{x_B,x_B'\in\partial B} f^P_{x_B} f^Q_{x_B'} \delta^{x_B x_B'} \Bigg(\sum_{x_1'',x_1'''\in\Lambda_1} c^R_{x_1''} c^S_{x_1'''} \delta^{x_1'' x_1'''} + \sum_{x_2'',x_2'''\in\Lambda_2} d^R_{x_2''} d^S_{x_2'''} \delta^{x_2'' x_2'''} \Bigg) \\
&+ \sum_j \sum_{x_B,x_B''\in\partial B} f^P_{x_B} f^R_{x_B''} \delta^{x_B x_B''} \Bigg(\sum_{x_1',x_1'''\in\Lambda_1} c^Q_{x_1'} c^S_{x_1'''} \delta^{x_1' x_1'''} \delta_{j_1 j_1'} + \sum_{x_2',x_2'''\in\Lambda_2} d^Q_{x_2'} d^S_{x_2'''} \delta^{x_2' x_2'''} \delta_{j_2 j_2'} \Bigg) \\
&+ \sum_j \sum_{x_B,x_B'''\in\partial B} f^P_{x_B} f^S_{x_B'''} \delta^{x_B x_B'''} \Bigg(\sum_{x_1'',x_1'''\in\Lambda_1} c^R_{x_1''} c^Q_{x_1'} \delta^{x_1' x_1''} \delta_{j_1 j_1'} + \sum_{x_2'',x_2'\in\Lambda_2} d^R_{x_2''} d^Q_{x_2'} \delta^{x_2' x_2''} \delta_{j_2 j_2'} \Bigg).
\end{aligned}
$$

Evaluating the sums over j- and x-indices, gives

$$
\begin{aligned}
H = \sum_{x,x'\in\Lambda} \Bigg[& c^P_x c^Q_x \left(d^R_{x'} d^S_{x'} + f^R_{x'} f^S_{x'} \right) + \frac{1}{M} c^P_x c^R_x \left(\frac{1}{M} d^Q_{x'} d^S_{x'} + f^Q_{x'} f^S_{x'} \right) \\
&+ \frac{1}{M} c^P_x c^S_x \left(\frac{1}{M} d^R_{x'} d^Q_{x'} + f^R_{x'} f^Q_{x'} \right) + d^P_x d^Q_x \left(c^R_{x'} d^S_{x'} + f^R_{x'} f^S_{x'} \right) + \frac{1}{M} d^P_x d^R_x \left(\frac{1}{M} c^Q_{x'} c^S_{x'} + f^Q_x f^S_{x'} \right) \\
&+ \frac{1}{M} d^P_x d^S_x \left(\frac{1}{M} c^R_{x'} c^Q_{x'} + f^R_{x'} f^Q_{x'} \right) \Bigg] + f^P_x f^Q_x \left(c^R_{x'} c^S_{x'} + d^R_{x'} d^S_{x'} \right) + f^P_x f^R_x \left(\frac{1}{M} c^Q_{x'} c^S_{x'} + \frac{1}{M} d^Q_{x'} d^S_{x'} \right) \\
&+ f^P_x f^S_x \left(\frac{1}{M} c^R_{x'} c^Q_{x'} + \frac{1}{M} d^R_{x'} d^Q_{x'} \right) \Bigg].
\end{aligned}
\tag{C.15}
$$

We are now in a position to combine terms,

$$
Z = A' + \frac{B'}{M} + \frac{C'}{M^2},
\tag{C.16}
$$

where

$$
\begin{aligned}
A' = \sum_{x,x'} \Bigg[& \left(c_x^P c_x^Q c_{x'}^R c_{x'}^S + d_x^P d_x^Q d_{x'}^R d_{x'}^S + f_x^P f_x^Q f_{x'}^R f_{x'}^S + f_x^P f_{x'}^Q f_x^R f_{x'}^S + f_x^P f_{x'}^Q f_{x'}^R f_x^S \right) \\
& + c_x^P c_x^Q \left(d_{x'}^R d_{x'}^S + f_{x'}^R f_{x'}^S \right) + d_x^P d_x^Q \left(c_{x'}^R c_{x'}^S + f_{x'}^R f_{x'}^S \right) + f_x^P f_x^Q \left(c_{x'}^R c_{x'}^S + d_{x'}^R d_{x'}^S \right) \Bigg], \\
= & \sum_{x,x'} \left[f_x^P f_{x'}^Q f_x^R f_{x'}^S + f_x^P f_{x'}^Q f_{x'}^R f_x^S \right] \\
& + \sum_{x} \left[c_x^P c_x^Q + d_x^P d_x^Q + f_x^P f_x^Q \right] \sum_{x'} \left[c_{x'}^R c_{x'}^S + d_{x'}^R d_{x'}^S + f_{x'}^R f_{x'}^S \right], \\
B' = \sum_{x,x'} \Big[& c_x^P c_{x'}^Q c_x^R c_{x'}^S + c_x^P c_{x'}^Q c_{x'}^R c_x^S + d_x^P d_{x'}^Q d_x^R d_{x'}^S + d_x^P d_{x'}^Q d_{x'}^R d_x^S \\
& + (c_x^P f_{x'}^Q + f_{x'}^P c_x^Q)(c_x^R f_{x'}^S + f_{x'}^R c_x^S) + (d_x^P f_{x'}^Q + f_{x'}^P d_x^Q)(d_x^R f_{x'}^S + f_{x'}^R d_x^S), \\
C' = \sum_{x,x'} \Big[& c_x^P d_{x'}^Q c_x^R d_{x'}^S + c_x^P d_{x'}^Q d_{x'}^R c_x^S + d_{x'}^P c_x^Q c_x^R d_{x'}^S + d_{x'}^P c_x^Q c_x^S d_{x'}^R \Big].
\end{aligned}
\tag{C.17}
$$

We have rearranged A' by pulling out a bracketed product which is exactly the product of the expectation of two-point correlators (eq. (6.55)), e.g.

$$
\frac{1}{M^4} \sum_{j_1} \sum_{j_2} \sum_{j_1'} \sum_{j_2'} \left(\langle \varphi^P_{j_1 j_2} \varphi^Q_{j_1 j_2} \varphi^R_{j_1' j_2'} \varphi^S_{j_1' j_2'} \rangle - \langle \varphi^P_{j_1 j_2} \varphi^Q_{j_1 j_2} \rangle \langle \varphi^P_{j_1 j_2} \varphi^Q_{j_1 j_2} \rangle \right) = A + \frac{B}{M} + \frac{C}{M^2},
\tag{C.18}
$$

where

$$
\begin{aligned}
A = & \sum_{x,x'} \left[f_x^P f_{x'}^Q f_x^R f_{x'}^S + f_x^P f_{x'}^Q f_{x'}^R f_x^S \right], \\
B = & \sum_{x,x'} \Big[c_x^P c_{x'}^Q c_x^R c_{x'}^S + c_x^P c_{x'}^Q c_{x'}^R c_x^S + d_x^P d_{x'}^Q d_x^R d_{x'}^S + d_x^P d_{x'}^Q d_{x'}^R d_x^S \\
& + (c_x^P f_{x'}^Q + f_{x'}^P c_x^Q)(c_x^R f_{x'}^S + f_{x'}^R c_x^S) + (d_x^P f_{x'}^Q + f_{x'}^P d_x^Q)(d_x^R f_{x'}^S + f_{x'}^R d_x^S) \Big], \\
C = & \sum_{x,x'} (c_x^P d_{x'}^Q + d_{x'}^P c_x^Q)(c_x^R d_{x'}^S + d_{x'}^R c_x^S).
\end{aligned}
$$

Appendix D

Slice-Coordinate Two-Point Functions

In this appendix we derive equation (6.22).

The momentum space two-point correlator is given by

$$\begin{aligned}
\tilde{C}_2(p) &= \frac{1}{V}\sum_z e^{-ip\cdot z} C_2(z) \qquad \text{(D.1)}\\
&= \frac{1}{V}\sum_z e^{-ip\cdot z}\langle\phi(z)\phi(0)\rangle\\
&= \frac{1}{V^2}\sum_x\sum_z e^{-ip\cdot z}\langle\phi(z+x)\phi(x)\rangle\\
&= \frac{1}{V^2}\sum_x\sum_y e^{-ip\cdot(\mathbf{y}-\mathbf{x})}\langle\phi(y)\phi(x)\rangle\\
&= \langle\left(\frac{1}{V}\sum_x e^{+ip\cdot x}\phi(x)\right)\left(\frac{1}{V}\sum_y e^{-ip\cdot y}\phi(y)\right)\rangle,
\end{aligned}$$

where $C_2(x)$ is the position space two-point correlator and we use the change of variables $y = z + x$ and translational invariance. Setting the momentum to zero in all but one direction, we find

$$\begin{aligned}
\tilde{C}_2(0,0,...,0,p) = \langle&\left(\frac{1}{V}\sum_x e^{+ipx}\sum_{x_1,x_2,...,x_{d-1}}\phi(x_1,x_2,...,x_{d-1},x)\right)\times \qquad \text{(D.2)}\\
&\left(\frac{1}{V}\sum_y e^{-ipy}\sum_{y_1,y_2,...,y_{d-1}}\phi(y_1,y_2,...,y_{d-1},y)\right)\rangle \qquad \text{(D.3)}\\
&= \langle\left(\frac{1}{L}\sum_x e^{+ipx}\Phi(x)\right)\left(\frac{1}{L}\sum_y e^{-ipy}\Phi(y)\right)\rangle. \qquad \text{(D.4)}
\end{aligned}$$

We can perform some rearrangement of the sums to express this in terms of the slice-coordinate two-point function:

$$\begin{aligned}
\tilde{C}_2(0,0,...,0,p) &= \langle \left(\frac{1}{L} \sum_x e^{+ipx} \Phi(x) \right) \left(\frac{1}{L} \sum_y e^{-ipy} \Phi(y) \right) \rangle \\
&= \langle \left(\frac{1}{L^2} \sum_x \sum_y e^{+ip(x-y)} \Phi(x)\Phi(y) \right) \rangle \\
&= \langle \left(\frac{1}{L^2} \sum_x \sum_z e^{-ipz} \Phi(x)\Phi(z+x) \right) \rangle \\
&= \langle \left(\frac{1}{L} \sum_z e^{-ipz} \Phi(0)\Phi(z) \right) \rangle \\
&= \frac{1}{L} \sum_z e^{-ipz} c_2(z),
\end{aligned} \tag{D.5}$$

where we have used translational invariance and the change of variables $y = z+x$. Lastly, we can perform a Fourier Transform to both sides of this equation get an expression for c_2,

$$\begin{aligned}
\sum_p e^{+ipx} \tilde{C}_2(0,0,...,0,p) &= \sum_p \sum_z e^{+ip(x-z)} c_2(z) \\
&= \frac{1}{L} \sum_z \sum_p e^{+ip(x-z)} c_2(z) \\
&= \frac{1}{L} \sum_z L \delta_{xz} c_2(z) \\
&= c_2(x).
\end{aligned}$$

Thus, equation (6.22) has been proven.

References

[1] Guido Cossu, Luigi Del Debbio, Andreas Jüttner, Ben Kitching-Morley, Joseph K. L. Lee, Antonin Portelli, Henrique Bergallo Rocha, and Kostas Skenderis and. Nonperturbative infrared finiteness in a superrenormalizable scalar quantum field theory. *Physical Review Letters*, 126(22), jun 2021. doi: 10.1103/physrevlett.126.221601. URL https://doi.org/10.1103%2Fphysrevlett.126.221601.

[2] Luigi Del Debbio, Elizabeth Dobson, Andreas Jüttner, Ben Kitching-Morley, Joseph K. L. Lee, Valentin Nourry, Antonin Portelli, Henrique Bergallo Rocha, and Kostas Skenderis. Renormalisation of the energy-momentum tensor in three-dimensional scalar $SU(N)$ theories using the Wilson flow. *Phys. Rev. D*, 103(11): 114501, 2021. doi: 10.1103/PhysRevD.103.114501.

[3] Ben Kitching-Morley and Andreas Jüttner. A numerical and theoretical study of multilevel performance for two-point correlator calculations. *PoS*, LATTICE2021: 133, 2022. doi: 10.22323/1.396.0133.

[4] Joseph K. L. Lee, Elizabeth Dobson, Andreas Jüttner, Ben Kitching-Morley, Valentin Nourry, Antonin Portelli, Henrique Bergallo Rocha, Kostas Skenderis, and Luigi Del Debbio. Renormalisation of the 3D SU(N) scalar energy-momentum tensor using the Wilson flow. *PoS*, LATTICE2021:313, 2022. doi: 10.22323/1.396.0313.

[5] Henrique Bergallo Rocha, Luigi Del Debbio, Elizabeth Dobson, Andreas Jüttner, Ben Kitching-Morley, Joseph K. L. Lee, Antonin Portelli, and Kostas Skenderis. Critical behaviour and phase structure of 3d Scalar+Gauge Field Theories in the adjoint representation. *PoS*, LATTICE2021:270, 2022. doi: 10.22323/1.396.0270.

[6] Ben Kitching-Morley and Andreas Jüttner. andreasjuettner/finite-size-scaling-analysis: V1.1 – code release, https://doi.org/10.5281/zenodo.4266114, November 2020. URL https://doi.org/10.5281/zenodo.4290508.

[7] Guido Cossu, Henrique Bergallo Rocha, Luigi Del Debbio, Andreas Jüttner, Ben Kitching-Morley, Joseph K.L. Lee, Antonin Portelli, and Kostas Skenderis. Non-perturbative infrared finiteness in super- renormalisable scalar quantum field theory – data release, November 2020. URL https://doi.org/10.5281/zenodo.4266114.

[8] Ben Kitching-Morley. Multievel notebook: V1.0 – code release, https://-doi.org/10.5281/zenodo.7311081, November 2022. URL https://doi.org/10.5281/zenodo.7311081.

[9] Ben Kitching-Morley. Multievel notebook: V1.0 – data release, https://-doi.org/10.5281/zenodo.7303520, November 2022. URL https://doi.org/10.5281/zenodo.7303520.

[10] William I. Jay and Ethan T. Neil. Bayesian model averaging for analysis of lattice field theory results. *Phys. Rev. D*, 103:114502, 2021. doi: 10.1103/PhysRevD.103.114502.

[11] Alexander Friedman. Über die krümmung des raumes. *Zeitschrift für Physik*, 10 (1):377–386, 1922.

[12] Alexander Friedmann. Über die möglichkeit einer welt mit konstanter negativer krümmung des raumes. *Zeitschrift für Physik*, 21(1):326–332, 1924.

[13] Georges Lemaître. 126. a homogeneous universe of constant mass and increasing radius accounting for the radial velocity of extra-galactic nebulae. In *A Source Book in Astronomy and Astrophysics, 1900–1975*, pages 844–848. Harvard University Press, 2013.

[14] Georges Lemaître. L'univers en expansion. In *Annales de la Société scientifique de Bruxelles*, volume 53, page 51, 1933.

[15] Howard Percy Robertson. Kinematics and world-structure. *The Astrophysical Journal*, 82:284, 1935.

[16] Howard P Robertson. Kinematics and world-structure ii. *The Astrophysical Journal*, 83:187, 1936.

[17] Howard P Robertson. Kinematics and world-structure iii. *The Astrophysical Journal*, 83:257, 1936.

[18] Arthur Geoffrey Walker. On milne's theory of world-structure. *Proceedings of the London Mathematical Society*, 2(1):90–127, 1937.

[19] Edwin Hubble. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15(3):168–173, March 1929. doi: 10.1073/pnas.15.3.168.

[20] Alexander Friedman. Über die krümmung des raumes. *Zeitschrift für Physik*, 10 (1):377–386, 1922.

[21] Alexander Friedmann. Über die möglichkeit einer welt mit konstanter negativer krümmung des raumes. *Zeitschrift für Physik*, 21(1):326–332, 1924.

[22] Y. Akrami et al. Planck 2018 results. I. Overview and the cosmological legacy of Planck. 7 2018.

[23] Arno A Penzias and Robert Woodrow Wilson. A measurement of excess antenna temperature at 4080 mc/s. *The Astrophysical Journal*, 142:419–421, 1965.

[24] Robert H. Brandenberger. Inflationary cosmology: Progress and problems. In *IPM School on Cosmology 1999: Large Scale Structure Formation*, 1 1999.

[25] Alan H. Guth. The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems. *Adv. Ser. Astrophys. Cosmol.*, 3:139–148, 1987. doi: 10.1103/PhysRevD.23.347.

[26] Jerome Martin and Robert H. Brandenberger. The TransPlanckian problem of inflationary cosmology. *Phys. Rev. D*, 63:123501, 2001. doi: 10.1103/PhysRevD.63.123501.

[27] Arvind Borde and Alexander Vilenkin. Eternal inflation and the initial singularity. *Phys. Rev. Lett.*, 72:3305–3308, May 1994. doi: 10.1103/PhysRevLett.72.3305. URL https://link.aps.org/doi/10.1103/PhysRevLett.72.3305.

[28] Fred C. Adams, Katherine Freese, and Alan H. Guth. Constraints on the scalar-field potential in inflationary models. *Phys. Rev. D*, 43:965–976, Feb 1991. doi: 10.1103/PhysRevD.43.965. URL https://link.aps.org/doi/10.1103/PhysRevD.43.965.

[29] Gerard 't Hooft. Dimensional reduction in quantum gravity. *Conf. Proc. C*, 930308:284–296, 1993.

[30] Leonard Susskind. The World as a hologram. *J. Math. Phys.*, 36:6377–6396, 1995. doi: 10.1063/1.531249.

[31] Stephen W Hawking. Particle creation by black holes. In *Euclidean quantum gravity*, pages 167–188. World Scientific, 1975.

[32] Jacob D. Bekenstein. Entropy bounds and black hole remnants. *Physical Review D*, 49(4):1912–1921, feb 1994. doi: 10.1103/physrevd.49.1912. URL https://doi.org/10.1103%2Fphysrevd.49.1912.

[33] Juan Martin Maldacena. The Large N limit of superconformal field theories and supergravity. *Adv. Theor. Math. Phys.*, 2:231–252, 1998. doi: 10.1023/A:1026654312961.

[34] Ofer Aharony, Steven S. Gubser, Juan Maldacena, Hirosi Ooguri, and Yaron Oz. Large n field theories, string theory and gravity. *Physics Reports*, 323(3-4):183–386, jan 2000. doi: 10.1016/s0370-1573(99)00083-6. URL https://doi.org/10.1016%2Fs0370-1573%2899%2900083-6.

[35] Harm J Boonstra, Kostas Skenderis, and Paul K Townsend. The domain-wall/QFT correspondence. *Journal of High Energy Physics*, 1999(01):003–003, jan 1999. doi: 10.1088/1126-6708/1999/01/003. URL https://doi.org/10.1088%2F1126-6708%2F1999%2F01%2F003.

[36] Ofer Aharony, Steven S. Gubser, Juan Martin Maldacena, Hirosi Ooguri, and Yaron Oz. Large N field theories, string theory and gravity. *Phys. Rept.*, 323: 183–386, 2000. doi: 10.1016/S0370-1573(99)00083-6.

[37] S. S. Gubser, Igor R. Klebanov, and Alexander M. Polyakov. Gauge theory correlators from noncritical string theory. *Phys. Lett. B*, 428:105–114, 1998. doi: 10.1016/S0370-2693(98)00377-3.

[38] Edward Witten. Anti-de Sitter space and holography. *Adv. Theor. Math. Phys.*, 2:253–291, 1998. doi: 10.4310/ATMP.1998.v2.n2.a2.

[39] Kostas Skenderis and Paul K. Townsend. Hidden supersymmetry of domain walls and cosmologies. *Phys. Rev. Lett.*, 96:191301, 2006. doi: 10.1103/PhysRevLett.96.191301.

[40] D. S. Salopek and J. R. Bond. Nonlinear evolution of long wavelength metric fluctuations in inflationary models. *Phys. Rev. D*, 42:3936–3962, 1990. doi: 10.1103/PhysRevD.42.3936.

[41] Kostas Skenderis and Paul K. Townsend. Gravitational stability and renormalization group flow. *Phys. Lett. B*, 468:46–51, 1999. doi: 10.1016/S0370-2693(99)01212-5.

[42] O. DeWolfe, D. Z. Freedman, S. S. Gubser, and A. Karch. Modeling the fifth-dimension with scalars and gravity. *Phys. Rev. D*, 62:046008, 2000. doi: 10.1103/PhysRevD.62.046008.

[43] Jan de Boer, Erik P. Verlinde, and Herman L. Verlinde. On the holographic renormalization group. *JHEP*, 08:003, 2000. doi: 10.1088/1126-6708/2000/08/003.

[44] D. Z. Freedman, Carlos Nunez, M. Schnabl, and K. Skenderis. Fake supergravity and domain wall stability. *Phys. Rev. D*, 69:104027, 2004. doi: 10.1103/PhysRevD.69.104027.

[45] Paul McFadden and Kostas Skenderis. Holography for Cosmology. *Phys. Rev. D*, 81:021301, 2010. doi: 10.1103/PhysRevD.81.021301.

[46] Andrew Strominger. Inflation and the dS / CFT correspondence. *JHEP*, 11:049, 2001. doi: 10.1088/1126-6708/2001/11/049.

[47] Horatiu Nastase and Kostas Skenderis. Holography for the very early universe and the classic puzzles of hot big bang cosmology. *Physical Review D*, 101(2), jan 2020. doi: 10.1103/physrevd.101.021901. URL https://doi.org/10.1103%2Fphysrevd.101.021901.

[48] Ioannis Papadimitriou and Kostas Skenderis. Correlation functions in holographic RG flows. *Journal of High Energy Physics*, 2004(10):075–075, nov 2004. doi: 10.1088/1126-6708/2004/10/075. URL https://doi.org/10.1088%2F1126-6708%2F2004%2F10%2F075.

[49] P L McFadden and K Skenderis. The holographic universe. *Journal of Physics: Conference Series*, 222:012007, apr 2010. doi: 10.1088/1742-6596/222/1/012007. URL https://doi.org/10.1088%2F1742-6596%2F222%2F1%2F012007.

[50] C. M. Hull. Timelike T duality, de Sitter space, large N gauge theories and topological field theory. *JHEP*, 07:021, 1998. doi: 10.1088/1126-6708/1998/07/021.

[51] Edward Witten. Quantum gravity in de sitter space, 2001. URL https://arxiv.org/abs/hep-th/0106109.

[52] Andrew Strominger. The dS / CFT correspondence. *JHEP*, 10:034, 2001. doi: 10.1088/1126-6708/2001/10/034.

[53] Juan Martin Maldacena. Non-Gaussian features of primordial fluctuations in single field inflationary models. *JHEP*, 05:013, 2003. doi: 10.1088/1126-6708/2003/05/013.

[54] Jaume Garriga, Kostas Skenderis, and Yuko Urakawa. Multi-field inflation from holography. *JCAP*, 01:028, 2015. doi: 10.1088/1475-7516/2015/01/028.

[55] Kostas Skenderis. Nonperturbative and numerical approaches to quantum gravity, string theory and holography. ICTS Bangalore online lectures, 2021. URL https://www.youtube.com/watch?v=i_Xabl4WslY.

[56] Nima Arkani-Hamed and Juan Maldacena. Cosmological collider physics, 2015. URL https://arxiv.org/abs/1503.08043.

[57] Ingmar Kanitscheider, Kostas Skenderis, and Marika Taylor. Precision holography for non-conformal branes. *JHEP*, 09:094, 2008. doi: 10.1088/1126-6708/2008/09/094.

[58] Claudio Corianò, Luigi Delle Rose, and Kostas Skenderis. Two-point function of the energy-momentum tensor and generalised conformal structure. *Eur. Phys. J. C*, 81(2):174, 2021. doi: 10.1140/epjc/s10052-021-08892-5.

[59] Thomas Appelquist and Robert D. Pisarski. High-temperature yang-mills theories and three-dimensional quantum chromodynamics. *Phys. Rev. D*, 23:2305–2317, May 1981. doi: 10.1103/PhysRevD.23.2305. URL https://link.aps.org/doi/10.1103/PhysRevD.23.2305.

[60] R. Jackiw and S. Templeton. How super-renormalizable interactions cure their infrared divergences. *Phys. Rev. D*, 23:2291–2304, May 1981. doi: 10.1103/PhysRevD.23.2291. URL https://link.aps.org/doi/10.1103/PhysRevD.23.2291.

[61] Gerard 't Hooft. A Planar Diagram Theory for Strong Interactions. *Nucl. Phys. B*, 72:461, 1974. doi: 10.1016/0550-3213(74)90154-0.

[62] A. Rassat, J.-L. Starck, P. Paykari, F. Sureau, and J. Bobin. Planck CMB anomalies: astrophysical and cosmological secondary effects and the curse of masking. *Journal of Cosmology and Astroparticle Physics*, 2014(08):006–006, aug 2014. doi: 10.1088/1475-7516/2014/08/006. URL https://doi.org/10.1088/1475-7516/2014/08/006.

[63] Richard Easther, Raphael Flauger, Paul McFadden, and Kostas Skenderis. Constraining holographic inflation with WMAP. *JCAP*, 09:030, 2011. doi: 10.1088/1475-7516/2011/09/030.

[64] Niayesh Afshordi, Claudio Corianò, Luigi Delle Rose, Elizabeth Gould, and Kostas Skenderis. From planck data to planck era: Observational tests of holographic cosmology. *Phys. Rev. Lett.*, 118:041301, Jan 2017. doi: 10.1103/PhysRevLett.118.041301. URL https://link.aps.org/doi/10.1103/PhysRevLett.118.041301.

[65] Niayesh Afshordi, Elizabeth Gould, and Kostas Skenderis. Constraining holographic cosmology using planck data. *Phys. Rev. D*, 95:123505, Jun 2017. doi: 10.1103/PhysRevD.95.123505. URL https://link.aps.org/doi/10.1103/PhysRevD.95.123505.

[66] Niayesh Afshordi, Claudio Coriano, Luigi Delle Rose, Elizabeth Gould, and Kostas Skenderis. From planck data to planck era: Observational tests of holographic cosmology. 2016. URL https://arxiv.org/abs/1607.04878v1.

[67] Christof Gattringer and Christian B. Lang. *Quantum chromodynamics on the lattice*, volume 788. Springer, Berlin, 2010. ISBN 978-3-642-01849-7, 978-3-642-01850-3. doi: 10.1007/978-3-642-01850-3.

[68] Hale F Trotter. On the product of semi-groups of operators. *Proceedings of the American Mathematical Society*, 10(4):545–551, 1959.

[69] Ulli Wolff. Monte Carlo errors with less errors. *Comput. Phys. Commun.*, 156:143–153, 2004. doi: 10.1016/S0010-4655(03)00467-3. [Erratum: Comput.Phys.Commun. 176, 383 (2007)].

[70] Neal Madras and Alan D. Sokal. The Pivot algorithm: a highly efficient Monte Carlo method for selfavoiding walk. *J. Statist. Phys.*, 50:109–186, 1988. doi: 10.1007/BF01022990.

[71] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth. Hybrid Monte Carlo. *Phys. Lett. B*, 195:216–222, 1987. doi: 10.1016/0370-2693(87)91197-X.

[72] Stephen L. Adler. Overrelaxation algorithms for lattice field theories. *Phys. Rev. D*, 37:458–471, Jan 1988. doi: 10.1103/PhysRevD.37.458. URL https://link.aps.org/doi/10.1103/PhysRevD.37.458.

[73] Frank R. Brown and Thomas J. Woch. Overrelaxed heat-bath and metropolis algorithms for accelerating pure gauge monte carlo calculations. *Phys. Rev. Lett.*, 58:2394–2396, Jun 1987. doi: 10.1103/PhysRevLett.58.2394. URL https://link.aps.org/doi/10.1103/PhysRevLett.58.2394.

[74] B. Bunk. Montecarlo methods and results for the electroweak phase transition. *Nuclear Physics B - Proceedings Supplements*, 42(1):566–568, 1995. ISSN 0920-5632. doi: https://doi.org/10.1016/0920-5632(95)00313-X. URL https://www.sciencedirect.com/science/article/pii/092056329500313X.

[75] Zoltán Fodor and Karl Jansen. Overrelaxation algorithm for coupled gauge-higgs systems. *Physics Letters B*, 331(1):119–123, 1994. ISSN 0370-2693. doi: https://doi.org/10.1016/0370-2693(94)90951-2. URL https://www.sciencedirect.com/science/article/pii/0370269394909512.

[76] K. Fredenhagen and M. Marcu. A modified heat bath method suitable for monte carlo simulations on vector and parallel processors. *Physics Letters B*, 193(4):486–488, 1987. ISSN 0370-2693. doi: https://doi.org/10.1016/0370-2693(87)91703-5. URL https://www.sciencedirect.com/science/article/pii/0370269387917035.

[77] Zoltan Fodor and Karl Jansen. Overrelaxation algorithm for coupled gauge Higgs systems. *Phys. Lett. B*, 331:119–123, 1994. doi: 10.1016/0370-2693(94)90951-2.

[78] A. M. Ferrenberg and R. H. Swendsen. New Monte Carlo Technique for Studying Phase Transitions. *Phys. Rev. Lett.*, 61:2635–2638, 1988. doi: 10.1103/PhysRevLett.61.2635.

[79] Wolfhard Janke. Monte carlo methods in classical statistical physics. In *Computational many-particle physics*, pages 79–140. Springer, 2008.

[80] B. Efron. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1–26, 1979. doi: 10.2307/2958830. URL http://dx.doi.org/10.2307/2958830.

[81] Kenneth G. Wilson. Confinement of Quarks. *Phys. Rev. D*, 10:2445–2459, 1974. doi: 10.1103/PhysRevD.10.2445.

[82] István Montvay and Gernot Münster. *Quantum fields on a lattice*. Cambridge University Press, 1997.

[83] Martin Luscher and Peter Weisz. Coordinate space methods for the evaluation of Feynman diagrams in lattice field theories. *Nucl. Phys. B*, 445:429–450, 1995. doi: 10.1016/0550-3213(95)00185-U.

[84] M. Laine and A. Rajantie. Lattice continuum relations for 3-D SU(N) + Higgs theories. *Nucl. Phys. B*, 513:471–489, 1998. doi: 10.1016/S0550-3213(97)00709-8.

[85] K. Binder. Critical Properties from Monte Carlo Coarse Graining and Renormalization. *Phys. Rev. Lett.*, 47:693–696, 1981. doi: 10.1103/PhysRevLett.47.693.

[86] Kurt Binder. Finite size scaling analysis of ising model block distribution functions. *Zeitschrift für Physik B Condensed Matter*, 43(2):119–140, 1981.

[87] Peter Boyle, Azusa Yamaguchi, Guido Cossu, and Antonin Portelli. Grid: A next generation data parallel c++ qcd library, 2015. URL https://arxiv.org/abs/1512.03487.

[88] Peter A. Boyle, Guido Cossu, Azusa Yamaguchi, and Antonin Portelli. Grid: A next generation data parallel C++ QCD library. *PoS*, LATTICE 2015:023, 2016. doi: 10.22323/1.251.0023.

[89] S. S. SHAPIRO and M. B. WILK. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4):591–611, dec 1965. doi: 10.1093/biomet/52.3-4.591. URL https://doi.org/10.1093/biomet/52.3-4.591.

[90] Martin Hasenbusch. Three-dimensional $O(N)$-invariant ϕ^4 models at criticality for $N \geq 4$. 12 2021.

[91] Palash B Pal. *An introductory course of statistical mechanics*. Alpha Science International, 2008.

[92] Francesco Delfino, Andrea Pelissetto, and Ettore Vicari. Three-dimensional antiferromagnetic CP^{N-1} models. *Phys. Rev. E*, 91(5):052109, 2015. doi: 10.1103/PhysRevE.91.052109.

[93] Hirotugu Akaike. A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723, 1974.

[94] Hirotugu Akaike. A bayesian analysis of the minimum aic procedure. In *Selected Papers of Hirotugu Akaike*, pages 275–280. Springer, 1998.

[95] Hirotogu Akaike. Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotugu akaike*, pages 199–213. Springer, 1998.

[96] Harold Jeffreys. *The theory of probability*. OUP Oxford, 1998.

[97] F. Feroz, M.P. Hobson, and M. Bridges. MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Mon. Not. Roy. Astron. Soc.*, 398:1601–1614, 2009. doi: 10.1111/j.1365-2966.2009.14548.x.

[98] Farhan Feroz and Mike P Hobson. Multimodal nested sampling: an efficient and robust alternative to markov chain monte carlo methods for astronomical data analyses. *Monthly Notices of the Royal Astronomical Society*, 384(2):449–463, 2008.

[99] Farhan Feroz, Michael P Hobson, Ewan Cameron, and Anthony N Pettitt. Importance nested sampling and the multinest algorithm. *arXiv preprint arXiv:1306.2144*, 2013.

[100] J. Buchner, A. Georgakakis, K. Nandra, L. Hsu, C. Rangel, M. Brightman, A. Merloni, M. Salvato, J. Donley, and D. Kocevski. X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue. *Astron. Astrophys.*, 564:A125, 2014. doi: 10.1051/0004-6361/201322971.

[101] Edward E Leamer and Edward E Leamer. *Specification searches: Ad hoc inference with nonexperimental data*, volume 53. John Wiley & Sons Incorporated, 1978.

[102] A Racine, AP Grieve, H Fluhler, and AFM Smith. Bayesian methods in practice: experiences in the pharmaceutical industry. *Applied Statistics*, pages 93–150, 1986.

[103] Edward I George and Robert E McCulloch. Variable selection via gibbs sampling. *Journal of the American Statistical Association*, 88(423):881–889, 1993.

[104] Robert E Kass and Adrian E Raftery. Bayes factors. *Journal of the american statistical association*, 90(430):773–795, 1995.

[105] Adrian E Raftery and Yingye Zheng. Discussion: Performance of bayesian model averaging. *Journal of the American Statistical Association*, 98(464):931–938, 2003.

[106] Sz. Borsanyi et al. Ab initio calculation of the neutron-proton mass difference. *Science*, 347:1452–1455, 2015. doi: 10.1126/science.1257050.

[107] Martin Hasenbusch. Eliminating leading corrections to scaling in the three-dimensional O(N) symmetric phi**4 model: N=3 and N=4. *J. Phys. A*, 34: 8221–8236, 2001. doi: 10.1088/0305-4470/34/40/302.

[108] Joseph Kin Lok Lee. *Towards Lattice Simulations of Scalar Holographic Cosmological Models (In preparation)*. PhD thesis, The University of Edinburgh, Edinburgh.

[109] Sergio Caracciolo, Giuseppe Curci, Pietro Menotti, and Andrea Pelissetto. The Energy Momentum Tensor on the Lattice: The Scalar Case. *Nucl. Phys. B*, 309: 612–624, 1988. doi: 10.1016/0550-3213(88)90332-X.

[110] Sergio Caracciolo, Giuseppe Curci, Pietro Menotti, and Andrea Pelissetto. The Energy Momentum Tensor for Lattice Gauge Theories. *Annals Phys.*, 197:119, 1990. doi: 10.1016/0003-4916(90)90203-Z.

[111] John C Collins. Finite improvement renormalizes the energy-momentum tensor. *Physical Review Letters*, 36(26):1518, 1976.

[112] R. Narayanan and H. Neuberger. Infinite N phase transitions in continuum Wilson loop operators. *JHEP*, 03:064, 2006. doi: 10.1088/1126-6708/2006/03/064.

[113] Martin Lüscher. Trivializing maps, the wilson flow and the hmc algorithm. *Communications in Mathematical Physics*, 293(3):899–919, 2010.

[114] Martin Lüscher. Properties and uses of the wilson flow in lattice qcd. *Journal of High Energy Physics*, 2010(8):1–18, 2010.

[115] Martin Lüscher and Peter Weisz. Perturbative analysis of the gradient flow in non-abelian gauge theories. *Journal of High Energy Physics*, 2011(2):1–23, 2011.

[116] Hiroshi Suzuki. Energy–momentum tensor from the yang–mills gradient flow. *Progress of Theoretical and Experimental Physics*, 2013(8):083B03, 2013.

[117] Masayuki Asakawa, Tetsuo Hatsuda, Etsuko Itou, Masakiyo Kitazawa, Hiroshi Suzuki, FlowQCD Collaboration, et al. Thermodynamics of s u (3) gauge theory from gradient flow on the lattice. *Physical Review D*, 90(1):011501, 2014.

[118] Hiroki Makino and Hiroshi Suzuki. Lattice energy–momentum tensor from the yang–mills gradient flowinclusion of fermion fields. *Progress of Theoretical and Experimental Physics*, 2014(6), 2014.

[119] Hiroki Makino and Hiroshi Suzuki. Lattice energy-momentum tensor from the Yang-Mills gradient flow – a simpler prescription. 4 2014.

[120] Masakiyo Kitazawa, Masayuki Asakawa, Tetsuo Hatsuda, Takumi Iritani, Etsuko Itou, and Hiroshi Suzuki. Measurement of thermodynamics using gradient flow. *PoS*, LATTICE2014:022, 2014. doi: 10.22323/1.214.0022.

[121] Masakiyo Kitazawa, Takumi Iritani, Masayuki Asakawa, Tetsuo Hatsuda, and Hiroshi Suzuki. Equation of State for SU(3) Gauge Theory via the Energy-Momentum Tensor under Gradient Flow. *Phys. Rev. D*, 94(11):114512, 2016. doi: 10.1103/PhysRevD.94.114512.

[122] Masakiyo Kitazawa, Takumi Iritani, Masayuki Asakawa, and Tetsuo Hatsuda. Correlations of the energy-momentum tensor via gradient flow in SU(3) Yang-Mills theory at finite temperature. *Phys. Rev. D*, 96(11):111502, 2017. doi: 10.1103/PhysRevD.96.111502.

[123] Takehiro Hirakida, Etsuko Itou, and Hiroaki Kouno. Thermodynamics for SU(2) pure gauge theory using gradient flow. *PoS*, LATTICE2018:167, 2018. doi: 10.22323/1.334.0167.

[124] Martin Lüscher. Future applications of the Yang-Mills gradient flow in lattice QCD. *PoS*, LATTICE2013:016, 2014. doi: 10.22323/1.187.0016.

[125] Alberto Ramos. The Yang-Mills gradient flow and renormalization. *PoS*, LATTICE2014:017, 2015. doi: 10.22323/1.214.0017.

[126] Christopher Monahan and Kostas Orginos. Locally smeared operator product expansions in scalar field theory. *Phys. Rev. D*, 91(7):074513, 2015. doi: 10.1103/PhysRevD.91.074513.

[127] Kazuo Fujikawa. The gradient flow in $\lambda\phi^4$ theory. *JHEP*, 03:021, 2016. doi: 10.1007/JHEP03(2016)021.

[128] Andrea Carosso, Anna Hasenfratz, and Ethan T. Neil. Nonperturbative Renormalization of Operators in Near-Conformal Systems Using Gradient Flows. *Phys. Rev. Lett.*, 121(20):201601, 2018. doi: 10.1103/PhysRevLett.121.201601.

[129] Paul Bergold and Caroline Lasser. Fourier series windowed by a bump function, 2019. URL https://arxiv.org/abs/1901.04365.

[130] G. Peter Lepage. The Analysis of Algorithms for Lattice Field Theory. In *Theoretical Advanced Study Institute in Elementary Particle Physics*, 6 1989.

[131] G. Parisi, R. Petronzio, and F. Rapuano. A Measurement of the String Tension Near the Continuum Limit. *Phys. Lett. B*, 128:418–420, 1983. doi: 10.1016/0370-2693(83)90930-9.

[132] Martin Lüscher and Peter Weisz. Locality and exponential error reduction in numerical lattice gauge theory. *JHEP*, 09:010, 2001. doi: 10.1088/1126-6708/2001/09/010.

[133] Michele Della Morte and Leonardo Giusti. A novel approach for computing glueball masses and matrix elements in yang-mills theories on the lattice. *Journal of High Energy Physics*, 2011(5), May 2011. ISSN 1029-8479. doi: 10.1007/jhep05(2011)056. URL http://dx.doi.org/10.1007/JHEP05(2011)056.

[134] Marco Cè, Leonardo Giusti, and Stefan Schaefer. Local factorization of the fermion determinant in lattice qcd. *Phys. Rev. D*, 95:034503, Feb 2017. doi: 10.1103/PhysRevD.95.034503. URL https://link.aps.org/doi/10.1103/PhysRevD.95.034503.

[135] Karl Jansen, Eike Hermann Müller, and Robert Scheichl. Multilevel Monte Carlo for quantum mechanics on a lattice. *Phys. Rev. D*, 102:114512, 2020. doi: 10.1103/PhysRevD.102.114512.

[136] Mattia Dalla Brida, Leonardo Giusti, Tim Harris, and Michele Pepe. Multi-level Monte Carlo computation of the hadronic vacuum polarization contribution to $(g_\mu - 2)$. *Phys. Lett. B*, 816:136191, 2021. doi: 10.1016/j.physletb.2021.136191.

[137] Miguel Francisco García Vera. *Investigating the large N limit of SU(N) Yang-Mills gauge theories on the lattice*. PhD thesis, Humboldt U., Berlin, 2017.

[138] W Lenz. Beitrag zum Verständnis der magnetischen Erscheinungen in festen Körpern. *Z. Phys.*, 21:613–615, 1920. URL http://cds.cern.ch/record/460663.

[139] Lars Onsager. Crystal statistics. 1. A Two-dimensional model with an order disorder transition. *Phys. Rev.*, 65:117–149, 1944. doi: 10.1103/PhysRev.65.117.

[140] Ulli Wolff. Collective monte carlo updating for spin systems. *Phys. Rev. Lett.*, 62: 361–364, Jan 1989. doi: 10.1103/PhysRevLett.62.361. URL https://link.aps.org/doi/10.1103/PhysRevLett.62.361.