Selected Problems in Computational Gravity

by

Jonah M. Miller

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Doctor of Philosophy
in
Physics

Guelph, Ontario, Canada

©Jonah M. Miller, August, 2017
ABSTRACT

SELECTED PROBLEMS IN GRAVITY

Jonah M. Miller
University of Guelph, 2017

Advisors:
Erik Schnetter
and Eric Poisson

Einstein’s theory of general relativity comprises a system of ten highly nonlinear, coupled differential equations. This makes problems in general relativity notoriously difficult to solve via analytic methods. Therefore, numerical simulations and other computational techniques are often required to make progress. In this work, we discuss three such problems: transition amplitudes and the arrow of time in causal dynamical triangulations (CDT), discontinuous Galerkin finite element (DGFE) methods for relativistic astrophysics, and the stability of Kerr black holes with Proca hair.
To Alexandra Fresch,

with whom I’ve shared a grand adventure. You came so far with me in both space and time. I hope I’ve helped and supported you as much as you’ve helped and supported me.

You are the Christoffel connection to my metric.
Thanks to my thesis advisor, Erik Schnetter, and to my advisory committee members, Luis Lehner, Eric Poisson, and Harald Pfeiffer. I could not have come this far without you and all the other people who have mentored me over the years: Thomas Schibli and Jeanne Clelland, who helped me discover myself as a researcher; Steve Carlip, who trained me in CDT; Niayesh Afshordi and Achim Kempf, who trained me in astrophysics, quantum field theory, and cosmology; Christoph Junghans, Brandon Wiggins, Wes Even, Josh Dolence, Chris Fryer, Ben Bergen, Robert Bird, Oleg Korobkin, Allen McPherson, Robert Pavel, and Geoff Womeldorff, who trained me in SPH; Will East, who trained me in Proca fields; and Saul Teukolsky, Larry Kidder, Mark Scheel, Christian Ott, Gab Allen, Ed Siedel, Dan Katz, Matthew Turk, Eliu Huerta, Ian Hinder, Barry Wardell, Scott Field, Philipp Moesta, Eloisa Bentivegna, Bruno Giacomazzo, David Radice, Roland Haas, and Nathan Goldbaum, who provided invaluable advice.

I would also like to thank all my friends and colleagues, who have shared their knowledge with me and provided intellectual stimulation and emotional support. These include but are not limited to: Josh Cooperman, Kyle Lee, Rajesh Kommu, Chien-Chung Lee, Farbod Kamiab, Yasaman Yazdi, Brian Yee, Sarah Kaiser, Tyler Jackson, Scott VanBommel, Stephen Green, Nestor Ortiz, Jorge Alejandro Lopez, John Ryan Westernacher-Schneider, Pablo Bosch, Phil Landry, Huan Yang, Raissa Mendez, Peter Zimmerman, Heather Fong, Trever Vincent, Adam Lewis, Alejandro Cardenas Avendano, Federico Guercilena, Nicolas de Brye, Daniel George, Glenn Hordemann, Hyun Lim, Julien Loiseau, and Jonathan Sharmin.

Finally, I would like to thank my family. I never call or write, so you do it for me.
# Table of Contents

List of Tables viii

List of Figures ix

1 Introduction 1
   1.1 The Paths to Computational Gravity ........................................ 2
   1.2 A Brief Summary of Topics ...................................................... 4

I Causal Dynamical Triangulations and the Arrow of Time 6

2 A First Look at Transition Amplitudes in (2+1)-Dimensional Causal Dynamical Triangulations 9
   2.1 Context ................................................................. 9
   2.2 Abstract ........................................................................ 10
   2.3 Introduction ............................................................... 11
   2.4 Causal dynamical triangulations with fixed boundaries .................. 16
   2.5 Transition amplitudes ...................................................... 32
   2.6 Conclusion ................................................................. 53
   Appendices ........................................................................ 58
   2.A On the consistency of the action $S_{R[T_c]}$ under composition ........ 58
   2.B On the algorithm for inserting spatial 2-sphere boundaries ............. 59
   2.C A derivation of the function $A_2^{S_L}(\tau)$ .................................. 62

3 A Second Look at Transition Amplitudes in (2+1)-Dimensional Causal Dynamical Triangulations 72
   3.1 Context ................................................................. 72
   3.2 Abstract ........................................................................ 73
   3.3 Euclidean from Lorentzian ...................................................... 73
   3.4 Causal dynamical triangulations ............................................... 77
## II Relativistic Astrophysics

### 4 An Operator-Based Discontinuous Galerkin Method Compatible with the BSSN Formulation of the Einstein Equations

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Context</td>
<td>132</td>
</tr>
<tr>
<td>4.2</td>
<td>Abstract</td>
<td>133</td>
</tr>
<tr>
<td>4.3</td>
<td>Introduction</td>
<td>133</td>
</tr>
<tr>
<td>4.4</td>
<td>Methods</td>
<td>140</td>
</tr>
<tr>
<td>4.5</td>
<td>Asymptotic Properties</td>
<td>178</td>
</tr>
<tr>
<td>4.6</td>
<td>Numerical Tests</td>
<td>189</td>
</tr>
<tr>
<td>4.7</td>
<td>Concluding Remarks</td>
<td>210</td>
</tr>
<tr>
<td>4.8</td>
<td>Acknowledgments</td>
<td>211</td>
</tr>
<tr>
<td>4.A</td>
<td>Legendre Polynomials, Collocation Points, and Gauss Lobatto Quadrature</td>
<td>213</td>
</tr>
<tr>
<td>4.B</td>
<td>Proof That $\xi = 1$ Satisfies Summation By Parts</td>
<td>218</td>
</tr>
<tr>
<td>4.C</td>
<td>Stability for the Wave Equation</td>
<td>223</td>
</tr>
<tr>
<td>4.D</td>
<td>Convergence for the Wave Equation</td>
<td>225</td>
</tr>
<tr>
<td>4.E</td>
<td>Making Contact with Standard Discontinuous Galerkin Methods</td>
<td>230</td>
</tr>
<tr>
<td>4.F</td>
<td>Calculating Self-Convergence</td>
<td>234</td>
</tr>
</tbody>
</table>

### 5 Nonlinear Stability of Kerr Black Holes with Proca Hair

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Context</td>
<td>236</td>
</tr>
<tr>
<td>5.2</td>
<td>Abstract</td>
<td>237</td>
</tr>
<tr>
<td>5.3</td>
<td>Introduction</td>
<td>237</td>
</tr>
<tr>
<td>5.4</td>
<td>HRR’s Construction of Kerr Black Holes with Proca Hair</td>
<td>240</td>
</tr>
<tr>
<td>5.5</td>
<td>Transformation to Horizon Penetrating Coordinates</td>
<td>248</td>
</tr>
<tr>
<td>5.6</td>
<td>Validation of our Coordinate Transformation</td>
<td>272</td>
</tr>
<tr>
<td>5.7</td>
<td>Next Steps</td>
<td>275</td>
</tr>
<tr>
<td>5.8</td>
<td>Acknowledgements</td>
<td>276</td>
</tr>
<tr>
<td>Appendix</td>
<td></td>
<td>278</td>
</tr>
<tr>
<td>5.A</td>
<td>On the Regularity of $\Lambda$</td>
<td>278</td>
</tr>
<tr>
<td>5.B</td>
<td>On the Regularity of $\mathfrak{P}$ and $\Xi$</td>
<td>279</td>
</tr>
</tbody>
</table>
5.C The Asymptotic Limit of SKSL Coordinates ............................. 281
5.D On the Regularity of Σ ...................................................... 283
5.E On the Regularity of Δ ...................................................... 283
5.F Calculating the Field Tensor .............................................. 285
5.G The Asymptotic Limit of CKSL Coordinates ......................... 285

6 Computational Gravity ......................................................... 287
   6.1 Gauge Freedom .......................................................... 288
   6.2 Discretuum and Continuum .......................................... 290
   6.3 Conclusion .............................................................. 293

References ............................................................... 294
### List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>$\chi^2$ Values as we vary the number of time slices</td>
<td>43</td>
</tr>
<tr>
<td>4.1</td>
<td>Notation for discontinuous Galerkin methods</td>
<td>141</td>
</tr>
<tr>
<td>4.2</td>
<td>Gauss-Lobatto quadrature</td>
<td>214</td>
</tr>
</tbody>
</table>
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The three types of 3-simplices employed in ((2 + 1))-dimensional CDT</td>
<td>18</td>
</tr>
<tr>
<td>2.2</td>
<td>An illustration of extrinsic curvature</td>
<td>23</td>
</tr>
<tr>
<td>2.3</td>
<td>Extrema of the Euclidean minisuperspace action</td>
<td>37</td>
</tr>
<tr>
<td>2.4</td>
<td>Ensemble averages vs. individual triangulations</td>
<td>40</td>
</tr>
<tr>
<td>2.5</td>
<td>Fixed boundaries remove the stalk</td>
<td>41</td>
</tr>
<tr>
<td>2.6</td>
<td>Fixed boundaries “squeeze” the spacetime</td>
<td>42</td>
</tr>
<tr>
<td>2.7</td>
<td>Ensembles that transition from minimal initial to nonminimal final boundaries</td>
<td>45</td>
</tr>
<tr>
<td>2.8</td>
<td>Ensembles with (N_{SL}^2(S_i^2) = 4) and (N_{SL}^2(S_f^2) \in {100, 300, 500, 700})</td>
<td>46</td>
</tr>
<tr>
<td>2.9</td>
<td>Varying the number of time slices for minimal initial and nonminimal final boundaries</td>
<td>47</td>
</tr>
<tr>
<td>2.10</td>
<td>Ensembles with (N_{SL}^2(S_i^2) = N_{SL}^2(S_f^2))</td>
<td>50</td>
</tr>
<tr>
<td>2.11</td>
<td>Ensembles with (N_{SL}^2(S_i^2) \neq N_{SL}^2(S_f^2), N_{SL}^2(S_i^2), N_{SL}^2(S_f^2) \neq 0)</td>
<td>51</td>
</tr>
<tr>
<td>2.12</td>
<td>Ensembles near the edge of validity for the Euclidean minisuperspace action</td>
<td>52</td>
</tr>
<tr>
<td>2.13</td>
<td>An example decomposition of a triangulated spatial (2)-sphere (not completely depicted) into four pseudofaces—labelled 1, 2, 3, and 4—and the six corresponding pseudoedges.</td>
<td>61</td>
</tr>
<tr>
<td>2.14</td>
<td>The psuedo-3-simplices employed when imposing boundary data</td>
<td>62</td>
</tr>
<tr>
<td>3.1</td>
<td>Causal 3-simplices employed in ((2+1))-dimensional CDT</td>
<td>79</td>
</tr>
<tr>
<td>3.2</td>
<td>(\langle N_{SL}^2(S_i^2) \rangle) for (N_{SL}^2(S_i^2) = 4), (N_{SL}^2(S_f^2) \in {4, 100})</td>
<td>87</td>
</tr>
<tr>
<td>3.3</td>
<td>(\langle N_{SL}^2(S_i^2) \rangle) for (N_{SL}^2(S_i^2) = 500), (N_{SL}^2(S_f^2) \in {500, 700, 900})</td>
<td>88</td>
</tr>
<tr>
<td>3.4</td>
<td>(\langle N_{SL}^2(S_i^2) \rangle) for (N_{SL}^2(S_i^2) = 600), (N_{SL}^2(S_f^2) \in {600, 800})</td>
<td>89</td>
</tr>
<tr>
<td>3.5</td>
<td>Eigenvectors of (\langle n_{SL}^2(\tau) n_{SL}(\tau') \rangle)</td>
<td>90</td>
</tr>
<tr>
<td>3.6</td>
<td>Eigenvalues of (\langle n_{SL}^2(\tau) n_{SL}(\tau') \rangle)</td>
<td>91</td>
</tr>
<tr>
<td>3.7</td>
<td>Best fit of the scale factor for Euclidean de Sitter space to (\langle N_{SL}^2 \rangle) for (N_{SL}^2(S_i^2) = N_{SL}^2(S_f^2) = 4)</td>
<td>96</td>
</tr>
<tr>
<td>3.8</td>
<td>Best, fit of the spacelike 2-volume for Lorentzian de Sitter space to (\langle N_{SL}^2 \rangle) for (N_{SL}^2(S_i^2) = 500), (N_{SL}^2(S_f^2) \in {500, 700, 900})</td>
<td>97</td>
</tr>
<tr>
<td>3.9</td>
<td>Best fit of the spacelike 2-volume for Lorentzian de Sitter space to (\langle N_{SL}^2 \rangle) for (N_{SL}^2(S_i^2) = 500), (N_{SL}^2(S_f^2) \in {500, 700, 900})</td>
<td>98</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.10</td>
<td>Eigenvectors of the two-point function for spacelike 2-volume for Euclidean de Sitter space: discrete vs. continuum</td>
<td>103</td>
</tr>
<tr>
<td>3.11</td>
<td>Eigenvalues of the two-point function for the spacelike 2-volume for Euclidean de Sitter space: discrete vs. continuum</td>
<td>104</td>
</tr>
<tr>
<td>3.12</td>
<td>Eigenvectors of the two-point function for spacelike 2-volume for Lorentzian de Sitter space: discrete vs. continuum</td>
<td>105</td>
</tr>
<tr>
<td>3.13</td>
<td>Eigenvalues of the two-point function for the spacelike 2-volume for Lorentzian de Sitter space: discrete vs. continuum</td>
<td>106</td>
</tr>
<tr>
<td>3.14</td>
<td>Spatial 2-volume as a function of time in our Lagrange-multiplier based model</td>
<td>111</td>
</tr>
<tr>
<td>3.15</td>
<td>Spectral dimension for $N_{SL}^2(S^2_i) = N_{SL}^2(S^2_f) \in {4, 600}$</td>
<td>114</td>
</tr>
<tr>
<td>4.1</td>
<td>A broken domain with $K = 3$</td>
<td>146</td>
</tr>
<tr>
<td>4.2</td>
<td>An example test function</td>
<td>149</td>
</tr>
<tr>
<td>4.3</td>
<td>Wide derivative operator for $P^k = 2$</td>
<td>165</td>
</tr>
<tr>
<td>4.4</td>
<td>Wide derivative operator for $P^k = 3$</td>
<td>166</td>
</tr>
<tr>
<td>4.5</td>
<td>Wide derivative operator for $P^k = 4$</td>
<td>166</td>
</tr>
<tr>
<td>4.6</td>
<td>Computational cost for DGFE vs FD</td>
<td>182</td>
</tr>
<tr>
<td>4.7</td>
<td>A two-dimensional slice of the collocation cells on a single CPU</td>
<td>183</td>
</tr>
<tr>
<td>4.8</td>
<td>Theoretical communication overhead under strong scaling for DGFE vs FD</td>
<td>185</td>
</tr>
<tr>
<td>4.9</td>
<td>Theoretical memory access overhead for DGFE vs FD</td>
<td>189</td>
</tr>
<tr>
<td>4.10</td>
<td>Second-order wave equation under $h$-refinement</td>
<td>192</td>
</tr>
<tr>
<td>4.11</td>
<td>Second-order wave equation under $p$-refinement</td>
<td>193</td>
</tr>
<tr>
<td>4.12</td>
<td>Pointwise error for the wave equation “apples to apples” DGFE vs FD</td>
<td>194</td>
</tr>
<tr>
<td>4.13</td>
<td>Pointwise error for the wave equation “realistic” DGFE vs FD</td>
<td>195</td>
</tr>
<tr>
<td>4.14</td>
<td>Pointwise error for the wave equation $4^{th}$- vs. $5^{th}$-order DGFE</td>
<td>197</td>
</tr>
<tr>
<td>4.15</td>
<td>Interpolation error with and without averaging</td>
<td>200</td>
</tr>
<tr>
<td>4.16</td>
<td>The robust stability test under $h$-refinement</td>
<td>201</td>
</tr>
<tr>
<td>4.17</td>
<td>The robust stability test under $p$-refinement</td>
<td>202</td>
</tr>
<tr>
<td>4.18</td>
<td>A two-dimensional slice of the three-dimensional gauge wave</td>
<td>203</td>
</tr>
<tr>
<td>4.19</td>
<td>“Pointwise“ convergence for the one-dimensional gauge wave</td>
<td>205</td>
</tr>
<tr>
<td>4.20</td>
<td>$L_2$ norm errors for the one-dimensional gauge wave</td>
<td>206</td>
</tr>
<tr>
<td>4.21</td>
<td>$L_2$-norm errors for the three-dimensional gauge wave</td>
<td>207</td>
</tr>
<tr>
<td>4.22</td>
<td>$L_2$-norm differences for the shifted gauge wave</td>
<td>209</td>
</tr>
<tr>
<td>4.23</td>
<td>Self-convergence for the shifted gauge wave</td>
<td>210</td>
</tr>
<tr>
<td>5.1</td>
<td>Components of the Proca one-form $A_\mu$ in HRR’s coordinates</td>
<td>248</td>
</tr>
<tr>
<td>5.2</td>
<td>Lapse $\alpha$ and volume form $\sqrt{\gamma}$ for HRR’s coordinate system</td>
<td>249</td>
</tr>
<tr>
<td>5.3</td>
<td>Example values of $\Lambda$ and $\Xi$</td>
<td>253</td>
</tr>
<tr>
<td>5.4</td>
<td>Example values of $\Sigma$ and $-i \ln \Psi$</td>
<td>258</td>
</tr>
<tr>
<td>5.5</td>
<td>Example values of $\xi$ and $\Delta$</td>
<td>259</td>
</tr>
<tr>
<td>5.6</td>
<td>Constant time slices in HRR and SKSL coordinates</td>
<td>271</td>
</tr>
</tbody>
</table>
5.7 Lapse $\alpha$ and volume form $\sqrt{\gamma}$ for CKSL coordinates \hfill 274
5.8 Shift in CKSL coordinates \hfill 275
5.9 Components of the Proca field in CKSL coordinates after the 3+1 split \hfill 277
Chapter 1

Introduction

Using a term like nonlinear science is like referring to the bulk of zoology as the study of non-elephant animals.

Stanisław Ulam

This work discusses three topics united by a need to solve problems in gravity and general relativity using computational techniques. Discrete quantum gravity and numerical relativistic astrophysics share many surprising connections. Although I do not explore these connections in detail when I attack these problems, these connections will appear if one steps back from the trees and observes the forest. I therefore propose a new name for the study of topics at the intersection of numerical relativity and quantum gravity: computational gravity.
1.1 The Paths to Computational Gravity

Einstein’s theory of general relativity (GR) [88, 87] elegantly combines space, time, and geometry to describe gravitational systems. GR has been wildly successful. It has passed an array of weak-field tests [235], its regime of validity ranges from compact object astrophysics [198] to cosmology [224], and the recent direct detection of gravitational waves predicted by GR [1, 70] have opened the door to gravitational wave astronomy.

When unpacked into the language of partial differential equations, the conceptually clean language of the geometry of curved manifolds becomes complex and non-linear. Since Einstein first presented his theory 102 years ago, exact [173, 206] and perturbative [174] solutions to this complex nonlinear system have proliferated and GR has thrived, both as a tool for astrophysics and cosmology and as an independent field. The field has now matured to the point where its practitioners routinely tackle complex, nonlinear problems both analytically and numerically.

Numerical relativity is the field of study where one solves problems in classical general relativity using numerical techniques. In numerical relativity (and most numerical approaches to PDE theory), the Einstein equations and their solutions are approximated by a discrete, finite number of degrees of freedom. Numerical simulations were critical in the recent direct detection and analysis of gravitational waves [1, 70] and they will likely play an important role in relativistic astrophysics for the foreseeable future.

\footnote{Although GR has been very successful in a cosmological context, cosmology is arguably the field where problems with GR are readily visible in the presence of the unknown dark sector.}
Research in quantum gravity arrived at numerical calculations from a rather different perspective. Naive quantizations of the Einstein field equations suffer from an ultraviolet divergence that does not appear amenable to power-counting renormalization techniques\textsuperscript{2} and the correct measure in path integral formulations is not known. One can cure both of these ills by introducing an ultraviolet cutoff and \textit{discretizing} the geometry.

Depending on the approach, the discrete theory may be physical or it may simply be an approximation technique. In the case of the latter, the connection to numerical relativity is clear: both approaches use a system with a finite number of degrees of freedom to approximate one with an infinite number. In either case, the combinatorics of this discrete geometry usually necessitate numerical simulations. In this work, I will discuss only one type of discrete quantum gravity—causal dynamical triangulations—descended from a particular lineage of approaches. However, I urge the reader to investigate other promising approaches such as spin foams and causal sets. See [171] and [204] for respective reviews.

As we shall see, there are a surprising number of parallels between numerical approaches to classical GR and to discrete quantum gravity. For example, both numerical approximations of GR and quantum theories of gravity must converge to the classical theory in the appropriate limit. The former must do so as a necessary (but not sufficient) condition for the validity of the approximation, while the latter must match observations at low-energies. I urge the reader to keep these connections in mind as they peruse the selected topics below.

\textsuperscript{2}There are promising approaches to resolving this divergence by searching for fixed-points in the renormalization group flow. For a review, see [163].
1.2 A Brief Summary of Topics

In this work, I present two related studies of lattice quantum gravity, one novel numerical method motivated by the needs of the relativistic astrophysics community, and a study of dynamics in an astrophysical context. The first two problems are presented as three peer-review papers collected here. The last is a work in progress and is not yet peer-reviewed.

Causal dynamical triangulations (CDT) is a sum-over-histories approach to quantum gravity motivated by the success of lattice field theory [16]. In CDT, spacetime itself becomes the lattice. Spacetimes are approximated, and thus regularized, as piecewise flat manifolds constructed out of pieces of Minkowski space. In [78] (chapter 2), my collaborator Josh Cooperman and I presented the first-ever studies of transition amplitudes in CDT. We also uncovered a tantalizing connection to the arrow of time, and evidence that the signature of the metric in CDT may not be fixed. In [77] (chapter 3), along with Kyle Lee, we re-examined this evidence and presented a simpler explanation. We believe that our case study in effective signature change in CDT can shine light on other speculative discussions of this topic such as those presented in [79] and [8].

One difficulty in utilizing simulations to make predictions is that, for fully (3+1)-dimensional general relativity, these simulations require a very long runtime. Discontinuous Galerkin finite element (DGFE) methods offer a faster alternative to the traditional algorithms [119]. In collaboration with Erik Schnetter, I constructed a generalization of DGFE methods compatible with a broad class of systems including the BSSN formulation of the Einstein equations. This work is presented in [160] and in chapter 4.
A minimal coupling of massive vector fields—known as Proca fields [182]—to gravity not only provides a toy model with which to explore dynamics of modified [131, 122] and unmodified gravity [166], but may be astrophysically relevant in the context of stringy models of dark matter [37]. Recently Herdeiro, Radu, and Runarsson discovered a family of stationary black holes with Proca hair, evading the no hair theorems [115]. I am numerically studying the nonlinear stability of these solutions to perturbation. In chapter 5, I present the first steps required for such a study. This study is a collaboration with William East.
Part I

Causal Dynamical Triangulations and the Arrow of Time
Chapters 2 and 3 contain [78] and [77] respectively, which are parts one and two of a two-part series. In this series, my collaborators and I studied the thick-sandwich problem in causal dynamical triangulations (CDT): Given the geometry of a spacelike hypersurface at some initial time and the geometry of one at some final time, what is the probability of the former transitioning into the latter? We provide some preliminary answers for CDT. In the process, we developed a formalism and the appropriate numerical tools to delve further into these questions.

Taking the continuum limit of a lattice field theory is the converse of the discretization procedure one undertakes when numerically solving a differential equation. In the latter case, one takes the continuum equation and finds a discrete system that well-approximates it. In CDT, conversely, one instead takes a discrete model of quantum general relativity, performs a numerical experiment, and then asks what continuum theory matches the experiment.\(^3\) For sufficiently long length and time scales (or sufficiently many discrete building blocks), the theory should match classical GR—one of the early triumphs of CDT was the demonstration that de Sitter spacetime dominates the vacuum path integral (with positive cosmological constant). However, between the classical limit and the ultraviolet cutoff, an effective field theory should emerge which may or may not be GR. Much of the research in CDT is aimed at understanding what this effective theory is, and progress has been made in this direction. For example, there is some evidence that the low-energy limit of CDT is Hořava-Lifshitz gravity [12, 11, 31].

\(^3\)The ultraviolet cutoff in CDT is not considered physical, but an artifact of the regularization procedure. Nevertheless, the discrete theory, not the continuum one, is the first-class citizen.
Effective field theories are not necessarily unique, however, and it is often difficult to extract physical meaning from them. [78] and [77] show one such example. In [78] (chapter 2), we performed numerical experiments that are well-described by an effective theory with a signature change in it—the physical spacetime transitions from Lorentzian signature to Euclidean. In [78], we attributed physical meaning to this effective action. However, in [77] (chapter 3), we show that a simpler effective theory without signature change works just as well. We believe the simpler theory is the one with physical meaning.
Chapter 2

A First Look at Transition Amplitudes in (2+1)-Dimensional Causal Dynamical Triangulations

2.1 Context

This work (first presented in [78]) was a collaboration between myself and Joshua Cooperman. The problem we solved was posed to us by Steve Carlip.

The bulk of the work was done between June 2012 and May 2013, before I received my bachelor’s degree in May 2013. While I was a Ph.D. student, the resulting paper underwent peer review from May to December 2013 and was published in January 2014. I was the corresponding author on this work. Based on the unpublished work of David Kamensky,
I developed the mathematical machinery and the novel algorithm presented here. I also wrote the code that implements it and ran all simulations. J. Cooperman performed much of the data analysis. Our contributions to the prose in the paper are about equal. As of this writing, the paper has received eight citations in the approximately three years since its publication.

2.2 Abstract

We study a lattice regularization of the gravitational path integral—causal dynamical triangulations—for $(2+1)$-dimensional Einstein gravity with positive cosmological constant in the presence of past and future spacelike boundaries of fixed intrinsic geometries. For spatial topology of a 2-sphere, we determine the form of the Einstein-Hilbert action supplemented by the Gibbons-Hawking-York boundary terms within the Regge calculus of causal triangulations. Employing this action we numerically simulate a variety of transition amplitudes from the past boundary to the future boundary. To the extent that we have so far investigated them, these transition amplitudes appear consistent with the gravitational effective action previously found to characterize the ground state of quantum spacetime geometry within the Euclidean de Sitter-like phase. Certain of these transition amplitudes convincingly demonstrate that the so-called stalks present in this phase are numerical artifacts of the lattice regularization, seemingly indicate that the quantization technique of causal dynamical triangulations differs in detail from that of the no-boundary proposal of Hartle and Hawking, and possibly represent the first numerical simulations of portions of temporally unbounded quantum spacetime geometry within the causal dynamical triangulations approach. We also uncover tantalizing evidence suggesting that Lorentzian not Euclidean de Sitter spacetime dominates the ground state on sufficiently large scales.
2.3 Introduction

A quantum theory’s ground state dictates much but not all of that theory’s structure: transition amplitudes above the ground state encode a rich structure themselves. Indeed, the information contained in transition amplitudes is most relevant for experimental tests of the quantum theory: for an experiment to have empirical import, it must strive to observe at least a modicum of change. As in all other quantum theories, one would like to compute transition amplitudes in a quantum theory of gravity. Such amplitudes are often computed in various quantum theories of gravity. In the case of \((2+1)\)-dimensional causal dynamical triangulations, Ambjørn et al considered transition amplitudes within a spherical symmetry truncation of one variant and Benedetti et al considered transition amplitudes in a more ordered variant [47]. However, we know of no calculations for the full theory in \(2+1\) spacetime dimensions or higher. We remedy this situation by presenting here the first numerical simulations of transition amplitudes in causal dynamical triangulations.

Causal dynamical triangulations is a path integral based approach to the nonperturbative quantization of classical metric theories of gravity. (See [17] for a review.) In rigorously defining this quantization, causal dynamical triangulations emulates the well established techniques of lattice quantum field theory, specifically, the introduction of a lattice regularization followed by application of finite size scaling and renormalization. The approach distinguishes itself in two related respects: its restriction on the spacetime geometries allowed to contribute to the gravitational path integral and its associated regularization of this path integral. One restricts to causal spacetime geometries, namely those admitting a global
foliation by spacelike hypersurfaces all of the same topology. One regulates causal spacetime geometries by causal triangulations, namely Lorentzian simplicial manifolds possessing such a foliated structure. When computing the gravitational path integral according to this prescription, one thus additionally specifies a topology for the leaves of the foliation.

The restriction to causal spacetime geometries—the approach’s key new feature—finds its primary motivation in the failures of previous (Euclidean) lattice quantum gravity programs to define physically sound quantum theories of gravity. (See, for instance, the review [153].) With models in $1 + 1$ dimensions implicating the absence of any Lorentzian causal structure in these failures, Ambjørn, Loll, and Jurkiewicz decided to formulate a regularization of the gravitational path integral directly in Lorentzian signature [21, 30]. Of course, a generic Lorentzian spacetime geometry does not admit a global foliation by spacelike hypersurfaces all of the same topology, a fact that significantly complicates its regularization by a Lorentzian simplicial manifold. Consequently, these authors elected to impose the restriction to causal spacetime geometries. Besides allowing for the implementation of the lattice regularization in Lorentzian signature, this causality condition provides for a well defined Wick rotation to Euclidean signature required for Monte Carlo simulations, the primary means of investigation of causal dynamical triangulations. One may view the causality condition as a compromise between the necessities of introducing causal structure and of conducting numerical studies.

Causal dynamical triangulations has thus far produced several promising results. Most studies in $2 + 1$ and $3 + 1$ dimensions have focused on the ground state emerging from
the quantization of Einstein gravity with positive cosmological constant for topologically spherical leaves of the foliation.\footnote{There are two notable exceptions for $2+1$ dimensions: Budd considered topologically toric leaves \cite{56}, and Anderson \textit{et al} considered projectable Hořava-Lifshitz gravity \cite{31}.} Within the so-called phase C of quantum spacetime geometry, which is present for both $2+1$ and $3+1$ dimensions, this ground state exhibits the correct semiclassical behavior on sufficiently large scales and novel quantum mechanical behavior on sufficiently small scales. In particular, on large scales the gravitational effective action of a simple minisuperspace model describes exceedingly well the quantum spacetime geometry \cite{9, 15, 14, 18, 23, 24, 26, 25}. The ensemble average spacetime geometry fits that of Euclidean de Sitter spacetime,\footnote{The spacetime is Euclidean, rather than Lorentizan, after Wick rotation.} the maximally symmetric extremum of this gravitational effective action \cite{9, 15, 14, 18, 23, 24, 26, 25, 31, 45, 137}. Deviations from the ensemble average spacetime geometry fit a straightforward quantization of this gravitational effective action with Euclidean de Sitter spacetime as the ground state \cite{15, 14}. Of course, these results accord with the semiclassical expectation that the ground state is the most symmetric configuration. On small scales the ensemble average spacetime geometry exhibits a dynamical dimensional reduction from the topological dimension of $2+1$ or $3+1$ to a dimension of approximately $2$ \cite{25, 27, 45, 137}. In $3+1$ dimensions there is also firm evidence for the presence of a second order phase transition at the B-C phase boundary, potentially pointing to a well defined continuum limit of causal dynamical triangulations \cite{19, 20}.

All of these results concern the ground state of quantum spacetime geometry. We wish to learn more about the nature of the causal dynamical triangulations quantization scheme...
by studying transition amplitudes. Given the foliation of each causal triangulation into spacelike hypersurfaces of a fixed topology, one most readily considers transition amplitudes between two leaves of this foliation. Such amplitudes constitute the analogue in the quantum theory of the so-called thick sandwich problem: specify the intrinsic geometries of initial and final spacelike hypersurfaces, allowing their extrinsic geometries to vary, and then compute the evolution from initial to final spacelike hypersurface. In principle, one could also consider other types of transition amplitudes specified by other types of boundary conditions. For instance, Warner, Catterall, and Renken explored Euclidean dynamical triangulations with a single boundary, introducing a generalized boundary action with the hope of discovering an enriched phase structure [231, 232]. We have not pursued these other possibilities for two primary reasons. First, their implementation in causal dynamical triangulations is rather more difficult than that of the thick sandwich problem. Second, the thick sandwich problem already affords the investigation of some interesting physical questions. Clearly, computing these transition amplitudes amounts to studying causal dynamical triangulations with fixed boundaries, a situation not previously explored for spacetime dimension of $2 + 1$ or higher.$^3$

We specifically consider the causal dynamical triangulations of $(2+1)$-dimensional Einstein gravity with positive cosmological constant for spacetime topology of the direct product of a 2-sphere $S^2$ and a line interval $I$. In section 2.4 we first introduce the formalism of causal dynamical triangulations, and we then derive the form of the action appropriate to

---

$^3$Technically, we do not compute any transition amplitudes but only explore properties of representative causal triangulations contributing to certain transition amplitudes. The nature of Markov chain Monte Carlo simulations forces this approach on us.
the boundary conditions under consideration. We also discuss the modified numerical implementation that we have devised to run simulations of the path integral with these boundary conditions. In section 2.5 we report the results of our numerical simulations of three classes of transition amplitudes: from minimal initial boundary to minimal final boundary, from minimal initial boundary to nonminimal final boundary, and from nonminimal initial boundary to nonminimal final boundary. The second of these classes captures within causal dynamical triangulations the setting of the Hartle-Hawking no-boundary wavefunction [113]. We demonstrate that the transition amplitudes of the first two classes and certain transition amplitudes of the third class are quantitatively consistent with the gravitational effective action previously found to characterize the ground state of quantum spacetime geometry within phase C. We propose that this may also be the case for all of the transition amplitudes of the third class. The transition amplitudes of the first class substantiate the hypothesis that the so-called stalks present in this phase are numerical artifacts of the lattice regularization. The transition amplitudes of the second class imply that the causal dynamical triangulations quantization scheme differs from that of the no-boundary proposal of Hartle and Hawking. Some of the transition amplitudes of the third class appear to constitute the first numerical simulations of portions of temporally unbounded quantum spacetime geometry within the causal dynamical triangulations approach. These last transition amplitudes also suggest that Lorentzian not Euclidean de Sitter spacetime may dominate the ground state of quantum spacetime geometry on sufficiently large scales. In section 2.6 we conclude by considering several extensions of our work and several questions that such
extensions could address.

2.4 Causal dynamical triangulations with fixed boundaries

2.4.1 Formalism

The path integral offers a powerful nonperturbative procedure for quantizing classical theories. To compute a transition amplitude $A[\xi]$ in the quantum theory, one evaluates an expression of the form

$$A[\xi] = \int_{q|_B = \xi} \mathcal{D}q e^{iS[q]}.$$  \hspace{1cm} (2.1)

Here, $q$ denotes the set of dynamical variables of the classical theory characterized by the action $S[q]$. One performs the path integration over all physically distinct configurations of the dynamical variables $q$ consistent with the boundary conditions $q|_B = \xi$ defining the transition amplitude $A[\xi]$. Of course, except in certain sufficiently simple circumstances, such path integrals are technically difficult not only to compute, but even to define.\(^4\)

The causal dynamical triangulations approach nevertheless provides a prescription for overcoming these technical difficulties in the case of classical metric theories of gravity. A transition amplitude $A[\gamma]$ for such a theory takes the form

$$A[\gamma] = \int_{g|_{\partial M} = \gamma} \mathcal{D}g e^{iS[g]},$$  \hspace{1cm} (2.2)

\(^4\)One major difficulty in defining these integrals is choosing the appropriate measure. For example, many metrics represent the same physical spacetime. So a physically correct measure on the space of metrics must take this multiplicity into account. The same difficulty appears in path integral formulations of any gauge theory such as quantum chromodynamics.
where $g$ is the metric characterizing a spacetime $\mathcal{M}$ with boundary $\partial \mathcal{M}$. To define rigorously the transition amplitude (2.2), causal dynamical triangulations invokes a restriction on and a regularization of the spacetime geometries contributing to the path integral. Only spacetime geometries possessing the causal structure of a global foliation by spacelike hypersurfaces all of the same topology are permitted. One regulates these spacetime geometries by a specific class of simplicial manifolds, causal triangulations. A causal triangulation is a Lorentzian simplicial manifold constructed by gluing together Minkowskian simplices in a manner consistent with the presence of this foliation. We often refer to the foliation’s leaves as time slices, and we enumerate them with a discrete Lorentzian time coordinate $t \in \{1, \ldots, T\}$. In $2+1$ dimensions, the case that we consider below, every causal triangulation is constructed from the set of Minkowskian 3-simplices depicted in figure 2.1. Each leaf of the foliation is triangulated by regular spacelike 2-simplices—equilateral triangles—having squared edge length $a^2$. The fixed spacelike edge length $a$ serves as the lattice spacing. Adjacent leaves are then connected by timelike edges having squared length $-\alpha a^2$ for $\alpha > \frac{1}{2}$ such that only the three types of Minkowskian 3-simplices are formed. Starting from a given causal triangulation $\mathcal{T}_c$ with a fixed number $T$ of time slices, one may obtain any other causal triangulation $\mathcal{T}'_c$ also with a fixed number $T$ of time slices by applying a finite sequence of the dynamical Pachner moves adapted to causal triangulations [21, 22].

The prescription of causal dynamical triangulations transforms the path integral (2.2)
Figure 2.1: The three types of 3-simplices employed in $(2+1)$-dimensional causal dynamical triangulations: (a) $(3, 1)$ 3-simplex, (b) $(2, 2)$ 3-simplex, (c) $(1, 3)$ 3-simplex. The first number in the ordered pair indicates the number of vertices on the initial time slice ($t = 0$), and the second number in the ordered pair indicates the number of vertices on the final time slice ($t = 1$). We denote by $N_3^{(3,1)}$ the number of $(3, 1)$ 3-simplices, by $N_3^{(2,2)}$ the number of $(2, 2)$ 3-simplices, and by $N_3^{(1,3)}$ the number of $(1, 3)$ 3-simplices in a causal triangulation.

into the path sum

$$\mathcal{A}[\Gamma] = \sum_{\mathcal{T}_c} \mu(\mathcal{T}_c)e^{iS_R[\mathcal{T}_c]},$$

(2.3)

where $\mu(\mathcal{T}_c)$ is the measure, and $S_R[\mathcal{T}_c]$ is the action expressed in the Regge calculus of causal triangulations. One takes the measure equal to $\frac{1}{\mathcal{C}(\mathcal{T}_c)}$, the inverse of the order of the automorphism group of the causal triangulation $\mathcal{T}_c$. This factor accounts for the discrete remnants of the group of diffeomorphisms acting on the causal triangulation $\mathcal{T}_c$. The action $S_R[\mathcal{T}_c]$ depends on the classical theory of gravity that one wishes to quantize as well as the transition amplitude that one wishes to study, the latter dependence coming from the specification of boundary conditions by appropriate boundary terms in the action.
2.4.2 Action

We are interested in the causal dynamical triangulations of $(2 + 1)$-dimensional Einstein gravity with positive cosmological constant. For spacetime manifolds $\mathcal{M}$ without boundary, this theory’s action is just the Einstein-Hilbert action

$$S_{EH}[g] = \frac{1}{16\pi G} \int_{\mathcal{M}} d^{3}x \sqrt{-g} (R - 2\Lambda). \quad (2.4)$$

As observed by Gibbons and Hawking and by York, if the spacetime $\mathcal{M}$ has a boundary $\mathcal{\partial M}$, then one must supplement the action (2.4) with a boundary term to ensure that variation of the total action leads to well posed equations of motion [98, 240]. The type of boundary conditions that one wants to enforce dictates the particular boundary term that one must include. If one wants to hold fixed the metric induced on the boundary $\mathcal{\partial M}$—the situation that we study below—then one must add to the action (2.4) the Gibbons-Hawking-York boundary term

$$S_{GHY}[g] = \frac{1}{8\pi G} \int_{\mathcal{\partial M}} d^{2}y \sqrt{\gamma} |K|. \quad (2.5)$$

Here, $\gamma$ is the determinant of the metric induced on the boundary $\mathcal{\partial M}$, $y$ denotes a set of coordinates on the boundary $\mathcal{\partial M}$, and $K$ is the trace of the extrinsic curvature of the boundary $\mathcal{\partial M}$. We thus employ the complete action

$$S[g] = S_{EH}[g] + S_{GHY}[g] \quad (2.6)$$
in the path integral (2.2).

To apply the prescription of causal dynamical triangulations, we require the Regge calculus expression for the action (2.6). Regge himself demonstrated that for a triangulation $\mathcal{T}$ the Einstein-Hilbert action (2.4) assumes the form \[^6\]

$$S_{EH}[\mathcal{T}] = \frac{1}{8\pi G} \sum_{h \in \mathcal{T}} V_1^{(h)} \delta_h - \frac{\Lambda}{8\pi G} \sum_{s \in \mathcal{T}} V_3^{(s)}.$$  

(2.7)

Here, $h$ is a hinge—a 1-simplex in 2+1 dimensions—of 1-volume $V_1^{(h)}$ and deficit angle $\delta_h$, and $V_3^{(s)}$ is the spacetime 3-volume of a 3-simplex $s$. Hartle and Sorkin later determined the form of the Gibbons-Hawking-York boundary term in Regge calculus [114]. These authors showed that

$$S_{GHY}[\mathcal{T}] = \frac{1}{8\pi G} \sum_{h \in \partial \mathcal{T}} V_1^{(h)} \psi_h,$$

(2.8)

where $h$ is hinge on the boundary $\partial \mathcal{T}$ of the triangulation $\mathcal{T}$ having 1-volume $V_1^{(h)}$, and $\psi_h$ is the angle between the two vectors normal to the two spacelike 2-simplices intersecting at the hinge $h$. See figure 2.2.

We now translate the expressions (2.7) and (2.8) into the Regge calculus of causal triangulations for 2-sphere spatial topology and line interval temporal topology.\[^7\] The reader not interested in the details of this construction may skip to equation (2.15) in which we

\[^6\]For an accessible introduction to Regge calculus, see the relevant chapter in [161].

\[^7\]Ambjørn et al performed this translation for the variant of $(2 + 1)$-dimensional causal dynamical triangulations considered in [28, 29].
report our result. For an arbitrary causal triangulation $T_c$ the Einstein-Hilbert action is

$$S_{EH}[T_c] = \frac{1}{8\pi G} \sum_{h^{SL} \in T_c} a \delta_{h^{SL}} + \frac{1}{8\pi G} \sum_{h^{TL} \in T_c} \sqrt{\alpha} a \delta_{h^{TL}} - \frac{\Lambda}{8\pi G} \sum_{s \in T_c} V^{(s)}_3$$  \tag{2.9}$$

since the 1-volume $V_1(h^{SL})$ of a spacelike hinge $h^{SL}$ is $a$, and the 1-volume $V_1(h^{TL})$ of a timelike hinge $h^{TL}$ is $\sqrt{\alpha} a$. Ambjørn et al derived the specific form of the Einstein-Hilbert action for the case of spacetime topology $S^2 \times S^1$, finding that

$$S_{EH}[T_c] = \frac{a}{8\pi G} \left[ \frac{2\pi}{i} N_{1}^{SL} - \frac{2}{i} \theta_{SL}^{(2,2)} N_{3}^{(2,2)} \right.$$

$$- \frac{4}{i} \theta_{SL}^{(3,1)} N_{1}^{SL} + 2\pi \sqrt{\alpha} N_{1}^{TL} - 4\sqrt{\alpha} \theta_{TL}^{(2,2)} N_{3}^{(2,2)}$$

$$- 3\sqrt{\alpha} \theta_{TL}^{(1,3)} N_{3}^{(1,3)} - 3\sqrt{\alpha} \theta_{TL}^{(3,1)} N_{3}^{(3,1)} \left. \right] \tag{2.10}$$

$$- \frac{\Lambda}{8\pi G} \left[ V_3^{(2,2)} N_{3}^{(2,2)} + V_3^{(1,3)} N_{3}^{(1,3)} + V_3^{(3,1)} N_{3}^{(3,1)} \right].$$

Here, $N_{1}^{SL}$ is the number of spacelike 1-simplices, $N_{1}^{TL}$ is the number of timelike 1-simplices, $\theta_{SL}^{(p,q)}$ is the Lorentzian dihedral angle about a spacelike edge of a $(p, q)$ 3-simplex, $\theta_{TL}^{(p,q)}$ is the Lorentzian dihedral angle about a timelike edge of a $(p, q)$ 3-simplex, and $V_3^{(p,q)}$ is the Lorentzian spacetime volume of a $(p, q)$ 3-simplex. We refer the reader to [22, 31] for explicit expressions for these Lorentzian dihedral angles and spacetime volumes. Within the first set of square brackets, the first three terms stem from the summation over spacelike hinges, and the last four terms stem from the summation over timelike hinges. The terms within the second set of square brackets stem from the summation over
3-simplices. As we demonstrate below, we must modify the action (2.10) to account for the presence of the boundaries.

We now derive the Gibbons-Hawking-York boundary term in this case. Given the spacetime topology $S^2 \times I$, the boundary $\partial T_c$ of a causal triangulation $T_c$ consists of two disconnected components: an initial or past spatial 2-sphere $S^2_i$ and a final or future spatial 2-sphere $S^2_f$. Starting from the prescription (2.8) of Hartle and Sorkin [114] and drawing on similar constructions by Dittrich and Loll [82] and Anderson et al [31], we propose the prescription

$$S_{GHY}[T_c] = \alpha \frac{8\pi G}{\pi} \sum_{h \in S^2_i} \frac{1}{i} \left[ \pi - 2\theta^{(3,1)}_{SL} - \theta^{(2,2)}_{SL} N^{(2,2)}_{3\uparrow}(h) \right]$$

$$+ \alpha \frac{8\pi G}{\pi} \sum_{h \in S^2_f} \frac{1}{i} \left[ \pi - 2\theta^{(1,3)}_{SL} - \theta^{(2,2)}_{SL} N^{(2,2)}_{3\downarrow}(h) \right].$$  \tag{2.11}

Here, $N^{(2,2)}_{3\uparrow}(h)$ is the number of future-directed $(2,2)$ 3-simplices attached to the hinge $h$, and $N^{(2,2)}_{3\downarrow}(h)$ is the number of past-directed $(2,2)$ 3-simplices attached to the hinge $h$. We justify this prescription as follows. In parallel transporting the vector normal to one component of the boundary $\partial T_c$ between two spacelike 2-simplices intersecting at the hinge $h$, the vector rotates through the angle

$$\frac{1}{i} \left[ 2\theta^{(3,1)}_{SL} + \theta^{(2,2)}_{SL} N^{(2,2)}_{3\uparrow}(h) \right].$$  \tag{2.12}

See figure 2.2. When this angle is $\frac{\pi}{4}$, the extrinsic curvature vanishes locally at the hinge.
Figure 2.2: An illustration of the construction of the extrinsic curvature for the hinge $h$ (black) indicating the angle $\psi_h$ through which the normal vector rotates for the case of $N_{3l}^{(2,2)}(h) = 1$.

$h$; this fact dictates the deficit angle-like form of our above prescription. One might have expected a relative negative sign in the Gibbons-Hawking-York boundary term (2.11) between the contributions from the two disconnected components of the boundary. Its absence stems from the following fact: the future-directed orientation of the vector normal to $\mathcal{S}^2_i$ and the past-directed orientation of the vector normal to $\mathcal{S}^2_j$ are accounted for in the past-directed and future-directed orientations of the $(2, 2)$ 3-simplices attached to the boundaries. Performing the summations over the hinges on the boundary $\partial\mathcal{T}_c$, we may
rewrite the Gibbons-Hawking-York boundary term as

\[
S_{GHY}[\mathcal{T}_c] = \frac{a}{8\pi G} \left[ \frac{\pi}{i} N^{SL}_1(S^2_i) - \frac{2}{i} \theta^{(3,1)}_{SL} N^{SL}_1(S^2_i) - \frac{1}{i} \theta^{(2,2)}_{SL} N^{(2,2)}_{3\uparrow} (S^2_i) \right]
\]

\[
+ \frac{a}{8\pi G} \left[ \frac{\pi}{i} N^{SL}_1(S^2_f) - \frac{2}{i} \theta^{(1,3)}_{SL} N^{SL}_1(S^2_f) - \frac{1}{i} \theta^{(2,2)}_{SL} N^{(2,2)}_{3\downarrow} (S^2_f) \right].
\] (2.13)

Before writing down the complete action for \((2 + 1)\)-dimensional causal dynamical triangulations with 2-sphere spatial topology and line interval temporal topology, we must account for the effect of the boundary on the Einstein-Hilbert term (2.10). In particular, the spacelike hinges on the boundary \(\partial \mathcal{T}_c\) no longer contribute to the Einstein-Hilbert term because full spacetime parallel transport about those hinges is no longer well defined. Accordingly, we modify the terms stemming from the summation over spacelike hinges—the first three terms in the first set of square brackets in equation (2.10)—of the above prescription for the Einstein-Hilbert term to

\[
\frac{a}{8\pi G} \left[ \frac{2\pi}{i} \left( N^{SL}_1 - N^{SL}_1 (S^2_i) - N^{SL}_1 (S^2_f) \right) - \frac{1}{i} \theta^{(2,2)}_{SL} \left( 2 N^{(2,2)}_3 - N^{(2,2)}_3 (S^2_i) - N^{(2,2)}_3 (S^2_f) \right) \right]
\]

\[
- \frac{1}{i} \theta^{(1,3)}_{SL} \left( 4 N^{SL}_1 - 2 N^{SL}_1 (S^2_i) - 2 N^{SL}_1 (S^2_f) \right) \right].
\] (2.14)

The numerical factors appearing in the second and third terms of this expression require some explanation. In the second term \(N^{(2,2)}_3 (S^2_i)\) and \(N^{(2,2)}_3 (S^2_f)\) enter with the factor 1 instead of 2 because only one of the two spacelike edges of these \((2, 2)\) 3-simplices attaches
to the boundary $\partial T_c$. In the third term $N^{SL}_1(S^2_i)$ and $N^{SL}_1(S^2_f)$ enter with the factor 2 instead of 4 because only two $(1,3)$ or $(3,1)$ 3-simplices attach to a hinge on the boundary $\partial T_c$.

The complete action is thus

$$S_R[T_c] = \frac{a}{8\pi G} \left[ \frac{2\pi}{i} \left( N^{SL}_1 - N^{SL}_1(S^2_i) - N^{SL}_1(S^2_f) \right) ight]$$

$$- \frac{1}{i} \theta_{SL}^{(2,2)} \left( 2N_{3}^{(2,2)} - N_{3}^{(2,2)}(S^2_i) - N_{3}^{(2,2)}(S^2_f) \right)$$

$$- \frac{1}{i} \theta_{SL}^{(1,3)} \left( 4N^{SL}_1 - 2N^{SL}_1(S^2_i) - 2N^{SL}_1(S^2_f) \right)$$

$$+ 2\pi \sqrt{\alpha} N^{TL}_1 - 4\sqrt{\alpha} \theta_{TL}^{(2,2)} N^{(2,2)}_3$$

$$- 3\sqrt{\alpha} \theta_{TL}^{(1,3)} N^{(1,3)}_3 - 3\sqrt{\alpha} \theta_{TL}^{(3,1)} N^{(3,1)}_3 \right]$$

$$- \frac{\Lambda}{8\pi G} \left[ V_3^{(2,2)} N^{(2,2)}_3 + V_3^{(1,3)} N^{(1,3)}_3 + V_3^{(3,1)} N^{(3,1)}_3 \right]$$

$$+ \frac{a}{8\pi G} \left[ \frac{\pi}{i} N^{SL}_1(S^2_i) - \frac{2}{i} \theta_{SL}^{(3,1)} N^{SL}_1(S^2_i) - \frac{1}{i} \theta_{SL}^{(2,2)} N^{(2,2)}_3 \right]$$

$$+ \frac{a}{8\pi G} \left[ \frac{\pi}{i} N^{SL}_1(S^2_f) - \frac{2}{i} \theta_{SL}^{(1,3)} N^{SL}_1(S^2_f) - \frac{1}{i} \theta_{SL}^{(2,2)} N^{(2,2)}_3 \right].$$

Although at first glance the action (2.15) does not appear to be real, it is in fact real for $\alpha > 0$ given the expressions for the Lorentzian dihedral angles. In appendix 2.A we demonstrate that the action (2.15) is consistent with the composition of two causal triangulations sharing a common boundary. We use the action (2.15) in the path sum (2.3) to compute transition amplitudes $\mathcal{A}[$\$\Gamma(S^2_i), \Gamma(S^2_f)\$]$ between the initial and final spacelike boundaries $S^2_i$ and $S^2_f$ with fixed intrinsic geometries $\Gamma(S^2_i)$ and $\Gamma(S^2_f)$. 


2.4.3 Numerics

Although the formalism of causal dynamical triangulations has turned the computation of the path sum (2.3) for the action (2.15) into a well defined problem in combinatorics, solving this problem analytically even in $2 + 1$ dimensions is difficult. To study the transition amplitudes defined by the path sum (2.3), we thus turn to numerical methods, in particular, Markov chain Monte Carlo simulations of representative paths. To perform such numerical simulations, we must transform the path sum (2.3) into a partition function with real as opposed to complex summands. As we noted above, the structure of causal triangulations is specifically adapted to such a transformation, allowing a simple Wick rotation from Lorentzian to Euclidean signature. This Wick rotation consists in analytically continuing the parameter $\alpha$ to $-\alpha$ in the lower half complex plane [22]. The path sum (2.3) thus becomes the partition function

$$\mathcal{Z}[\Gamma] = \sum_{\mathcal{T}_c} \mu(\mathcal{T}_c) e^{-S^{(E)}_R[\mathcal{T}_c]}$$ (2.16)
for the Euclidean action

\[ S_R^{(E)}[\mathcal{T}_c] = \frac{ia}{8\pi G} \left[ \frac{2\pi}{i} \left( N^{SL}_1 - N^{SL}_1(S^2_i) - N^{SL}_1(S^2_f) \right) \right. \]
\[ - \frac{1}{i} \vartheta^{(2,2)}_{SL} \left( 2N^{(2,2)}_3 - N^{(2,2)}_3(S^2_i) - N^{(2,2)}_3(S^2_f) \right) \]
\[ - \frac{1}{i} \vartheta^{(1,3)}_{SL} \left( 4N^{SL}_1 - 2N^{SL}_1(S^2_i) - 2N^{SL}_1(S^2_f) \right) \]
\[ - 2\pi i\sqrt{-\alpha}N^T_L + 4i\sqrt{-\alpha}\vartheta^{(2,2)}_{TL}N^{(2,2)}_3 \]
\[ + 3i\sqrt{-\alpha}\vartheta^{(1,3)}_{TL}N^{(1,3)}_3 + 3i\sqrt{-\alpha}\vartheta^{(3,1)}_{TL}N^{(3,1)}_3 \]  \hspace{1cm} (2.17)
\[ - \frac{i\Lambda}{8\pi G} \left[ V^{(2,2)}_3 N^{(2,2)}_3 + V^{(1,3)}_3 N^{(1,3)}_3 + V^{(3,1)}_3 N^{(3,1)}_3 \right] \]
\[ + \frac{ia}{8\pi G} \left[ \frac{\pi}{i} N^{SL}_1(S^2_i) - \frac{2}{i} \vartheta^{(3,1)}_{SL} N^{SL}_1(S^2_i) - \frac{1}{i} \vartheta^{(2,2)}_{SL} N^{(2,2)}_3(S^2_i) \right] \]
\[ + \frac{ia}{8\pi G} \left[ \frac{\pi}{i} N^{SL}_1(S^2_f) - \frac{2}{i} \vartheta^{(3,1)}_{SL} N^{SL}_1(S^2_f) - \frac{1}{i} \vartheta^{(2,2)}_{SL} N^{(2,2)}_3(S^2_f) \right] , \]

Here, \( \vartheta^{(p,q)}_{SL} \) is the Euclidean dihedral angle about a (still) spacelike edge of a \( (p, q) \) 3-simplex, \( \vartheta^{(p,q)}_{TL} \) is the Euclidean dihedral angle about a (formerly) timelike edge of a \( (p, q) \) 3-simplex, and \( V^{(p,q)}_3 \) is the Euclidean spacetime 3-volume of a \( (p, q) \) 3-simplex. We refer the reader to [22, 31] for explicit expressions for these Euclidean dihedral angles and spacetime volumes. When the initial and final spatial 2-spheres \( S^2_i \) and \( S^2_f \) are identified, in which case the spacetime topology is \( S^2 \times S^1 \), the action (2.17) assumes the simple form

\[ S_R^{(E)}[\mathcal{T}_c] = -k_0N_0 + k_3N_3 \]  \hspace{1cm} (2.18)
for the couplings

\begin{align}
  k_0 &= 2\pi a k \\
  k_3 &= \frac{a^3 \lambda}{4\sqrt{2}} + 2\pi a k \left[ \frac{3}{\pi} \cos^{-1} \left( \frac{1}{3} \right) - 1 \right]
\end{align}

(2.19a) \quad (2.19b)

with \( k = \frac{1}{8\pi G}, \lambda = \frac{\Lambda}{8\pi G}, \) and \( \alpha = -1 \) [22, 23]. In 2+1 dimensions the numerical value of \( \alpha \) is irrelevant (so long as \( |\alpha| > \frac{1}{2} \)) because topological constraints render its value redundant [22, 45]. We thus set \( \alpha = -1 \) in the following. When referring to an ensemble of causal triangulations with fixed initial and final boundaries \( S^2_i \) and \( S^2_f \), we employ the couplings \( k_0 \) and \( k_3 \) instead of the couplings \( k \) and \( \lambda \) to make contact with previous work. By the given values of \( k_0 \) and \( k_3 \), we mean the values dictated by the relations (2.19) for the values of \( k \) and \( \lambda \) actually characterizing the given ensemble.

We perform numerical simulations as follows. The partition function (2.16) involves a sum over all causal triangulations with the specified spacetime topology including those with arbitrarily large numbers \( T \) of time slices and \( N_3 \) of 3-simplices. For a given value of the coupling \( k_0 \), however, the partition function (2.16) defines a critical surface for a critical value \( k_3^c \) of the coupling \( k_3 \) on which the amplitude \( \mu(T_c)e^{-S^{(E)}_{R}[T_c]} \) is peaked for causal triangulations with particular fixed numbers \( T \) of time slices and \( N_3 \) of 3-simplices. Since we cannot simulate causal triangulations of arbitrarily large size, we work on a particular critical surface on which \( T \) is precisely fixed and \( N_3 \) is approximately fixed. We thus
numerically simulate the partition function

\[
Z[\Gamma] = \sum_{T_c, T_f, N_3} \mu(T_c) e^{-S_R^{(E)}(T_c)}
\]

(2.20)

for the action (2.17). The partition function (2.20) at fixed \(T\) and \(N_3\) is related by a Legendre transform to the partition function (2.16) without these constraints [17]:

\[
Z[\Gamma] = \sum_{T, N_3} e^{-k_0 N_3} Z[\Gamma].
\]

(2.21)

Following standard techniques, we generate multiple ensembles at different values of \(T\) and \(N_3\) and explore how physical observables finite size scale towards the continuum.\(^8\)

Evidently, the dependence of \(T\) and \(N_3\) on \(k_0\) is quite weak, so we can essentially select \(T\) and \(N_3\) independently of \(k_0\) to within the numerical precision of \(k_0\). Accordingly, we first select the number \(T\) of time slices, the target number \(\bar{N}_3\) of 3-simplices, and the value of the coupling \(k_0\).\(^9\)

We next construct the initial and final boundary 2-spheres \(S^2_i\) and \(S^2_f\) to each possess the particular fixed intrinsic geometries \(\Gamma(S^2_i)\) and \(\Gamma(S^2_f)\). To incorporate the initial

---

\(^8\)For a discussion of finite size scaling see, e.g., [168].

\(^9\)One might be concerned that the transition amplitudes under consideration are not well defined with the number \(T\) of time slices fixed. For the classical thick sandwich problem one does not specify any such property of the spacetime interpolating between the initial and final spacelike hypersurfaces. Indeed, the thick sandwich problem is generically not well posed if one specifies data beyond the intrinsic geometries of the initial and final spacelike hypersurfaces. We believe that this concern is not necessarily a problem within the quantum theory defined by causal dynamical triangulations. Previous results—specifically, the formation of so-called stalks—suggest that the temporal extent of the quantum spacetime geometry is determined dynamically.
and final boundaries into a causal triangulation with $T$ time slices, we create a minimal causal triangulation over the time slices labelled by the Euclidean discrete time coordinate $\tau \in \{2, \ldots, T - 1\}$. Within a minimal causal triangulation each time slice is a minimally triangulated 2-sphere—the surface of a tetrahedron—and adjacent time slices are connected by the minimal number of 3-simplices. We then employ an algorithm, described in appendix 2.B, that connects the initial and final 2-spheres, labelled by $\tau = 1$ and $\tau = T$, to their adjacent time slices.\footnote{We gratefully acknowledge David Kamensky for developing this algorithm.} We finally apply the spacetime 3-volume-increasing Pachner moves, always holding the boundary intrinsic geometries fixed, to raise the total number $N_3$ of 3-simplices to the target value $\bar{N}_3$. Fortuitously, none of the Pachner moves can change the boundaries’ intrinsic geometries.\footnote{Currently, we cannot design a boundary 2-sphere with an arbitrarily chosen intrinsic geometry; nevertheless, the algorithm described in appendix 2.B can take as input any boundary 2-sphere.}

With the conditions for the simulation of a particular transition amplitude established, we now tune the coupling $k_3$ to its critical value $k_3^c$ so that we work on the critical surface on which the partition function (2.20) is well defined. Portions of this critical surface may well fall into different phases of the partition function (2.20). Below we study exclusively phase C, the de Sitter-like phase of causal dynamical triangulations for 2-sphere spatial topology. Floating point imprecision essentially guarantees that we cannot tune precisely to the critical surface, so we add to the action (2.17) a Lagrange multiplier term of the form $\epsilon |N_3 - \bar{N}_3|$ for Lagrange multiplier $\epsilon \ll 1$. Effectively, this term smears the critical surface into a thin critical volume. In the simulations reported below we have set $\epsilon = 0.02$.

Once this initialization process is complete, we run a standard Metropolis-Hastings al-
algorithm based on the partition function (2.20) for the action (2.17). We employ the Pachner moves to update our simulations at each step, thereby moving through the space of causal triangulations with fixed boundaries, fixed $T$, and approximately fixed $N_3$. After a period of thermalization, we generate an ensemble of causal triangulations representative of their relative weightings in the partition function (2.20) by sampling every five hundred sweeps, one sweep consisting of one attempted Pachner move for every 3-simplex in the current causal triangulation. Once we have generated an ensemble, containing of order $10^5$ members in the results reported below, we estimate the expectation values of observables as ensemble averages.

Our computer code is a version of that reported in [137] modified to account for initial and final 2-sphere boundaries with fixed intrinsic geometries.\textsuperscript{12} We can readily disable its fixed boundary functionality to allow for periodic boundary conditions, and we have verified that we recover results quantitatively in agreement with those reported previously.\textsuperscript{13} We believe that the results presented in the next section provide further evidence that our modified code is working properly. Although all of these results are for phase C, we have run our code with fixed boundaries for values of the coupling $k_0$ that should fall within phase A of the partition function (2.20), and we do indeed find that this is the case.

\textsuperscript{12}Our implementation uses the standard Mersenne Twister algorithm [155] as implemented in Steel Bank Common Lisp [162] for random number generation. The Mersenne Twister algorithm is demonstrably robust and has a period of $2^{19937} − 1$.

\textsuperscript{13}We have compared our results to those reported in [137], which uses the same codebase and to those reported in, e.g., [15, 14], which was developed independently.
2.5 Transition amplitudes

We now report on our numerical simulation of a variety of transition amplitudes by the methods of causal dynamical triangulations. As discussed above, we exclusively consider the transition amplitudes $Z[\Gamma(S^2_i), \Gamma(S^2_f)]$ from an initial spatial 2-sphere $S^2_i$ of fixed intrinsic geometry $\Gamma(S^2_i)$ to a final spatial 2-sphere $S^2_f$ of fixed intrinsic geometry $\Gamma(S^2_f)$, and we work at fixed numbers $T$ of time slices and $N_3$ of 3-simplices. Specifically, we explore three classes of such transition amplitudes: from minimally triangulated 2-sphere to minimally triangulated 2-sphere, from minimally triangulated 2-sphere to nonminimally triangulated 2-sphere, and from nonminimally triangulated 2-sphere to nonminimally triangulated 2-sphere. Thus far we have limited our consideration to the dependence of the transition amplitudes $Z[\Gamma(S^2_i), \Gamma(S^2_f)]$ on the discrete spatial 2-volumes $N^{SL}_2(S^2_i)$ and $N^{SL}_2(S^2_f)$ of the initial and final boundaries as measured by the number $N^{SL}_2$ of spacelike 2-simplices composing each triangulated 2-sphere. That is, we have averaged the transition amplitudes $Z[\Gamma(S^2_i), \Gamma(S^2_f)]$ over all of the geometrical degrees of freedom of $S^2_i$ and $S^2_f$ except for their discrete spatial 2-volumes $N^{SL}_2(S^2_i)$ and $N^{SL}_2(S^2_f)$, obtaining the transition amplitudes $Z[N^{SL}_2(S^2_i), N^{SL}_2(S^2_f)]$.

We numerically simulate the geometry-averaged transition amplitudes $Z[N^{SL}_2(S^2_i), N^{SL}_2(S^2_f)]$ as follows. We first employ a separate Markov chain Monte Carlo algorithm to generate a finite set of pairs of initial and final boundary 2-spheres. Each 2-sphere has the desired discrete spatial 2-volume but otherwise random intrinsic geometry.
drawn from the uniform distribution. For each pair of initial and final boundary 2-spheres, we then implement the numerical algorithm described above to generate an ensemble of representative causal triangulations. Finally, we compute ensemble average observables by averaging over the ensembles of representative causal triangulations for every pair of initial and final boundary 2-spheres. In particular, we have employed sets consisting of ten pairs of initial and final boundary 2-spheres.

We have elected to study the transition amplitudes $Z[N_{2}^{SL}(S_{i}^{2}), N_{2}^{SL}(S_{f}^{2})]$ for two primary reasons. First of all, the numerical simulation of the full transition amplitudes $Z[\Gamma(S_{i}^{2}), \Gamma(S_{f}^{2})]$ necessitates methods for both characterization and generation of specific boundary geometries. We are in the process of developing such methods, but our numerical implementation does not currently possess these capabilities. Moreover, most of the analytical computations of these transition amplitudes have been performed within a minisuperspace truncation of the metric degrees of freedom. Restricting consideration to the geometry-averaged transition amplitudes $Z[N_{2}^{SL}(S_{i}^{2}), N_{2}^{SL}(S_{f}^{2})]$ thus facilitates comparison of our results to those in the literature.

To study these transition amplitudes, we primarily perform numerical measurements of the observable $N_{2}^{SL}(\tau)$, the number $N_{2}^{SL}$ of spacelike 2-simplices as a function of the discrete Euclidean time coordinate $\tau$. The value $N_{2}^{SL}$ for a particular value of $\tau$ is not an observable, even in the presence of fixed spacelike boundaries, since the value of $\tau$ is merely a labeling convention, typically differing from one representative causal triangulation.

---

14Drawing the intrinsic geometries of the boundary 2-spheres from the uniform distribution represents an assumption on our part, which we deemed acceptable for an initial investigation. We believe that the ensuing results—in particular, their agreement with previous results—justify our choice.
tion to another. The set of values $N^{SL}_2(\tau)$ for all values of $\tau$ is an observable: only given the dynamics of $N^{SL}_2$ can one choose a consistent convention for the label $\tau$ across all representative causal triangulations.\(^{15}\) This observable essentially gives the discrete time evolution of the discrete spatial 2-volume.

By focusing on the observable $N^{SL}_2(\tau)$, we are anticipating that a minisuperspace truncation of the metric degrees of freedom will prove a good description of our results. As we remarked briefly above, there is now ample evidence that this truncation accurately describes the ground state geometry within phase C. In particular, the ensemble average spacetime geometry of this ground state on sufficiently large scales, characterized by the observable $N^{SL}_2(\tau)$, is remarkably well described by the gravitational effective action [9, 15, 14, 18, 23, 24, 26, 25]

$$S^{(E)}_{\text{eff}}[N^{SL}_2(\tau)] = c_1 \sum_{\tau=1}^{T} \left[ \frac{1}{N^{SL}_2(\tau)} \left( \frac{\Delta N^{SL}_2(\tau)}{\Delta \tau} \right)^2 + c_2 \right]. \quad (2.22)$$

Here, $c_1$ and $c_2$ are phenomenological couplings, and $\frac{\Delta}{\Delta \tau}$ represents an appropriate discrete time derivative. The effective action (2.22) is the discrete analogue of the Einstein-Hilbert action

$$S^{(E)}_{\text{EH}}[\rho(\eta)] = \frac{1}{2G} \int d\eta \sqrt{g_{\eta\eta}} \left[ g^{\eta\eta} \left( \frac{d\rho(\eta)}{d\eta} \right)^2 + 1 - \Lambda \rho^2(\eta) \right]. \quad (2.23)$$

\(^{15}\)There may be some gauge redundancy in the observable $N^{SL}_2(\tau)$, but this observable does contain physical information.
for a Euclidean minisuperspace model described by the line element

\[ ds^2 = g_{\eta\eta}d\eta^2 + \rho^2(\eta) \left( d\theta^2 + \sin^2\theta d\phi^2 \right) \] (2.24)

for scale factor \( \rho \) as a function of the global time coordinate \( \eta \). The discrete analogue of the Lagrange multiplier term \( \Lambda \rho^2(\eta) \) in the action (2.23) does not appear in the effective action (2.22) because each causal triangulation within a given ensemble is characterized by a fixed number \( \bar{N}_3 \) of 3-simplices. We retain the constant factor \( g_{\eta\eta} \) for convenience in making a correspondence with the discrete effective action (2.22). The maximally symmetric spacetime characterized by the line element (2.24) is Euclidean de Sitter spacetime for which

\[ \rho(\eta) = l_{dS} \cos \left( \frac{\sqrt{g_{\eta\eta}\eta}}{l_{dS}} \right) \] (2.25)

for the de Sitter radius \( l_{dS} = \sqrt{\frac{1}{3}} \). The solution (2.25) describes a 3-sphere of radius \( l_{dS} \).

The ensemble average \( \langle N_{2}^{SL}(\tau) \rangle \) is exceedingly well fit by the discrete analogue of the time evolution of the spatial 2-volume \( V_2(\eta) = 4\pi \rho^2(\eta) \) of Euclidean de Sitter spacetime [9, 15, 14, 18, 23, 24, 26, 25, 31, 45, 137], explicitly

\[ N_{2}^{SL}(\tau) = \frac{2}{\pi} \frac{\langle N_3^{(1.3)} \rangle \cos^2 \left( \frac{\tau}{\bar{s}_0 \langle N_3^{(1.3)} \rangle^{1/3}} \right)}{\bar{s}_0 \langle N_3^{(1.3)} \rangle^{1/3}}. \] (2.26)

We demonstrate the explicit correspondence between the function (2.26) and the scale factor of Euclidean de Sitter spacetime in appendix 2.C. The parameter \( \bar{s}_0 \) is fit to the particular
ensemble of causal triangulations, serving to relate the dimensionless de Sitter radius to the scale of the discrete time coordinate $\tau$.

Assuming for the moment that this minisuperspace truncation similarly applies to the transition amplitudes treated below, we consider briefly the semiclassical expectations for these transition amplitudes. Several authors have argued that one can compute the Lorentzian path integral (2.2) by performing an appropriate Euclidean path integral. (See, for instance, [113].) When considering the Euclidean path integral within the minisuperspace truncation, one typically calculates the semiclassical approximation to this path integral by the method of steepest descents. For the transition amplitudes of interest, this path integral is

$$A[\rho_i, \rho_f] = \int_{\rho(\eta_i)=\rho_i}^{\rho(\eta_f)=\rho_f} \mathcal{D}\rho(\eta) e^{-S_{E\text{H}}^{(E)}[\rho(\eta)]}. \quad (2.27)$$

We consider the semiclassical predictions of the path integral (2.27) for each of the three classes of transition amplitudes explored below.

For the transition amplitudes from minimally triangulated 2-sphere to minimally triangulated 2-sphere, corresponding to $\rho_i = 0$ and $\rho_f = 0$, Euclidean de Sitter spacetime is the extremum of the action (2.23) that dominates the steepest descents approximation of the path integral (2.27). See figure 2.3(a). This accords of course with previous results for the ground state within phase C. For the transition amplitudes from minimally triangulated 2-sphere to nonminimally triangulated 2-sphere, corresponding to the Hartle-Hawking no-boundary scenario of $\rho_i = 0$, the dominant extrema of the action (2.23) depend on the relative values of $\rho_f$ and $l_{\text{dS}}$. If $\rho_f < l_{\text{dS}}$, then there are two real extrema of the action
Figure 2.3: Extrema of the action (2.23) for $\rho_i < l_{dS}$ and $\rho_f < l_{dS}$ depicted for one fewer spacelike dimension. (a) $\rho_i = 0$, $\rho_f = 0$ (b) $\rho_i = 0$, $\rho_f \neq 0$ (c) $\rho_i = 0$, $\rho_f \neq 0$ (d) $\rho_i \neq 0$, $\rho_f \neq 0$

(2.23) corresponding to the two portions of Euclidean de Sitter spacetime with final scale factor $\rho_f$, one containing less and one containing more than half of the spacetime 3-volume of Euclidean de Sitter spacetime. See figures 2.3(b) and 2.3(c). If $\rho_f = l_{dS}$, then these two extrema coincide. If $\rho_f > l_{dS}$, then there are two dominant complex extrema of the action (2.23), yielding an oscillatory transition amplitude. The envelope of these oscillations gives a probability distribution for $\rho_f$ consistent with that of Lorentzian de Sitter spacetime. For the transition amplitude from nonminimally triangulated 2-sphere to nonminimally triangulated 2-sphere, the dominant extrema of the action depend on the relative values of $\rho_i$, $\rho_f$, and $l_{dS}$. In particular, if $\rho_i < l_{dS}$ and $\rho_f < l_{dS}$, then the dominant extremum of the action (2.23) corresponds to a portion of Euclidean de Sitter spacetime with initial scale factor $\rho_i$ and final scale factor $\rho_f$. See figure 2.3(d). If $\rho_i > l_{dS}$ and $\rho_f > l_{dS}$, then the dominant extrema of the action (2.23) are complex, again yielding an oscillatory transition amplitude. The envelope of these oscillations gives a probability distribution for $\rho_i$ and $\rho_f$ consistent with that of Lorentzian de Sitter spacetime.
Given that the Markov chain Monte Carlo method yields an ensemble of representative paths contributing to the path integral and that the semiclassical contributions dominate the path integral, we expect our numerical simulations to output paths close to the semiclassical expectation. In analyzing transition amplitudes within causal dynamical triangulations, we attempt to determine if this is the case.

### 2.5.1 Minimal initial and final boundaries

We first take both the initial boundary $S_i^2$ and the final boundary $S_f^2$ as the minimal triangulation of the 2-sphere, namely the surface of the regular tetrahedron. (Recall that the causality condition on causal triangulations dictates that every time slice has the topology of a 2-sphere.) The minimal triangulation of the 2-sphere represents the best approximation to a no-boundary boundary condition within causal dynamical triangulations. For ensembles characterized by a sufficiently large number $\tilde{N}_3$ of 3-simplices, the nonzero spatial extent of the minimal boundary is presumably negligible, as our numerical simulations attest.

Before displaying the results of our simulations, we recall the findings of two simulations with periodic boundary conditions for the Euclidean discrete time coordinate $\tau$. One may regard these as simulations of vacuum persistence amplitudes as opposed to transition amplitudes. In figure 2.4(a) we display $\langle N_{2}^{SL}(\tau) \rangle$, the coherent ensemble average number $\langle N_{2}^{SL} \rangle$ of spacelike 2-simplices as a function of $\tau$, for an ensemble characterized by $T = 64$, $\tilde{N}_3 = 30850$, and $k_0 = 1.00$. Points indicate numerical measurements of $\langle N_{2}^{SL} \rangle$; error bars indicating the standard deviation $\sigma(\langle N_{2}^{SL} \rangle)$ are not visible on the scale of the
plot. Clearly, $\langle N_2^{SL}(\tau) \rangle$ possesses a distinctly regular form. There is a central accumulation of discrete spatial 2-volume spanning a significant portion of the time slices. The time slices beyond this accumulation constitute the so-called stalk in which each time slice has a near minimal number of spacelike 2-simplices. In figure 2.4(a), for instance, the central accumulation spans the time slices for which $\tau \in \{-14, \ldots, 14\}$, and the stalk spans the time slices for which $\tau \in \{-32, \ldots, -15\} \cup \{15, \ldots, 32\}$. This form of $\langle N_2^{SL}(\tau) \rangle$ is the characteristic feature of the de Sitter-like phase C. The curve in figure 2.4(a) is a fit of the function (2.26) to $\langle N_2^{SL}(\tau) \rangle$ within the central accumulation of discrete spatial 2-volume. Within the stalk we fit a constant function to $\langle N_2^{SL}(\tau) \rangle$, matching appropriately at the junctions with the central accumulation. In appendix 2.C we explain the details of this and related fits. The stalk is believed to be an artifact of the causality condition; interestingly, however, the effective action (2.22) evidently provides a good description of the stalk too [9, 18].

In figure 2.4(b) we display $N_2^{SL}(\tau)$ for a single representative causal triangulation from an ensemble characterized by $T = 28$, $\bar{N}_3 = 30850$, and $k_0 = 1.00$. Note that the central accumulation of discrete 2-volume in figure 2.4(a) extends over approximately twenty-eight time slices. In this case there are evidently too few time slices for the de Sitter-like phase to form. Instead, representative causal triangulations have an approximately uniform distribution of discrete spatial 2-volume over all of the time slices.

---

16When the Euclidean discrete time coordinate $\tau$ is periodically identified, the center of discrete spatial 2-volume of the central accumulation may occur at any value of $\tau$. Accordingly, to coherently average $N_2^{SL}(\tau)$ over an ensemble of causal triangulations, we temporally align each causal triangulation’s center of discrete spatial 2-volume [31].
We now consider a set of ensembles for which both the initial boundary $S^2_i$ and the final boundary $S^2_f$ are fixed to the minimal triangulation of the 2-sphere. In particular, we set $\bar{N}_3 = 30850$ and $k_0 = 1.00$ for successively fewer numbers $T$ of time slices. In figure 2.5 we depict $\langle N^SL_2(\tau) \rangle$ for three such ensembles with $T = 65$, $T = 37$, and $T = 29$ fit to the function (2.26). For these numbers $T$ of time slices, representative causal triangulations still possess a stalk—technically, two disconnected stalks, one on either side of the central accumulation—although its extent diminishes as $T$ decreases. We also note that we must still coherently average $N^SL_2(\tau)$ despite the fixed boundaries, which do not temporally align the centers of discrete spatial 2-volume in these cases.

In figure 2.6 we display $\langle N^SL_2(\tau) \rangle$ for four more such ensembles with $T = 25$, $T = 21$, $T = 17$, and $T = 13$ fit to the function (2.26). For these numbers $T$ of time slices, representative causal triangulations no longer possess stalks, and the fixed boundaries serve to
temporally align the centers of discrete spatial 2-volume, effectively performing the coherent averaging automatically. One can thus completely remove the stalk, demonstrating that it is indeed an artifact of the numerical implementation of the causality condition. Moreover, as one further reduces the number of time slices, $\langle N_2^{SL}(\tau) \rangle$ dynamically readjusts to fit the function (2.26) for different values of $\tilde{s}_0$.

Visually, all of the fits of the function (2.26) to $\langle N_2^{SL}(\tau) \rangle$ appear quite good, but we wish to ascertain the quality of these fits quantitatively. To this end we have determined the
Figure 2.6: Ensemble average number $\langle N_{2}^{SL} \rangle$ of spacelike 2-simplices as a function of the discrete time coordinate $\tau$ for $\bar{N}_3 = 30850$ and $k_0 = 1.00$. (a) $T = 25$ (b) $T = 21$ (c) $T = 17$ (d) $T = 13$

goodness of fit, as measured by the chi-squared per degree of freedom $\chi^2_{pdf}$—essentially the fit residual normalized by the number of time slices—for a variety of ensembles.\(^\text{17}\) In table 2.1 we list the value of $\chi^2_{pdf}$ for each of these ensembles. In appendix 2.C we explain the details of this statistical analysis.\(^\text{18}\)

\(^{17}\)Typically a $\chi^2_{pdf}$ value should be of order 1. If it is significantly above 1 the model fits the data poorly. If it is significantly below 1, then the model overfits the data. We find our $\chi^2_{pdf}$ values to be significantly less than 1. Since our model only has two parameters, it is unlikely to be overfit. However, we made a simplifying assumption in our analysis— that $\langle N_{2}^{SL} \rangle$ on a given discrete time slice is uncorrelated to $\langle N_{2}^{SL} \rangle$ at another time. If this assumption is wrong, then our $\chi^2_{pdf}$ values could be wrong as well. A more detailed analysis is necessary to understand this behaviour.

\(^{18}\)One may notice a significant range in the $\chi^2_{pdf}$ values listed in table 2.1. They range in order of magnitude
the fit parameter $T$ of time slices, the fractional temporal extent characterized by the spacetime topology, the target number $\bar{N}$ of the function (2.26) to $\chi^2$. A significant deviation from the fit function is visible in the bulk. This is the scenario for which $O$ from Table 2.1: The goodness of fit, as measured by the chi-squared per degree of freedom $\chi^2$ of the function (2.26) to $\langle N_2^{SL}(\tau) \rangle$ for a variety of ensembles of $N(T_c)$ causal triangulations characterized by the spacetime topology, the target number $\bar{N}$ of 3-simplices, the number $T$ of time slices, the fractional temporal extent $\Delta \tau_{stalk}/T$ of the stalk, the best fit value of the fit parameter $\bar{s}_0$, and the error $\epsilon(\bar{s}_0)$ in the fit parameter $\bar{s}_0$. If no error $\epsilon(\bar{s}_0)$ is reported, then $\epsilon(\bar{s}_0) < 10^{-6}$. For the $\chi^2$ values, smaller numbers are better.

We draw two conclusions from the values of $\chi^2_{pdf}$. First, at fixed target number $\bar{N}$ of 3-simplices and coupling $k_0$, there is a particular number $T^*$ of time slices that optimizes the fit of the function (2.26) to $\langle N_2^{SL}(\tau) \rangle$. For instance, we find that $T^* \approx 21$ for $\bar{N} = 30850$ and $k_0 = 1.00$. This optimization occurs for an ensemble of causal triangulations characterized by a sufficiently small $T$ such that stalks are absent. As we increase $T$ above $T^*$, gradually developing stalks, the goodness of fit first worsens for sufficiently short stalks and then improves for sufficiently long stalks. Second, one can achieve an

from $O(10^{-1})$ to $O(10^{-4})$. This is because the quality of fits varies dramatically. For example, in figure 2.6(d) a significant deviation from the fit function is visible in the bulk. This is the scenario for which $\chi^2_{pdf} \sim O(10^{-1})$
equivalent goodness of fit for a notably smaller target number $\bar{N}_3$ of 3-simplices by working with the optimal number $T^*$ of time slices. The values of the chi-squared per degree of freedom $\chi^2_{\text{pdf}}$ normalized by the number $N(T_c)$ of causal triangulations make this particularly evident. This fact could prove useful for future statistically detailed studies of causal dynamical triangulations.

2.5.2 Minimal initial boundary and nonminimal final boundary

We next take the initial boundary $S^2_i$ as the minimal triangulation of the 2-sphere and the final boundary $S^2_f$ as a nonminimal triangulation of the 2-sphere. In particular, for $T = 29$, $\bar{N}_3 = 30850$, and $k_0 = 1.00$, we consider transition amplitudes to final boundaries with $N_{2}^{SL}(S^2_f) \in \{4, 100, 300, 500, 700\}$. (The first such transition amplitude is of course one of those considered previously.) In figure 2.7 we display $\langle N_{2}^{SL}(\tau) \rangle$ for each of these transition amplitudes.

We wish to determine if the plots of $\langle N_{2}^{SL}(\tau) \rangle$ in figure 2.7 accord with the semiclassical expectations discussed above, which suggest that each transition amplitude represented in figure 2.7 should be dominated by a portion of Euclidean de Sitter spacetime. (Clearly, these transition amplitudes have $\rho_f < l_{dS}$. ) If this is the case, then we expect the function (2.26) to fit $\langle N_{2}^{SL}(\tau) \rangle$ for a restriction on the upper limit of the discrete time coordinate $\tau$. While the function (2.26) fits exceedingly well to $\langle N_{2}^{SL}(\tau) \rangle$ for $N_{2}^{SL}(S^2_f) = 4$, we readily observe that this does not naively hold true for $N_{2}^{SL}(S^2_f) > 4$. If, however, we identify the early time behavior of $\langle N_{2}^{SL}(\tau) \rangle$ as that of a stalk, then we can test whether or not a
Figure 2.7: Ensemble average number $\langle N_{2}^{\text{SL}} \rangle$ of spacelike 2-simplices as a function of the discrete time coordinate $\tau$ for $\bar{N}_3 = 30850$ and $k_0 = 1.0$.

The restriction of the function (2.26) fits the late time behavior of $\langle N_{2}^{\text{SL}}(\tau) \rangle$.

There still remains one further issue to address: in attempting these fits, what value do we take for the ensemble average number $\langle N_{3}^{(1,3)} \rangle$ of $(1, 3)$ 3-simplices, which sets the overall scaling of the function (2.26)? The value of $\langle N_{3}^{(1,3)} \rangle$ appearing in the function (2.26) represents the total number of $(1, 3)$ 3-simplices within all of Euclidean de Sitter spacetime. We thus propose taking $\langle N_{3}^{(1,3)} \rangle$ as the effective ensemble average number $\langle N_{3}^{(1,3)} \rangle_{\text{eff}}$ of $(1, 3)$ 3-simplices that would be present without the restriction on the upper limit of $\tau$. We compute $\langle N_{3}^{(1,3)} \rangle_{\text{eff}}$ as twice the number of $(1, 3)$ 3-simplices to the past of the time slice with maximal discrete spatial 2-volume. In figure 2.8 we display the fits of the function (2.26) to $\langle N_{2}^{\text{SL}}(\tau) \rangle$ for each of these transition amplitudes.

Although the function (2.26) fits $\langle N_{2}^{\text{SL}}(\tau) \rangle$ rather well, there is a disconcerting deviation in the fit close to the final boundary. The deviation grows with the discrete spatial
2-volume $N_2^{SL}(S_f^2)$ of the final boundary, which also corresponds to a growth in the temporal extent of the stalk. To explore the origin of this deviation, we consider the transition amplitudes for $N_2^{SL}(S_f^2) = 500$ for two successively smaller numbers $T$ of time slices keeping $\bar{N}_3 = 30850$ and $k_0 = 1.00$. In figure 2.9 we display $\langle N_2^{SL}(\tau) \rangle$ for $T = 29$ (as above), $T = 25$, and $T = 21$ fit to the function (2.26).

Clearly, as the temporal extent of the stalk diminishes, the deviation in the fit close to the final boundary also diminishes. Based on these findings, we suggest the following in-
Figure 2.9: Ensemble average number $\langle N_{2}^{SL} \rangle$ of spacelike 2-simplices as a function of the discrete time coordinate $\tau$ for $\bar{N}_3 = 30850$, $k_0 = 1.00$, and $N_{2}^{SL}(S_f^2) = 500$. (a) $T = 29$ (b) $T = 25$ (c) $T = 21$

interpretation. The deviation in the fit close to the final boundary indicates the beginning of the formation of another stalk, albeit composed of time slices with nonminimal spatial extent. This possibility is consistent with the behavior exhibited in figure 2.8, particularly for the case of $N_{2}^{SL}(S_f^2) = 700$ and in figure 2.9. Moreover, in the cases of periodic temporal boundary conditions and of minimal initial and final boundary conditions, when a stalk has formed, there are transitional regions at the junctions with the central accumulation of discrete spatial 2-volume where the fit of the function (2.26) deviates in a similar manner. See,
for instance, figures 2.4(a) and 2.5. We are currently investigating further this phenomenon and its proffered explanation.

As the plots of \( \langle N^{SL}_2(\tau) \rangle \) in figures 2.8 and 2.9 demonstrate, the partition function (2.20) for minimal initial boundary and nonminimal final boundary is dominated by the extremum of the action (2.17) corresponding to the portion of Euclidean de Sitter spacetime containing more than half of the discrete spacetime 3-volume of Euclidean de Sitter spacetime. Interestingly, Hartle and Hawking’s no-boundary proposal predicts just the opposite: that the portion of Euclidean de Sitter spacetime containing less than half of the spacetime 3-volume should dominate the path integral [113]. Halliwell and Louko pointed out that the proposal of Hartle and Hawking is not unique: there are other consistent choices of steepest descent integration contours, one of which yields the result that we have obtained [111]. We are currently attempting to determine to which analytic minisuperspace quantization the technique of causal dynamical triangulations corresponds. We hope that this investigation illuminates the relation of the causal dynamical triangulations quantization scheme to other proposals for the quantization of classical theories of gravity.

2.5.3 Nonminimal initial and final boundaries

We finally take both the initial boundary \( S^2_i \) and the final boundary \( S^2_f \) as nonminimal triangulations of the 2-sphere. In particular, for \( T = 29, \tilde{N}_3 = 30850 \), and \( k_0 = 1.00 \), we first consider transition amplitudes characterized by \( N^{SL}_2(S^2_i) = N^{SL}_2(S^2_f) = 4 \), \( N^{SL}_2(S^2_i) = N^{SL}_2(S^2_f) = 100 \), \( N^{SL}_2(S^2_i) = N^{SL}_2(S^2_f) = 500 \), \( N^{SL}_2(S^2_i) = N^{SL}_2(S^2_f) = 700 \),
and $N^{SL}_2(S^2_i) = N^{SL}_2(S^2_f) = 900$. (The first such transition amplitude is of course one of those considered previously.) In figure 2.10 we display $\langle N^{SL}_2(\tau) \rangle$ for each of these five transition amplitudes.

Also for $T = 29$, $\tilde{N}_3 = 30850$, and $k_0 = 1.00$, we consider the transition amplitudes characterized by \{\[N^{SL}_2(S^2_i) = 100, N^{SL}_2(S^2_f) = 300\], \[N^{SL}_2(S^2_i) = 300, N^{SL}_2(S^2_f) = 700\], and \[N^{SL}_2(S^2_i) = 500, N^{SL}_2(S^2_f) = 700\].\} In figure 2.11 we display $\langle N^{SL}_2(\tau) \rangle$ for each of these three transition amplitudes.

Clearly, for the transition amplitudes represented in figures 2.10 and 2.11, $\langle N^{SL}_2(\tau) \rangle$ only appears to conform to a portion of Euclidean de Sitter spacetime if the sum $N^{SL}_2(S^2_i) + N^{SL}_2(S^2_f)$ is sufficiently small. We confirm this quantitatively by fitting the function (2.26) to $\langle N^{SL}_2(\tau) \rangle$ for two such cases. In performing these fits, we again require the appropriate values for the ensemble average number $\langle N_{3}^{(1,3)} \rangle$ of $(1,3)$ 3-simplices. Instead of proceeding as in the previous subsection, we circumvent this issue by employing the variant of the function (2.26) appropriate to a portion of Euclidean de Sitter spacetime. We derive this variant in appendix 2.C; we could have elected to use this variant in the previous subsection as well. We display the results of these fits in figure 2.12. In the case of \{\[N^{SL}_2(S^2_i) = 100, N^{SL}_2(S^2_f) = 300\], there is clearly a deviation in the fit close to both the initial and final boundaries similar to what we found in the previous subsection.

If the sum $N^{SL}_2(S^2_i) + N^{SL}_2(S^2_f)$ is insufficiently small, then $\langle N^{SL}_2(\tau) \rangle$ exhibits a markedly different dynamics, possibly hyperbolic in nature. For the chosen values of $T$, $\tilde{N}_3$, and $k_0$,
Figure 2.10: Ensemble average number \( \langle N_{2}^{SL} \rangle \) of spacelike 2-simplices as a function of the discrete time coordinate \( \tau \) for \( T = 29, \bar{N}_{3} = 30850 \), and \( k_{0} = 1.00 \). (a) \( N_{2}^{SL}(S_{i}^{2}) = 4 \), \( N_{2}^{SL}(S_{f}^{2}) = 4 \) (b) \( N_{2}^{SL}(S_{i}^{2}) = 100 \), \( N_{2}^{SL}(S_{f}^{2}) = 100 \) (c) \( N_{2}^{SL}(S_{i}^{2}) = 500 \), \( N_{2}^{SL}(S_{f}^{2}) = 500 \) (d) \( N_{2}^{SL}(S_{i}^{2}) = 700 \), \( N_{2}^{SL}(S_{f}^{2}) = 700 \) (e) \( N_{2}^{SL}(S_{i}^{2}) = 900 \), \( N_{2}^{SL}(S_{f}^{2}) = 900 \)
there is a transition from the typical dynamics of phase C to the possibly hyperbolic dynamics when \( N_2^{SL}(S_i^2) + N_2^{SL}(S_f^2) \approx 600 \). We have determined the approximate location of this transition by additionally considering the transition amplitudes characterized by \( N_2^{SL}(S_i^2) = 300 \) and by \( \{ N_2^{SL}(S_i^2) = 100, N_2^{SL}(S_f^2) = 500 \} \), both of which show a mixture of these two dynamical behaviors.

Our above discussion of semiclassical expectations for these transition amplitudes suggests an interpretation of these results. The ensembles characterized by
Figure 2.12: Ensemble average number $\langle N_{2}^{SL} \rangle$ of spacelike 2-simplices as a function of the discrete time coordinate $\tau$ for $T = 29$, $N_{3} = 30850$, and $k_{0} = 1.00$. (a) $N_{2}^{SL}(S_{i}^{2}) = 100$, $N_{2}^{SL}(S_{f}^{2}) = 100$ (b) $N_{2}^{SL}(S_{i}^{2}) = 100$, $N_{2}^{SL}(S_{f}^{2}) = 300$

$N_{2}^{SL}(S_{i}^{2}) = N_{2}^{SL}(S_{f}^{2}) = 100$ and by $\{N_{2}^{SL}(S_{i}^{2}) = 100, N_{2}^{SL}(S_{f}^{2}) = 300\}$, which exhibit the typical dynamics of phase C, correspond to the regime in which both the initial and final scale factors are less than the de Sitter length for the given discrete spacetime 3-volume. This is the situation depicted in figure 2.3(d). Presumably then, the other five ensembles considered here, which exhibit possibly hyperbolic dynamics, correspond to the regime in which both the initial and final scale factors exceed the de Sitter length for the given discrete spacetime 3-volume. Recall that in this case the amplitude for the scale factor oscillates within an envelope consistent with the probability distribution of scale factors in Lorentzian de Sitter spacetime. Naively, this appears to be the result that we have achieved: $\langle N_{2}^{SL}(\tau) \rangle$ appears to conform to a portion of Lorentzian de Sitter spacetime. We are currently investigating whether or not this is quantitatively the case.

Supposing that a discretization of a portion of the spatial 2-volume as a function of the global time coordinate of Lorentzian de Sitter spacetime does indeed fit well to $\langle N_{2}^{SL}(\tau) \rangle$,
what conclusions should we draw? This finding would imply that the quantization scheme of causal dynamical triangulations functions similarly to that of the no-boundary proposal of Hartle and Hawking: one obtains transition amplitudes dominated by Lorentzian space-times via a Euclidean path integral [113]. Possibly then, as in the no-boundary proposal, Lorentzian not Euclidean de Sitter spacetime dominates the ground state of quantum spacetime geometry on sufficiently large scales. Regardless of the viability of these conclusions, the latter five transition amplitudes appear to represent numerical simulations of portions of temporally unbounded quantum spacetime geometry, a first within the causal dynamical triangulations approach.

2.6 Conclusion

We have initiated an investigation of transition amplitudes within the causal dynamical triangulations of (2 + 1)-dimensional Einstein gravity with positive cosmological constant. We have focused on transition amplitudes from a past spacelike hypersurface of fixed intrinsic geometry to a future spacelike hypersurface of fixed intrinsic geometry, these hypersurfaces being leaves of the foliation of causal triangulations. Our study thus required a generalization of previous numerical investigations of causal dynamical triangulations to the setting of fixed initial and final boundary geometries.

Specifically, we explored three classes of such transition amplitudes for spatial topology of a 2-sphere: minimal boundary–minimal boundary, minimal boundary–nonminimal
boundary, and nonminimal boundary–nonminimal boundary. When averaged over all geometrical degrees of freedom of each boundary except for the discrete spatial 2-volume, these transition amplitudes are all evidently compatible with the gravitational effective action (2.22) previously demonstrated to describe the ground state of phase C. We are currently investigating the extent to which this compatibility holds quantitatively at the level of deviations from the ensemble average.

Each class of transition amplitude has afforded new insights into the approach of causal dynamical triangulations. The minimal boundary–minimal boundary transition amplitudes, as well as the minimal boundary–nonminimal boundary transition amplitudes, definitively demonstrate that the stalk regions of quantum spacetime geometry, previously observed in all numerical simulations within phase C, are indeed artifacts of the numerical implementation. Nevertheless, we are interested in understanding the formation of stalks, particularly as they appear in the minimal boundary–nonminimal boundary transition amplitudes. To this end we are exploring whether the minisuperspace model defined by the effective action (2.22) can explain the dynamical formation of stalks.

The minimal boundary–nonminimal boundary transition amplitudes—and possibly the nonminimal boundary–nonminimal boundary transition amplitudes—provide for direct comparisons of the quantizations of causal dynamical triangulations and of certain continuum approaches. In particular, these transition amplitudes indicate that the technique of causal dynamical triangulations does not correspond precisely to that of Hartle and Hawking’s no-boundary proposal. We are currently attempting to ascertain whether the quantization
scheme of causal dynamical triangulations corresponds to one of the variants of this proposal put forward by Halliwell and Louko [111]. We also wish to determine whether the nonminimal boundary–nonminimal boundary transition amplitudes—apparently representing numerical simulations of portions of temporally unbounded quantum spacetime geometry—match the quantitative expectations of these scenarios.

Our study of transition amplitudes opens the door to a multitude of interesting new explorations of causal dynamical triangulations. As discussed above, we have only studied geometry-averaged transition amplitudes, that is, those depending solely on the discrete spatial $2$-volumes of the bounding geometries. Presumably, there is considerably more information within transition amplitudes that probe the full dependence on the boundaries’ intrinsic geometries. To study such transition amplitudes, one would require the ability first to characterize the geometrical degrees of freedom of triangulated boundary geometries and then to design triangulated boundary geometries with chosen characteristics. Some of the techniques employed by Sachs, in combination with our Markov chain Monte Carlo algorithm for generating randomly triangulated $2$-spheres, may serve this purpose [190]. We are particularly interested in the possibility of observing effects that are not describable within a minisuperspace truncation of the metric degrees of freedom. Specifically, by designing appropriate boundary $2$-spheres, we hope to investigate the absence or presence of propagating degrees of freedom in the quantum theory.

Our techniques might also prove useful in further studying the causal dynamical triangulations of projectable Hořava-Lifshitz gravity as initiated by Anderson et al [31]. These
authors discovered an auxiliary phase of quantum spacetime geometry—their so-called phase E—but they did not have a method for coherently ensemble averaging the discrete spatial 2-volume as a function of the discrete time coordinate in this phase. By fixing the intrinsic geometries of the initial and final boundaries, one could employ the boundaries to temporally align each causal triangulation in an ensemble, allowing for coherent averaging in phase E. Furthermore, following on the final aspiration of the previous paragraph, one could attempt to determine whether or not there is a propagating scalar mode within this quantization of Hořava-Lifshitz gravity. Such investigations would work towards ascertaining the relationship between the causal dynamical triangulations of Einstein gravity and quantum Hořava-Lifshitz gravity [12], which have recently been shown to be equivalent in $1 + 1$ dimensions [11].

Acknowledgments

We wish to thank Steve Carlip for his guidance and input throughout the course of this project. We are grateful to Christian Anderson, David Kamensky, and especially Rajesh Kommu for allowing us to employ parts of their computer codes. JHC also wishes to thank Kyle Lee for several useful conversations. JHC acknowledges support from the Department of Energy under grant DE-FG02-91ER40674. JMM acknowledges support from the National Science Foundation under REU grant PHY-1004848 at the University of California, Davis. This work utilized the Janus supercomputer, which is supported by the National
Science Foundation (award number CNS-0821794) and the University of Colorado, Boulder. The Janus supercomputer is a joint effort of the University of Colorado, Boulder, the University of Colorado, Denver, and the National Center for Atmospheric Research.
Appendix

2.A On the consistency of the action $S_R[\mathcal{T}_c]$ under composition

We demonstrate that our prescription (2.11) for the Gibbons-Hawking-York boundary term is consistent with the prescription (2.10) for the Einstein-Hilbert action within the causal dynamical triangulations of $(2 + 1)$-dimensional Einstein gravity with positive cosmological constant for 2-sphere spatial topology. We make such a demonstration by verifying that the prescription (2.11) reproduces the prescription (2.10) under the composition of two spacetime regions sharing a common boundary. Hartle and Sorkin employed this criterion in deriving the form (2.8) of the Gibbons-Hawking-York boundary term in Regge calculus [114], so we certainly expect our prescription (2.11) to preserve this criterion.

Consider two causal triangulations $\mathcal{T}_c$ and $\mathcal{T}_c'$ both with 2-sphere spatial topology and line interval temporal topology. The boundary $\partial \mathcal{T}_c$ consists of an initial 2-sphere $S^2_i$ and a final 2-sphere $S^2_f$, and the boundary $\partial \mathcal{T}_c'$ consists of an initial 2-sphere $S'_i$ and a final
2-sphere $S'^2$. To compose the two causal triangulations $T_c$ and $T'_c$, we take the 2-spheres $S^2_f$ and $S'^2_i$ to have the same intrinsic geometry, and we orient these 2-spheres $S^2_f$ and $S'^2_i$ to have coincident normal vectors. We may thus identify these two 2-spheres as $S^2_C$, the common boundary. The Gibbons-Hawking-York boundary term for $S^2_C$ from the two causal triangulations $T_c$ and $T'_c$ is

$$
\frac{a}{8\pi G} \left[ \pi i N^1_{i}(S^2_f) - \frac{2}{i} \theta_{SL}^{(1,3)} N^S_{i}(S^2_f) - \frac{1}{i} \theta_{SL}^{(2,2)} N^2_{i}(S^2_f) \right]
+ \frac{a}{8\pi G} \left[ \pi i N^1_{i}(S'^2_i) - \frac{2}{i} \theta_{SL}^{(3,1)} N^S_{i}(S'^2_i) - \frac{1}{i} \theta_{SL}^{(2,2)} N^2_{i}(S'^2_i) \right].
\tag{2.28}
$$

These two contributions to the Gibbons-Hawking-York boundary term combine to give

$$
\frac{a}{8\pi G} \left[ \frac{2\pi}{i} N^1_{i}(S^2_C) - \frac{4}{i} \theta_{SL}^{(1,3)} N^S_{i}(S^2_C) - \frac{2}{i} \theta_{SL}^{(2,2)} N^2_{i}(S^2_C) \right],
\tag{2.29}
$$

precisely the contribution to the Einstein-Hilbert action stemming from the spacelike hinges on the common boundary $S^2_C$.

2.B On the algorithm for inserting spatial 2-sphere boundaries

We describe the algorithm for incorporating an arbitrary initial or final spacelike boundary $S^2_B$ of 2-sphere topology into a minimal causal triangulation. Recall that a minimal causal triangulation consists of minimally triangulated 2-spheres, each the surface of a spa-
tial tetrahedron, for every time slice with adjacent time slices connected by the minimal number of timelike edges. This construction results in there being fourteen 3-simplices between adjacent time slices: four \((3, 1)\) 3-simplices connecting spacelike 2-simplices in the initial slice to vertices in the final slice, four \((1, 3)\) 3-simplices connecting spacelike 2-simplices in the final slice to vertices in the initial slice, and six \((2, 2)\) 3-simplices filling in the gaps between these other 3-simplices to ensure the correct topology.

To insert the arbitrarily triangulated 2-sphere \(S^2_B\) into this triangulation, we first decompose \(S^2_B\) into four pseudofaces and six pseudoedges. Each pseudoface is a simply-connected set of spacelike 2-simplices, and each pseudoedge is a piecewise curve of spacelike 1-simplices. See figure 2.13. Clearly, each pseudoface is bounded by three pseudoedges. Together the pseudofaces and pseudoedges form a pseudotetrahedron. The decomposition into pseudofaces and pseudoedges is not unique except for the minimally triangulated 2-sphere. This nonuniqueness does not impact our algorithm, which works for any such decomposition.

We next delete either the initial or final time slice from the minimal causal triangulation. We wish to replace this time slice with the arbitrarily triangulated 2-sphere \(S^2_B\) that we have decomposed into a pseudotetrahedron. To make this replacement, we require a method for correctly connecting \(S^2_B\) to the next-to-initial or next-to-final time slice of the minimal causal triangulation. We achieve this connection by replacing the fourteen 3-simplices formerly connecting the initial and next-to-initial or final and next-to-final time slices with fourteen pseudo-3-simplices. In particular, we employ pseudo-\((1, 3)\) 3-simplices, pseudo-
Figure 2.13: An example decomposition of a triangulated spatial 2-sphere (not completely depicted) into four pseudofaces—labelled 1, 2, 3, and 4—and the six corresponding pseudoedges.

(2, 2) 3-simplices, and pseudo-(3, 1) 3-simplices in place of the (1, 3) 3-simplices, (2, 2) 3-simplices, and (3, 1) 3-simplices. A pseudo-\((p, q)\) 3-simplex is a complex of \((p, q)\) 3-simplices having the topology of a single \((p, q)\) 3-simplex and constructed to have the spacelike pseudoedges matching those of the pseudofaces to which it will connect. This algorithm results in a minimal causal triangulation having either an initial or final arbitrarily triangulated 2-sphere boundary \(S^2_0\).
Figure 2.14: The three types of pseudo-3-simplices employed in (2+1)-dimensional causal dynamical triangulations with fixed 2-sphere boundaries: (a) pseudo-(3, 1) 3-simplex, (b) pseudo-(2, 2) 3-simplex, (c) pseudo-(3, 1) 3-simplex. The jagged curves are spacelike pseudoedges, and the straight lines are timelike edges.

2.C A derivation of the function $N_2^{SL}(\tau)$

We demonstrate that the function

$$N_2^{SL}(\tau) = \frac{2}{\pi} \frac{\langle N_3^{(1,3)} \rangle}{s_0 (N_3^{(1,3)})^{1/3}} \cos^2 \left( \frac{\tau}{s_0 (N_3^{(1,3)})^{1/3}} \right),$$

previously given in equation (2.26), is the discrete analogue of the spatial 2-volume $V_2$ as a function of the global time coordinate $\eta$ of Euclidean de Sitter spacetime. This demonstration relies crucially on application of the appropriate finite size scaling towards the continuum limit. Dimensional analysis suggests that one approaches the continuum limit by taking the lattice spacing $a$ to zero and letting the number $N_3$ of 3-simplices increase without bound while holding the product $a^3 N_3$ constant. That the gravitational effective action (2.22) describes the ensemble average spacetime geometry on sufficiently large scales
supports the conclusion of this naive analysis [15, 14, 23, 24, 26, 25, 45]. In particular, we expect the condition

$$V_3 = C_3 a^3 N_3,$$  
(2.31)

relating the continuum spacetime 3-volume $V_3$ to the lattice spacing $a$ and the number $N_3$ of 3-simplices to hold in the continuum limit. Here, $C_3$ is the effective discrete spacetime 3-volume of a 3-simplex. Accordingly, we expect discrete quantities having the associated dimensions $a^p$ to finite size scale towards the continuum as $N_3^{-p/3}$. Specifically, we expect the discrete time coordinate $\tau$ to scale as $\tau N_3^{-1/3}$, matching onto the dimensionless continuum time coordinate $\eta V_3^{-1/3}$, and we expect the number $N_2^{SL}$ of spacelike 2-simplices to scale as $N_2^{SL} N_3^{-2/3}$, matching onto the dimensionless continuum spatial 2-volume $V_2 V_3^{-2/3}$.

With these considerations we are now prepared to derive the function (2.30). We start from the definition

$$V_3 = \int_{-\pi l_{ds}/2\sqrt{g_{\eta\eta}}}^{+\pi l_{ds}/2\sqrt{g_{\eta\eta}}} d\eta \sqrt{g_{\eta\eta}} V_2(\eta),$$  
(2.32)

which determines the spacetime 3-volume $V_3$ of Euclidean de Sitter spacetime, and the definition

$$N_3 = N_3^{(1,3)} + N_3^{(2,2)} + N_3^{(3,1)},$$  
(2.33)

which determines the number $N_3$ of 3-simplices comprising a causal triangulation $\mathcal{T}_c$. For spacetime topology $S^2 \times S^1$ one may recast the definition (2.33) as

$$N_3 = 2(1 + \xi) \sum_{\tau=1}^{T} N_2^{SL}(\tau),$$  
(2.34)
where
\[ \xi = \frac{N_3^{(2,2)}}{N_3^{(1,3)} + N_3^{(3,1)}}. \] (2.35)

The condition (2.31) then becomes
\[ \int d\eta \sqrt{g_{\eta\eta}} V_2(\eta) = 2C_3a^3(1 + \xi) \sum_{\tau=1}^{T} N_{2}^{SL}(\tau), \] (2.36)

yielding the identification
\[ d\eta \sqrt{g_{\eta\eta}} V_2(\eta) = 2C_3a^3(1 + \xi)N_{2}^{SL}(\tau). \] (2.37)

Given the scale factor (2.25) describing Euclidean de Sitter spacetime, the spatial 2-volume as a function of the global time coordinate is
\[ V_2(\eta) = \frac{2V_3}{\pi l_{dS}} \cos^2 \left( \frac{\sqrt{g_{\eta\eta}}}{l_{dS}} \eta \right) \] (2.38)
in terms of the spacetime 3-volume \( V_3 = 2\pi^2l_{dS}^3 \). Solving equation (2.37) for \( N_{2}^{SL}(\tau) \) and substituting for \( V_2(\eta) \) from the expression (2.38), we obtain the relation
\[ N_{2}^{SL}(\tau) = \frac{d\eta \sqrt{g_{\eta\eta}}}{2a^2C_3(1 + \xi)} \frac{2V_3}{\pi l_{dS}} \cos^2 \left( \frac{\sqrt{g_{\eta\eta}}}{l_{dS}} \eta \right). \] (2.39)
Using the condition (2.31) again, we may rewrite equation (2.39) as

\[ N_2^{SL}(\tau) = \frac{d\eta\sqrt{g_{\eta\eta}}}{(1 + \xi)} \frac{N_3}{\pi l_{dS}^2} \cos^2 \left( \frac{\sqrt{g_{\eta\eta}}}{l_{dS}} \eta \right). \]  

(2.40)

Replacing \( \eta \) with \( \tau \) according to the above scaling correspondence, equation (2.40) becomes

\[ N_2^{SL}(\tau) = \frac{1}{\pi} \frac{\Delta \tau \sqrt{g_{\eta\eta}} V_3^{1/3} N_3}{l_{dS} N_3^{1/3} (1 + \xi)} \cos^2 \left( \frac{\sqrt{g_{\eta\eta}} V_3^{1/3} \tau}{l_{dS} N_3^{1/3}} \right). \]  

(2.41)

Defining

\[ \frac{1}{s_0} = \frac{V_3^{1/3} \sqrt{g_{\eta\eta}}}{l_{dS}}, \]  

(2.42)

we finally determine that

\[ N_2^{SL}(\tau) = \frac{1}{\pi} \frac{N_3}{s_0 (1 + \xi) N_3^{1/3}} \cos^2 \left( \frac{\tau}{s_0 N_3^{1/3}} \right) \]  

(2.43)

since \( \Delta \tau = 1 \). In terms of \( N_3^{(3,1)} \) and the modified parameter \( \tilde{s}_0 = 2^{1/3} s_0 (1 + \xi)^{1/3} \),

\[ N_2^{SL}(\tau) = \frac{2}{\pi} \frac{N_3^{(3,1)}}{\tilde{s}_0 \left( N_3^{(3,1)} \right)^{1/3}} \cos^2 \left[ \frac{\tau}{\tilde{s}_0 \left( N_3^{(3,1)} \right)^{1/3}} \right]. \]  

(2.44)

The parameter \( \tilde{s}_0 \) is effectively the dimensionless de Sitter radius. As we only expect the relation (2.44) to hold at the level of the ensemble average, we obtain the function (2.30).

We readily generalize the result (2.44) to a portion of Euclidean de Sitter spacetime.
We now start from the definition

\[ V_3 = \int_{\eta_i}^{\eta_f} d\eta \sqrt{|g_{\eta\eta}|} V_2(\eta), \]  

which determines the spacetime 3-volume \( V_3 \) of the portion of Euclidean de Sitter spacetime between global times \( \eta_i \) and \( \eta_f \) as

\[ V_3 = 2\pi l^2 \sqrt{|g_{\eta\eta}|} \left\{ \eta_f - \eta_i + \frac{l_{dS}}{\sqrt{g_{\eta\eta}}} \sin \left[ \frac{\sqrt{g_{\eta\eta}(\eta_f - \eta_i)}}{l_{dS}} \right] \cos \left[ \frac{\sqrt{g_{\eta\eta}(\eta_f + \eta_i)}}{l_{dS}} \right] \right\}. \]  

(2.46)

Proceeding precisely as above, we eventually find that

\[ N_{SL}^{2}(\tau) = \frac{2N_3^{(1,3)} \cos^2 \left[ \frac{\tau}{\bar{s}_0 \left( N_3^{(1,3)} \right)^{1/3}} \right]}{\tau_f - \tau_i + \bar{s}_0 \left( N_3^{(1,3)} \right)^{1/3} \sin \left[ \frac{\tau_f - \tau_i}{\bar{s}_0 \left( N_3^{(1,3)} \right)^{1/3}} \right] \cos \left[ \frac{\tau_f + \tau_i}{\bar{s}_0 \left( N_3^{(1,3)} \right)^{1/3}} \right]} \]  

(2.47)

for the values \( \tau_i \) and \( \tau_f \) of the discrete time coordinate corresponding to \( \eta_i \) and \( \eta_f \).

We fit the function (2.30) to the coherent ensemble average number \( \langle N_{SL}^{2}(\tau) \rangle \) of spacelike 2-simplices as a function of the discrete time coordinate, measured directly from Markov chain Monte Carlo data, as follows. In the case of periodic boundary conditions and for minimal boundary–minimal boundary transition amplitudes when stalks are
present, we employ in particular the function

\[ N_{2}^{SL}(\tau) = \begin{cases} 
A & \text{for } -\frac{T}{2} \leq \tau < -\tau_m \\
\frac{2}{\pi} \frac{\langle N_2^{(1,3)} \rangle}{\bar{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \cos^2 \left( \frac{\tau}{\bar{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) & \text{for } -\tau_m \leq \tau \leq \tau_m \\
A & \text{for } \tau_m < \tau \leq \frac{T}{2} 
\end{cases} \]  

(2.48)

for the value

\[ \tau_m = \bar{s}_0 \langle N_3^{(1,3)} \rangle^{1/3} \cos^{-1} \sqrt{\frac{\pi A \bar{s}_0}{2 \langle N_3^{(1,3)} \rangle^{2/3}}} \]  

(2.49)

of the discrete time coordinate that matches the discrete spatial 2-volume \( A \) of the stalk to that of the central accumulation. For minimal boundary–minimal boundary transition amplitudes when stalks are not present, we employ the function (2.48) with \( \tau_m = \frac{T}{2} \), and we enforce that the function (2.48) passes through the boundary values of \( \langle N_2^{SL}(\tau) \rangle \). The value of \( \bar{s}_0 \) is thus determined to be the root of the transcendental equation

\[ 4 = \frac{2}{\pi} \frac{\langle N_2^{(1,3)} \rangle}{\bar{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \cos^2 \left( \frac{\frac{T}{2}}{\bar{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) \]  

(2.50)

since the boundary value of \( \langle N_2^{SL}(\tau) \rangle \) is precisely 4 in this case. We determine whether or not stalks are present by considering both of the above fits, taking the goodness of the fit, described below, as our indicator.

For minimal boundary–nonminimal boundary transition amplitudes when stalks are
present, we employ in particular the function

\[
\mathcal{N}_2^{SL}(\tau) = \begin{cases} 
A & \text{for } -\frac{T}{2} \leq \tau < \tau_m' \\
\frac{2}{\pi} \frac{\langle N_{1,3}^{(1,3)} \rangle_{\text{eff}}}{\bar{s}_0 \langle N_{3}^{(1,3)} \rangle_{\text{eff}}^{1/3}} \cos^2 \left( \frac{\pi - \tau}{\bar{s}_0 \langle N_{3}^{(1,3)} \rangle_{\text{eff}}^{1/3}} \right) & \text{for } \tau_m' \leq \tau \leq \tau_m'' \\
B & \text{for } \tau_m'' < \tau \leq \frac{T}{2}
\end{cases}
\]  

(2.51)

for the values

\[
\tau_m' = \tau_s - \bar{s}_0 \langle N_{3}^{(1,3)} \rangle_{\text{eff}}^{1/3} \cos^{-1} \sqrt{\frac{\pi A \bar{s}_0}{2 \langle N_{1,3}^{(1,3)} \rangle_{\text{eff}}^{2/3}}} 
\]  

(2.52a)

\[
\tau_m'' = \tau_s + \bar{s}_0 \langle N_{3}^{(1,3)} \rangle_{\text{eff}}^{1/3} \cos^{-1} \sqrt{\frac{\pi B \bar{s}_0}{2 \langle N_{1,3}^{(1,3)} \rangle_{\text{eff}}^{2/3}}} 
\]  

(2.52b)

of the discrete time coordinate that match the discrete spatial 2-volumes \( A \) and \( B \) of the stalks to those of the central accumulation. For minimal boundary–nonminimal boundary transition amplitudes when stalks are not present, we employ the function (2.51) with \( \tau_m' = -\frac{T}{2} \) and \( \tau_m'' = \frac{T}{2} \), and we enforce that the function (2.51) passes through the boundary values of \( \langle N_2^{SL}(\tau) \rangle \). The values of \( \bar{s}_0 \) and \( \tau_s \) are thus determined to be the roots of the transcendental equations

\[
4 = \frac{2}{\pi} \frac{\langle N_{3}^{(1,3)} \rangle_{\text{eff}}}{\bar{s}_0 \langle N_{3}^{(1,3)} \rangle_{\text{eff}}^{1/3}} \cos^2 \left( \frac{T}{2} + \tau_s \right) \]  

(2.53a)

\[
X = \frac{2}{\pi} \frac{\langle N_{3}^{(1,3)} \rangle_{\text{eff}}}{\bar{s}_0 \langle N_{3}^{(1,3)} \rangle_{\text{eff}}^{1/3}} \cos^2 \left( \frac{T}{2} - \tau_s \right) \]  

(2.53b)

since \( \langle N_2^{SL}(S^2) \rangle = 4 \) and \( \langle N_2^{SL}(S^2_f) \rangle = X \) in this case.
For nonminimal boundary–nonminimal boundary transition amplitudes when stalks are present, we employ in particular the function

\[ N_{2}^{SL}(\tau) = \begin{cases} 
A & \text{for } -\frac{T}{2} \leq \tau < \tau_{m}' \\
2(N_{3}^{(1,3)})\cos^{2}\left(\frac{\tau-\tau_{s}}{\tilde{s}_{0}(N_{3}^{(1,3)})^{1/3}}\right) & \text{for } \tau_{m}' \leq \tau \leq \tau_{m}'' \\
B & \text{for } \tau_{m}'' < \tau \leq \frac{T}{2}
\end{cases} \]  

(2.54)

for the values

\[ \tau_{m}' = \tau_{s} - \tilde{s}_{0}(N_{3}^{(1,3)})^{1/3} \]  

(2.55a)

\[ \tau_{m}'' = \tau_{s} + \tilde{s}_{0}(N_{3}^{(1,3)})^{1/3} \]  

(2.55b)

of the discrete time coordinate that match the discrete spatial 2-volumes \( A \) and \( B \) of the stalks to those of the central accumulation. For nonminimal boundary–nonminimal boundary transition amplitudes when stalks are not present and \( N_{2}^{SL}(S_{i}^{2}) = N_{2}^{SL}(S_{f}^{2}) \), we employ the function (2.47) with \( \tau_{m}' = -\frac{T}{2} \) and \( \tau_{m}'' = \frac{T}{2} \), and we enforce that the function (2.54) passes through the boundary values of \( \langle N_{2}^{SL}(\tau) \rangle \). The value of \( \tilde{s}_{0} \) is thus determined to be
the root of the transcendental equation

\[ X = \frac{2\langle N^{(1,3)}_3 \rangle \cos^2 \left( \frac{\tau_f - \tau_i}{s_0 \langle N^{(1,3)}_3 \rangle^{1/3}} \right) \sin \left( \frac{\tau_f - \tau_i}{s_0 \langle N^{(1,3)}_3 \rangle^{1/3}} \right) \cos \left( \frac{\tau_f + \tau_i}{s_0 \langle N^{(1,3)}_3 \rangle^{1/3}} \right) \tau_f - \tau_i + s_0 \langle N^{(1,3)}_3 \rangle^{1/3} \} \]  

(2.56)

since \( N^{SL}_2(S^2_i) = N^{SL}_2(S^2_f) = X \) in this case.

To determine the best fit of the function (2.30) to \( \langle N^{SL}_2(\tau) \rangle \), we minimize the chi-squared function

\[ \chi^2(s_0, \ldots) = \sum_{\tau=1}^{T} \frac{\left[ \langle N^{SL}_2(\tau) \rangle - \mathcal{N}^{SL}_2(\tau) \right]^2}{\sigma^2(\langle N^{SL}_2(\tau) \rangle)} \]  

(2.57)

with the ellipsis indicating the potential inclusion of the fit parameters \( A, B, \) and \( \tau_s \). Here,

\[ \sigma^2(\langle N^{SL}_2(\tau) \rangle) = \frac{\langle (N^{SL}_2(\tau))^2 \rangle - \langle N^{SL}_2(\tau) \rangle^2}{N(T_c)} \]  

(2.58)

is the sample variance in \( \langle N^{SL}_2(\tau) \rangle \). We determine the error \( \epsilon(s_0) \) in the fit parameter \( s_0 \) by solving for the two values of \( s_0 \) satisfying

\[ \chi^2(s_0, \ldots) = \chi^2_{\text{min}}(s_0, \ldots) + 1, \]  

(2.59)

one less than and one greater than the value of \( s_0 \) at the minimum of the chi-squared function (2.57). We set the fit parameters \( A, B, \) and \( \tau_s \) to their values at the minimum of the chi-squared function (2.57) in making this determination. Our measure of the goodness of
fit of the function (2.30) to $\langle N^S_L(\tau) \rangle$ is the minimum chi-squared per degree of freedom

$$
\chi^2_{\text{pdf}} = \frac{\chi^2_{\text{min}}(\tilde{s}_0,\ldots)}{T - n},
$$

(2.60)

where $n$ is the number of fit parameters.
Chapter 3

A Second Look at Transition Amplitudes in (2+1)-Dimensional Causal Dynamical Triangulations

3.1 Context

This work (first presented in [77]) is a collaboration between myself, Joshua Cooperman, and Kyle Lee. The bulk of the work was done between September 2014 and April 2015. The manuscript was submitted for peer review October 2016 and published May 2017. J. Cooperman was the corresponding author. The results presented here are an extension of those presented in chapter 2, and thus I developed the algorithm and implementation used. I also ran all simulations. J. Cooperman performed the data analysis and
generated the plots. All three authors drew the main scientific conclusions and constructed the main argument in section 3.7 together in discussions via Skype and all three authors were heavily involved in the writing of the manuscript. As of yet, this paper has no citations; however it is an important part of the story presented in chapter 2.

3.2 Abstract

Studying transition amplitudes in \((2 + 1)\)-dimensional causal dynamical triangulations, Cooperman and Miller discovered speculative evidence for Lorentzian quantum geometries emerging from its Euclidean path integral [78]. On the basis of this evidence, Cooperman and Miller conjectured that Lorentzian de Sitter spacetime, not Euclidean de Sitter space, dominates the ground state of the quantum geometry of causal dynamical triangulations on large scales, a scenario akin to that of the Hartle-Hawking no-boundary proposal in which Lorentzian spacetimes dominate a Euclidean path integral [113]. We argue against this conjecture: we propose a more straightforward explanation of their findings, and we proffer evidence for the Euclidean nature of these seemingly Lorentzian quantum geometries. This explanation reveals another manner in which the Euclidean path integral of causal dynamical triangulations behaves correctly in its semiclassical limit—the implementation and interaction of multiple constraints. This work was performed in collaboration with Joshua Cooper-erman and Kyle Lee.

3.3 Euclidean from Lorentzian

One often studies a Poincaré-invariant quantum field theory defined on Minkowski spacetime via Wick rotation to a Euclidean-invariant statistical field theory defined on Euclidean space. Within the path integral formulation, the Wick rotation transforms a Lorentzian path integral, which involves complex probability amplitudes for each Lorentzian field configuration, into a statistical partition function, which involves real probabilities
for each Euclidean field configuration. Absent the complications of complex probability amplitudes for Lorentzian field configurations, calculations typically prove considerably more tractable. Provided that the statistical field theory satisfies the Osterwalder-Schrader axioms, one can recover the Lorentzian theory from the Euclidean theory through the Osterwalder-Schrader reconstruction theorem [164, 165]. One thus defines the Lorentzian theory in terms of the Euclidean theory.

The tempting prospect that a quantum theory of gravity could be similarly defined led to the development of various approaches taking as their starting point the partition function

\[ \mathcal{Z}[\gamma] = \int_{\mathcal{M}} d\mu(g) e^{-S_{cl}(g) / \hbar} \]  

(3.1)

over Euclidean geometries specified by a metric tensor \( g \).

One should, however, be skeptical of these approaches’ applicability to gravity: a typical spacetime, even satisfying the Einstein equations, does not permit a global Wick rotation from Lorentzian to Euclidean signature. Nevertheless, such approaches—collectively called Euclidean quantum gravity—work not only sensibly, but even successfully in sundry circumstances [97]. We briefly mention two notable examples. First, one can derive the thermodynamic behavior of black holes from the partition function (3.1). Gibbons and Hawking computed the black hole entropy [98], and Hartle and Hawking computed the black hole radiance [112]. Second, Hartle and Hawking developed a quantum theory of gravity in the minisuperspace truncation from the partition function (3.1), their so-called
no-boundary proposal [113]. These authors defined a wave function for the universe having a remarkable property: Lorentzian geometries dominate the partition function (3.1) owing to the necessity of deforming an integration contour into the complex plane. Consequently, there is no need for an Osterwalder-Schrader reconstruction: the partition function (3.1) directly defines a Lorentzian quantum theory of gravity. Initial attempts to construct a complete nonperturbative quantum theory of gravity on the basis of the partition function (3.1) did not fare so well [153]. Two approaches, quantum Regge calculus and Euclidean dynamical triangulations, both grounded upon lattice regularization of the partition function (3.1), were extensively studied [153]. Neither of the quantum theories of gravity so defined exhibited a sufficiently rich phase structure to support a continuum limit. More recently, an approach based on exact renormalization group analysis of the partition function (3.1) has shown promise [186].

Causal dynamical triangulations emerged from the failures of quantum Regge calculus and Euclidean dynamical triangulations [30]. This newer approach takes as its starting point the Lorentzian path integral

$$\mathcal{A}[\gamma] = \int_{\mathcal{M}} \mu(g) e^{iS_{cl}[g]/\hbar}$$

(3.2)

over Lorentzian geometries specified by a metric tensor $g$. One chooses to restrict the path integration to appropriately causal Lorentzian geometries, namely, those admitting a global foliation by spacelike hypersurfaces all of fixed topology. One then introduces

---

1See [148], however.
a lattice regularization—causal triangulations—of these causal Lorentzian geometries. As Ambjørn, Jurkiewicz, and Loll demonstrated, this restriction allows for a well-defined Wick rotation of any Lorentzian causal triangulation to a corresponding Euclidean causal triangulation [21, 22]. This Wick rotation enables the use of Monte Carlo methods to study the resulting partition function.

Having implemented this Wick rotation, one could have wondered if the resulting partition function behaves conventionally, such as that of a field theory satisfying the Osterwalder-Schrader axioms, or unconventionally, such as that of the Hartle-Hawking no-boundary proposal. On the basis of Monte Carlo simulations of certain transition amplitudes within the causal dynamical triangulations of (2 + 1)-dimensional Einstein gravity, Cooperman and Miller conjectured that its partition function behaves unconventionally [78]. Specifically, these authors suggested that geometries resembling Lorentzian de Sitter spacetime—not, as previously thought, Euclidean de Sitter space—on sufficiently large scales dominate this partition function. Independently, Ambjørn et al argued for a signature change transition within the causal dynamical triangulations of (3+1)-dimensional Einstein gravity [8]. We now argue, contrary to the conjecture of Cooperman and Miller, that the partition function of causal dynamical triangulations behaves conventionally. Specifically, by reinterpreting these Monte Carlo simulations, we maintain that geometries resembling Euclidean de Sitter space on sufficiently large scales indeed dominate this partition function. In the process of making this argument, we provide further evidence that the partition function of causal dynamical triangulations behaves correctly in its semiclassical limit.
We introduce the formalism of causal dynamical triangulations, specializing to the case of $2 + 1$ dimensions for spherical topology with initial and final spacelike boundaries, in section 3.4. After recalling the relevant results from [78] and presenting new related results, we restate the conjecture of Cooperman and Miller in section 3.5. We present a first analysis of all of these results in section 3.6, which offers evidence in support of their conjecture. We present a more careful analysis in section 3.7, which leads to our argument refuting the conjecture of Cooperman and Miller. We conclude in section 3.8 by echoing Cooperman’s call for the proof of an Osterwalder-Schrader-type theorem for causal dynamical triangulations [75]. Three appendices supplement aspects of sections 3.6 and 3.7.

3.4 Causal dynamical triangulations

Within a path integral quantization of a classical metric theory of gravity, one formally defines a transition amplitude as

$$A[\gamma] = \frac{\int_{g|_{\partial M}=\gamma} d\mu(g) e^{iS_{cl}[g]}/\hbar}{\int_{\gamma} d\mu(g) e^{iS_{cl}[g]}/\hbar}.$$  (3.3)

The right hand side of equation (3.3) encodes the following instructions for computing the transition amplitude $A[\gamma]$: integrate over all spacetime metric tensors $g$ that induce the metric tensor $\gamma$ on the boundary $\partial M$ of the spacetime manifold $M$, weighting each metric tensor $g$ by the product of the measure $d\mu(g)$ and the exponential $e^{iS_{cl}[g]/\hbar}$. $S_{cl}[g]$ is the
action specifying the classical metric theory of gravity, including boundary terms enforcing
the condition $g|_{\partial M} = \gamma$.

Within the causal dynamical triangulations approach to such a quantization,\(^2\) one restricts the path integration in equation (3.3) to so-called causal spacetime metric tensors $g_c$, those admitting a global foliation by spacelike hypersurfaces all of a fixed spatial topology $\Sigma$. The manifold $M$ therefore has the topology $\Sigma \times I$, the direct product of $\Sigma$ and a real interval $I$. By invoking this restriction, one considers transition amplitudes $\mathcal{A}_\Sigma[\gamma]$ formally defined as

$$\mathcal{A}_\Sigma[\gamma] = \int_{M \cong \Sigma \times I, g_c|_{\partial M} = \gamma} d\mu(g_c) e^{iS_{cl}[g_c]/\hbar}. \quad (3.4)$$

To regularize the transition amplitudes $\mathcal{A}_\Sigma[\gamma]$, one replaces the path integration over all causal metric tensors $g_c$ in equation (3.4) with a path summation over all causal triangulations $T_c$. A causal triangulation $T_c$ is a piecewise-Minkowski simplicial manifold possessing a global foliation by spacelike hypersurfaces all of the topology $\Sigma$. One constructs a causal triangulation $T_c$ by appropriately gluing together $N_D$ causal $D$-simplices, each a simplicial piece of $D$-dimensional Minkowski spacetime with spacelike edges of squared invariant length $a^2$ and timelike edges of squared invariant length $-\alpha a^2$ for positive constant $\alpha$. $a$ is the lattice spacing. We depict the three types of causal 3-simplices (tetrahedra) in figure 3.1. Causal $D$-simplices necessarily assemble into a manifold of topology $\Sigma \times I$, and their skeleton distinguishes a foliation of this manifold into spacelike hypersurfaces.

\(^2\)See [21, 22, 30] for the original formulation and [17] for a comprehensive review.
Figure 3.1: Causal 3-simplices employed in $(2 + 1)$-dimensional causal dynamical triangulations extending from time slice $\tau = 0$ to time slice $\tau = 1$. (a) $(3, 1)$ 3-simplex, (b) $(2, 2)$ 3-simplex, (c) $(1, 3)$ 3-simplex. We have adapted this figure from [78].

We refer to the leaves of this distinguished foliation as a causal triangulation’s time slices, and we enumerate a causal triangulation’s $T$ time slices with a discrete time coordinate $\tau$.

By invoking this regularization, one considers regularized transition amplitudes $A_\Sigma[\Gamma]$ defined as

$$A_\Sigma[\Gamma] = \sum_{\substack{T_c \cong \Sigma \times I \\\text{s.t.} \\partial T_c = \Gamma}} \mu(T_c) e^{iS_{cl}[T_c]/\hbar}. \quad (3.5)$$

$\Gamma$ is the triangulation of the boundary $\partial T_c$ of the causal triangulation $T_c$, $\mu(T_c)$ is the measure, equal to the inverse of the order of the automorphism group of the causal triangulation $T_c$, and $S_{cl}[T_c]$ is the translation of the action $S_{cl}[g]$ into the Regge calculus of causal triangulations. For $D > 2$ dimensions, analytic calculations of the transition amplitudes $A_\Sigma[\Gamma]$, even for the simplest nontrivial cases, are not currently possible. To study the quantum theory of gravity defined by the transition amplitudes $A_\Sigma[\Gamma]$, one therefore employs numerical techniques, specifically Monte Carlo methods. To enable the application of such methods, one first performs a Wick rotation of each causal triangulation by analytically continuing $\alpha$ to $-\alpha$ through the lower-half complex plane. This Wick rotation transforms the transition
The amplitude $A_{\Sigma}[^{E}\Gamma]$ is into the partition function

$$Z_{\Sigma}[^{E}\Gamma] = \sum_{\{T_c\} \geq \Sigma \times 1, T_c|_{\partial T_c} = \Gamma} \mu(T_c) e^{-S_{cl}[T_c]}$$

(3.6)

in which $S_{cl}[T_c]$ is the resulting real-valued Euclidean action. Since one can only numerically simulate finite causal triangulations, one chooses to consider the partition function (3.6) for fixed numbers $\tilde{T}$ of time slices and $\tilde{N}_D$ of causal $D$-simplices. Accordingly, Monte Carlo methods produce ensembles of causal triangulations representative of those contributing to the (canonical) partition function

$$Z_{\Sigma}[^{E}\Gamma] = \sum_{\{T_c\} \geq \Sigma \times 1, T_c|_{\partial T_c} = \Gamma} \mu(T_c) e^{-S_{cl}[T_c]}$$

(3.7)

related by Laplace transform to the (grand canonical) partition function (3.6).

We take the action $S_{cl}[g]$ as that of $(2 + 1)$-dimensional Einstein gravity:

$$S_{cl}[g] = \frac{1}{16\pi G_0} \int_M d^3x \sqrt{-g} \left( R - 2\Lambda_0 \right) + \frac{1}{8\pi G_0} \int_{\partial M} d^2y \sqrt{|\gamma|} K.$$  

(3.8)

The first term in the action (3.8)—the bulk term—is the Einstein-Hilbert action in which $G_0$ is the bare Newton constant, $R$ is the Ricci scalar of the metric tensor $g$, and $\Lambda_0$ is a positive bare cosmological constant. The second term in the action (3.8)—the boundary term—is the Gibbons-Hawking-York action in which $K$ is the trace of the extrinsic curvature of the
metric tensor $\gamma$ [98, 240]. We choose to consider a spacetime manifold $M$ isomorphic to the direct product $S^2 \times I$ of a 2-sphere $S^2$ and a real interval $I$. In this case the boundary $\partial \mathcal{T}_c$ consists of two disconnected components: an initial spacelike 2-sphere $S^2_i$ and a final spacelike 2-sphere $S^2_f$. Drawing on previous results of Hartle and Sorkin [114], Ambjørn et al [22], and Anderson et al [31], Cooperman and Miller derived the form of the action $S^{(E)}_{cl}[\mathcal{T}_c]$ arising from the action (3.8) for this case:

$$
S^{(E)}_{cl}[\mathcal{T}_c] = \frac{i a}{8 \pi G_0} \left[ \frac{2 \pi}{i} \left( N^{SL}_1 - N^{SL}_1(S^2_i) - N^{SL}_1(S^2_f) \right) - \frac{1}{i} \varphi^{(2,2)}_{SL} \left( 2 N^{(2,2)}_3 - N^{(2,2)}_{3^\uparrow}(S^2_i) - N^{(2,2)}_{3^\uparrow}(S^2_f) \right) - \frac{1}{i} \varphi^{(1,3)}_{SL} \left( 4 N^{SL}_1 - 2 N^{SL}_1(S^2_i) - 2 N^{SL}_1(S^2_f) \right) - 2 \pi i \sqrt{-\alpha} N^{SL}_1 + 4 i \sqrt{-\alpha} \varphi^{(2,2)}_{TL} N^{(2,2)}_3 + 3 i \sqrt{-\alpha} \varphi^{(1,3)}_{TL} N^{(1,3)}_3 + 3 i \sqrt{-\alpha} \varphi^{(3,1)}_{TL} N^{(3,1)}_3 \right) - \frac{i \Lambda_0}{8 \pi G_0} \left[ \varphi^{(2,2)}_3 N^{(2,2)}_3 + \varphi^{(1,3)}_3 N^{(1,3)}_3 + V^{(3,1)}_3 N^{(3,1)}_3 \right] + \frac{i a}{8 \pi G_0} \left[ \frac{\pi}{i} N^{SL}_1(S^2_i) - \frac{2}{i} \varphi^{(3,1)}_{SL} N^{SL}_1(S^2_i) - \frac{1}{i} \varphi^{(2,2)}_{SL} N^{(2,2)}_{3^\uparrow}(S^2_i) \right] + \frac{i a}{8 \pi G_0} \left[ \frac{\pi}{i} N^{SL}_1(S^2_f) - \frac{2}{i} \varphi^{(3,1)}_{SL} N^{SL}_1(S^2_f) - \frac{1}{i} \varphi^{(2,2)}_{SL} N^{(2,2)}_{3^\uparrow}(S^2_f) \right].
$$

The first five lines derive from the Ricci scalar part of the bulk term of equation (3.8), the fourth line derives from the cosmological constant part of the bulk term of equation (3.8), and the last two lines derive from the initial and final boundary terms of equation (3.8). $N^{SL}_1$ is the number of spacelike 1-simplices (edges) and $N^{TL}_1$ is the number of timelike 1-simplices. $N^{(2,2)}_{3^\uparrow}(S^2)$ is the number of future-directed $(2, 2)$ 3-simplices attached to the
2-sphere $S^2$, and $N^{(2,2)}_{34}(S^2)$ is the number of past-directed $(2, 2)$ 3-simplices attached to the 2-sphere $S^2$. $\vartheta_{SL}^{(p,q)}$ is the Euclidean dihedral angle about a spacelike 1-simplex of a $(p, q)$ 3-simplex, $\vartheta_{TL}^{(p,q)}$ is the Euclidean dihedral angle about a timelike 1-simplex of a $(p, q)$ 3-simplex, and $V_3^{(p,q)}$ is the Euclidean spacetime 3-volume of a $(p, q)$ 3-simplex. We refer the reader to [22, 31] for explicit expressions for these Euclidean dihedral angles and spacetime 3-volumes.

If the initial and final boundary 2-spheres $S^2_i$ and $S^2_f$ are identified, yielding periodic boundary conditions in the temporal direction, then the action $S_{cl}(E)[T_c]$ simplifies considerably [22]:

$$S_{cl}^{(E)}[T_c] = -k_0 N_0 + k_3 N_3.$$  (3.10)

$N_0$ is the number of 0-simplices (vertices), $N_3$ is the number of 3-simplices, and the bare couplings $k_0$ and $k_3$ are the following dimensionless combinations of $G_0$, $\Lambda_0$, and $a$:

$$k_0 = 2\pi ak$$  (3.11a)

$$k_3 = \frac{a^3 \lambda}{4\sqrt{2}} + 2\pi ak \left[\frac{3}{\pi} \cos^{-1} \left(\frac{1}{3}\right) - 1\right]$$  (3.11b)

with

$$k = \frac{1}{8\pi G_0}$$  (3.12a)

$$\lambda = \frac{\Lambda_0}{8\pi G_0}$$  (3.12b)
We set $\alpha = 1$ because the value of $\alpha$ (once the Wick rotation has been performed) is irrelevant in $2+1$ dimensions.

When referring to an ensemble of causal triangulations with fixed initial and final boundary 2-spheres $S^2_i$ and $S^2_f$, we employ the couplings $k_0$ and $k_3$ instead of the couplings $k$ and $\lambda$ of equation (3.9) to facilitate contact with previous work. By the given values of $k_0$ and $k_3$, we mean the values dictated by the relations (3.11) for the values of $k$ and $\lambda$ actually characterizing the given ensemble. An ensemble of causal triangulations is therefore characterized by the number $\bar{T}$ of time slices, the number $\bar{N}_3$ of 3-simplices, the value of the coupling $k_0$, and the triangulations $\Gamma(S^2_i)$ and $\Gamma(S^2_f)$ of the initial and final boundary 2-spheres $S^2_i$ and $S^2_f$. As explained, for instance in [78], we must tune the coupling $k_3$ to its critical value $k^c_3$ to ensure that causal triangulations constructed from $\bar{N}_3$ 3-simplices dominate the partition function (3.7) for the action (3.9). The value $k^c_3$ is therefore not independent of the other quantities characterizing an ensemble of causal triangulations.

The triangulations $\Gamma(S^2_i)$ and $\Gamma(S^2_f)$ completely characterize the geometries of the initial and final boundary 2-spheres $S^2_i$ and $S^2_f$, constituting a sizeable amount of boundary data on which the partition function (3.7) depends. Cooperman and Miller restricted attention to only one aspect of the geometries of the triangulations $\Gamma(S^2_i)$ and $\Gamma(S^2_f)$: their discrete spatial 2-volumes as measured by the numbers $N_{2}^{SL}(S^2_i)$ and $N_{2}^{SL}(S^2_f)$ of spacelike 2-simplices (equilateral triangles) comprising the 2-spheres $S^2_i$ and $S^2_f$. The dependence of the partition function (3.7) on $N_{2}^{SL}(S^2_i)$ and $N_{2}^{SL}(S^2_f)$ is not merely the simplest to consider: in the absence of a physically relevant characterization of the geometries of the triangula-
tions \( \Gamma(S^2_i) \) and \( \Gamma(S^2_f) \), the dependence of the partition function (3.7) on other aspects of these geometries is difficult to study meaningfully. To probe only the dependence on the initial and final numbers \( N_{S^2}^\text{SL}(S^2_i) \) and \( N_{S^2}^\text{SL}(S^2_f) \) of spacelike 2-simplices, Cooperman and Miller proceeded as follows. They generated \( N \) random triangulations \( \Gamma(S^2_i) \) of the 2-sphere \( S^2_i \) constructed from precisely \( N_{S^2}^\text{SL}(S^2_i) \) spacelike 2-simplices and \( N \) random triangulations \( \Gamma(S^2_f) \) of the 2-sphere \( S^2_f \) constructed from precisely \( N_{S^2}^\text{SL}(S^2_f) \) spacelike 2-simplices; they randomly paired the former \( N \) triangulations with the latter \( N \) triangulations to form \( N \) pairs of initial and final boundary triangulations \( \Gamma(S^2_i) \) and \( \Gamma(S^2_f) \); they generated an ensemble of causal triangulations for each of these \( N \) pairs; and they combined these \( N \) ensembles into a single averaged ensemble. 3

By choosing to consider the dependence of the partition function (3.7) only on \( N_{S^2}^\text{SL}(S^2_i) \) and \( N_{S^2}^\text{SL}(S^2_f) \), Cooperman and Miller emulated virtually all previous studies of causal dynamical triangulations in \( 2 + 1 \) dimensions (and in \( 3 + 1 \) dimensions) in probing the ground state of the quantum geometry defined by an ensembles of causal triangulations. Prior investigations examined the spacetime manifold structure \( S^2 \times S^1 \) for which the temporal direction is periodically identified. Such studies probe the ground state of quantum geometry in the sense that there are no boundary conditions to induce excitations of the quantum geometry. Although Cooperman and Miller explored transition amplitudes with the spacetime manifold structure \( S^2 \times I \) in [78], their averaging over all geometrical degrees of freedom of the boundary 2-spheres except for their discrete spatial 2-volumes results in

---

3Technically, the procedure of Cooperman and Miller assumes a constant measure over all causal triangulations with initial and final boundary triangulations \( \Gamma(S^2_i) \) and \( \Gamma(S^2_f) \) constructed respectively from precisely \( N_{S^2}^\text{SL}(S^2_i) \) and \( N_{S^2}^\text{SL}(S^2_f) \) spacelike 2-simplices (for given values of \( T, \bar{N}_3, \text{ and } k_0 \)) [78].
boundary conditions that do not induce excitations of the quantum geometry.

Monte Carlo methods do not give us access to the partition function (3.7) itself; they yield only a representative sample of causal triangulations contributing to the path summation defining the partition function (3.7). This fact poses no problem of principle: we do have access to the expectation values of observables in the quantum state defined by the partition function (3.7). One computes the expectation value $\mathbb{E}[O]$ of an observable $O$ in this quantum state as follows:

$$\mathbb{E}[O] = \frac{1}{Z_{\Sigma}[\Gamma]} \sum_{\substack{\mathcal{T}_c \cong \Sigma \times I \\
\mathcal{T}_c|_{\partial \mathcal{T}_c} = \Gamma \\
T(\mathcal{T}_c) = T \\
N_D(\mathcal{T}_c) = N_D \}} \mu(\mathcal{T}_c) e^{-S_{cl}(\mathcal{T}_c)} \mathcal{O}[\mathcal{T}_c].$$  \hspace{1cm} (3.13)

We approximate the expectation value $\mathbb{E}[O]$ by its average

$$\langle O \rangle = \frac{1}{N(\mathcal{T}_c)} \sum_{j=1}^{N(\mathcal{T}_c)} O[\mathcal{T}_c^{(j)}]$$  \hspace{1cm} (3.14)

over an ensemble of $N(\mathcal{T}_c)$ causal triangulations generated by Monte Carlo methods. The Metropolis algorithm behind these simulations guarantees that

$$\mathbb{E}[O] = \lim_{N(\mathcal{T}_c) \to \infty} \langle O \rangle.$$  \hspace{1cm} (3.15)

Numerical measurements of certain observables’ ensemble averages have revealed that the model defined by the partition function (3.7) for the action (3.10) exhibits two phases of
quantum geometry separated by a first-order phase transition: the decoupled phase, labeled A, for coupling $k_0 > k_0^c$ and the condensed phase, labeled C, for coupling $k_0 < k_0^c$ [23, 137]. Cooperman and Miller found that phase C also exists within the model defined by the partition function (3.7) for the action (3.9) [78]. We restrict attention to values of the coupling $k_0$ that fall within phase C as only the quantum geometry defined by ensembles of causal triangulations with phase C possesses physical properties. We explore these properties in sections 3.5, 3.6, and 3.7.

### 3.5 Evidence and conjecture

We now review and expand upon the evidence that led Cooperman and Miller to formulate their conjecture. Following several previous authors [9, 10, 15, 14, 18, 23, 24, 26, 25, 31, 137], Cooperman and Miller performed measurements of the number $N_{SL}^2(\tau)$ of spacelike 2-simplices as a function of the discrete time coordinate $\tau$ labeling the distinguished foliation’s time slices [78]. $N_{SL}^2(\tau)$ quantifies the evolution of discrete spatial 2-volume in the distinguished foliation.

Cooperman and Miller first considered the following two ensembles of causal triangulations. For $\bar{T} = 29$, $\bar{N}_3 = 30850$, and $k_0 = 1.00$, we display the ensemble average $\langle N_{SL}^2(\tau) \rangle$ for $N_{SL}^2(S_1^2) = N_{SL}^2(S_2^2) = 4$ in figure 3.2(a) and for $N_{SL}^2(S_1^2) = N_{SL}^2(S_2^2) = 100$ in figure 3.2(b).\(^4\) The plot in figure 3.2(a) shows the behavior of $\langle N_{SL}^2(\tau) \rangle$ previously un-

\(^4\)We average over ensembles containing on the order of $10^5$ causal triangulations to obtain sufficient statistics for the measurements reported in figures 3.5 and 3.6 below. We choose $\bar{N}_3$ on the order of $3 \times$
Figure 3.2: Ensemble average number $\langle N_{2}^{\text{SL}} \rangle$ of spacelike 2-simplices as a function of the discrete time coordinate $\tau$ for $\bar{T} = 29$, $\bar{N}_3 = 30850$, and $k_0 = 1.00$. (a) $N_{2}^{\text{SL}}(S_1^2) = N_{2}^{\text{SL}}(S_f^2) = 4$ (b) $N_{2}^{\text{SL}}(S_1^2) = N_{2}^{\text{SL}}(S_f^2) = 100$. We have taken this data from [78].

Understood as characteristic of phase C [13, 15, 14, 18, 23, 24, 26, 25, 31, 46, 76, 78, 137]: $\langle N_{2}^{\text{SL}}(\tau) \rangle$ smoothly increases from its minimal value of 4 at the initial boundary 2-sphere $S_1^2$ to its maximal value at the central time slice and symmetrically decreases from its maximal value to its minimal value of 4 at the final boundary 2-sphere $S_f^2$. As several authors have previously demonstrated [13, 15, 14, 23, 24, 26, 25, 31, 46, 76, 78, 137], and as we demonstrate once more in section 3.6, the ground state solution—Euclidean de Sitter space—of a minisuperspace model based on Euclidean Einstein gravity accurately describes the shape of $\langle N_{2}^{\text{SL}}(\tau) \rangle$. The plot in figure 3.2(b) shows that the characteristic behavior of $\langle N_{2}^{\text{SL}}(\tau) \rangle$ continues to be manifest even for boundary 2-spheres with nonminimal discrete spatial 2-volumes. Cooperman and Miller demonstrated, moreover, that a portion of Euclidean de Sitter space accurately describes the shape of $\langle N_{2}^{\text{SL}}(\tau) \rangle$ in this case as well [78].

Cooperman and Miller next increased further the discrete spatial 2-volumes of the initial $10^5$, which gives physically reliable results without overly long computational times. Note that the minimal piecewise-Euclidean simplicial 2-sphere is constructed from four 2-simplices.
Figure 3.3: Ensemble average number $\langle N_{2}^{\text{SL}} \rangle$ of spacelike 2-simplices as a function of the discrete time coordinate $\tau$ for $\bar{T} = 29, \bar{N}_3 = 30850,$ and $k_0 = 1.00$. (a) $N_{2}^{\text{SL}}(S_i^2) = N_{2}^{\text{SL}}(S_f^2) = 500$ (b) $N_{2}^{\text{SL}}(S_i^2) = N_{2}^{\text{SL}}(S_f^2) = 700$ (c) $N_{2}^{\text{SL}}(S_i^2) = N_{2}^{\text{SL}}(S_f^2) = 900$. We have taken this data from [78].

and final boundary 2-spheres. For $\bar{T} = 29, \bar{N}_3 = 30850, k_0 = 1.00$, we display $\langle N_{2}^{\text{SL}}(\tau) \rangle$ for $N_{2}^{\text{SL}}(S_i^2) = N_{2}^{\text{SL}}(S_f^2) = 500$ in figure 3.3(a), for $N_{2}^{\text{SL}}(S_i^2) = N_{2}^{\text{SL}}(S_f^2) = 700$ in figure 3.3(b), and for $N_{2}^{\text{SL}}(S_i^2) = N_{2}^{\text{SL}}(S_f^2) = 900$ in figure 3.3(c).\(^5\)

We considered two further ensembles of causal triangulations. For $\bar{T} = 29, \bar{N}_3 = 30850, k_0 = 1.00$, we display $\langle N_{2}^{\text{SL}}(\tau) \rangle$ for $N_{2}^{\text{SL}}(S_i^2) = N_{2}^{\text{SL}}(S_f^2) = 500$ in figure 3.3(a), for $N_{2}^{\text{SL}}(S_i^2) = N_{2}^{\text{SL}}(S_f^2) = 700$ in figure 3.3(b), and for $N_{2}^{\text{SL}}(S_i^2) = N_{2}^{\text{SL}}(S_f^2) = 900$ in figure 3.3(c). As Cooperman and Miller remarked, the shape of $\langle N_{2}^{\text{SL}}(\tau) \rangle$ for those ensembles represented in figures 3.3 and 3.4 is possibly of a hyperbolic sinusoidal character. They hypothesized accordingly that a portion of Lorentzian de Sitter spacetime might accurately describe the shape of $\langle N_{2}^{\text{SL}}(\tau) \rangle$ for these ensembles [78]. We test this hypothesis in section 3.6.

\(^5\)The ensemble of causal triangulations characterized by $\bar{T} = 29, \bar{N}_3 = 30850, k_0 = 1.00$, and $N_{2}^{\text{SL}}(S_i^2) = N_{2}^{\text{SL}}(S_f^2) = 300$ is very close to the transition of $\langle N_{2}^{\text{SL}}(\tau) \rangle$ from being concave-down to being concave-up for these values of $\bar{T}, \bar{N}_3$ and $k_0$. We have not yet performed Monte Carlo simulations for sufficiently long computer times to determine on which side of the transition this ensemble falls.
ensemble average connected 2-point function $\langle n_{2}^{\text{SL}}(\tau) n_{2}^{\text{SL}}(\tau') \rangle$ of deviations $n_{2}^{\text{SL}}(\tau)$ in the number $N_{2}^{\text{SL}}(\tau)$ of spacelike 2-simplices from the ensemble average $\langle N_{2}^{\text{SL}}(\tau) \rangle$ defined as

$$
\langle n_{2}^{\text{SL}}(\tau) n_{2}^{\text{SL}}(\tau') \rangle = \frac{1}{N(T_{c})} \sum_{j=1}^{N(T_{c})} \left[ n_{2}^{\text{SL}}(\tau) \right]_{j} \left[ n_{2}^{\text{SL}}(\tau') \right]_{j}
$$

for

$$
\left[ n_{2}^{\text{SL}}(\tau) \right]_{j} = \left[ N_{2}^{\text{SL}}(\tau) \right]_{j} - \langle N_{2}^{\text{SL}}(\tau) \rangle.
$$

$\langle n_{2}^{\text{SL}}(\tau) n_{2}^{\text{SL}}(\tau') \rangle$ is a $\bar{T} \times \bar{T}$ real symmetric matrix, which we diagonalize to obtain its eigenvectors $\eta_{j}(\tau)$ and associated eigenvalues $\lambda_{j}$. For the (Euclidean-like) ensemble $\mathcal{E}_{E}$ of causal triangulations characterized by $\bar{T} = 21$, $\bar{N}_{3} = 30850$, $k_{0} = 1.00$, and $N_{2}^{\text{SL}}(S_{i}^{2}) = N_{2}^{\text{SL}}(S_{f}^{2}) = 4$, we display the first three eigenvectors $\eta_{j}(\tau)$ and the eigenvalues $\lambda_{j}$ of $\langle n_{2}^{\text{SL}}(\tau) n_{2}^{\text{SL}}(\tau') \rangle$ in figures 3.5(a) and 3.6(a).\(^{6}\) The plots in figures 3.5(a) and 3.6(a)\(^{6}\)

---

\(^{6}\)We employ the ensemble $\mathcal{E}_{E}$ characterized by $\bar{T} = 21$, $\bar{N}_{3} = 30850$, $k_{0} = 1.00$, and $N_{2}^{\text{SL}}(S_{i}^{2}) = N_{2}^{\text{SL}}(S_{f}^{2}) = 4$ as a point of comparison for two reasons. First, our analysis of $\langle n_{2}^{\text{SL}}(\tau) \rangle$ for the ensemble of causal triangulations characterized by $\bar{T} = 29$, $\bar{N}_{3} = 30850$, $k_{0} = 1.00$, and $N_{2}^{\text{SL}}(S_{i}^{2}) = N_{2}^{\text{SL}}(S_{f}^{2}) = 4$ indicates the presence of a stalk, resulting in the first eigenvector $\eta_{1}(\tau)$ possessing three rather than two nodes.
Figure 3.5: First three eigenvectors $\eta_j(\tau)$ of the ensemble average connected 2-point function $\langle n_2^{SL}(\tau) n_2^{SL}(\tau') \rangle$ of deviations $n_2^{SL}$ in the number of spacelike 2-simplices as a function of the discrete time coordinate $\tau$ for $N_3 = 30850$ and $k_0 = 1.00$ (a) $\bar{T} = 21$ and $N_2^{SL}(S_2) = N_2^{SL}(S_2^f) = 4$ (ensemble $E_E$) (b) $\bar{T} = 29$ and $N_2^{SL}(S_2) = N_2^{SL}(S_2^f) = 600$ (ensemble $E_L$). We do not indicate the scale of the eigenvectors $\eta_j(\tau)$ as their normalization is arbitrary.

show the behavior of $\langle n_2^{SL}(\tau) n_2^{SL}(\tau') \rangle$ previously understood as characteristic of phase C [15, 14, 76]. As Ambjørn et al [15, 14] and Cooperman [76] have previously demonstrated in the case of 3+1 dimensions, and as we demonstrate for the first time in 2+1 dimensions in section 3.6, the connected 2-point function of linear gravitational perturbations propagating on Euclidean de Sitter space accurately describes the shape of $\langle n_2^{SL}(\tau) n_2^{SL}(\tau') \rangle$, both its eigenvectors $\eta_j(\tau)$ and its eigenvalues $\lambda_j$.

For the (Lorentzian-like) ensemble $E_L$ of causal triangulations characterized by $\bar{T} = 29$, see [14] for an explanation. Second, our analysis in section 3.6 of $\langle N_2^{SL}(\tau) \rangle$ for the ensemble $E_E$ yields a quality of fit comparable to that for the ensemble $E_L$ characterized by $\bar{T} = 29$, $N_3 = 30850$, $k_0 = 1.00$, and $N_2^{SL}(S_2^f) = N_2^{SL}(S_2^f) = 600$, which we anonymously introduced with figure 3.4(a) and formally introduce with figures 3.5(b) and 3.6(b).
Figure 3.6: Eigenvalues $\lambda_j$ of the ensemble average connected 2-point function $\langle n_2^{\text{SL}}(\tau) n_2^{\text{SL}}(\tau') \rangle$ of deviations $n_2^{\text{SL}}$ in the number of spacelike 2-simplices as a function of the discrete time coordinate $\tau$ for $\tilde{N}_3 = 30850$ and $k_0 = 1.00$ (a) $\bar{T} = 21$ and $N_2^{\text{SL}}(S^2_i) = N_2^{\text{SL}}(S^2_f) = 4$ (ensemble $\mathcal{E}_E$) (b) $\bar{T} = 29$ and $N_2^{\text{SL}}(S^2_i) = N_2^{\text{SL}}(S^2_f) = 600$ (ensemble $\mathcal{E}_L$).

$\tilde{N}_3 = 30850$, $k_0 = 1.00$, and $N_2^{\text{SL}}(S^2_i) = N_2^{\text{SL}}(S^2_f) = 600$, we display the first three eigenvectors $\eta_j(\tau)$ and the associated eigenvalues $\lambda_j$ of $\langle n_2^{\text{SL}}(\tau) n_2^{\text{SL}}(\tau') \rangle$ in figures 3.5(b) and 3.6(b).

The shapes of the eigenvectors $\eta_j(\tau)$ for the ensemble $\mathcal{E}_L$ differ subtly yet notably from the shapes of the eigenvectors $\eta_j(\tau)$ for the ensemble $\mathcal{E}_E$. The spectrum of eigenvalues $\lambda_j$ for the ensemble $\mathcal{E}_L$ also differs subtly yet notably from the spectrum of eigenvalues $\lambda_j$ for the ensemble $\mathcal{E}_E$. We hypothesize accordingly that linear gravitational perturbations propagating on a portion of Lorentzian de Sitter spacetime might accurately describe the shape of $\langle n_2^{\text{SL}}(\tau) n_2^{\text{SL}}(\tau') \rangle$, both its eigenvectors $\eta_j(\tau)$ and its eigenvalues $\lambda_j$, for the ensemble $\mathcal{E}_L$.

We test this hypothesis in section 3.6.

These finding led Cooperman and Miller to formulate the following conjecture: geometries resembling Lorentzian de Sitter spacetime, not Euclidean de Sitter space, on suf-
ficiently large scales dominate the partition function (3.6) for the action (3.10) defining the ground state of \((2 + 1)\)-dimensional causal dynamical triangulations for spherical spatial topology [78]. Cooperman and Miller also suggested that their conjecture’s scenario might arise \textit{via} a mechanism similar to that of the Hartle-Hawking no-boundary proposal in which complex geometries contribute to the partition function [113]. We subject their conjecture to a first test in section 3.6, obtaining evidence in its favor; however, we argue for a more straightforward explanation of the above findings in section 3.7, refuting their conjecture.

### 3.6 Analysis and support

We now perform a preliminary test of the conjecture of Cooperman and Miller by analyzing the measurements of \(\langle N_{2}^{SL}(\tau) \rangle\) and \(\langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle\) reported in section 3.5 on the basis of their conjecture. To connect their conjecture with these measurements, we attempt to describe these measurements within a simple yet nontrivial model inspired by their conjecture: a minisuperspace truncation of \((2 + 1)\)-dimensional Einstein gravity having either Lorentzian de Sitter spacetime or Euclidean de Sitter space as its ground state. Several authors have previously employed this model’s Euclidean version [8, 9, 10, 13, 15, 14, 17, 18, 23, 24, 26, 25, 31, 45, 46, 76, 78, 137], which Ambjörn, Jurkiewicz, and Loll first suggested [23, 24, 26, 25]. We specify the model’s metric tensor \(g\) by the line element

\[
\begin{align*}
\mathrm{d}s^2 &= \pm \omega^2 \mathrm{d}t^2 + \rho^2(t) \left( \mathrm{d}\theta^2 + \sin^2 \theta \, \mathrm{d}\phi^2 \right) \\
&= \pm \omega^2 \mathrm{d}t^2 + \rho^2(t) \left( \mathrm{d}\theta^2 + \sin^2 \theta \, \mathrm{d}\phi^2 \right) 
\end{align*}
\] (3.18)
for positive constant $\omega$ and scale factor $\rho(t)$ with upper sign ($+$) for Euclidean signature and the lower sign ($-$) for Lorentzian signature. For the line element (3.18), expressed in terms of the spatial 2-volume

$$V_2(t) = \int_0^\pi d\theta \int_0^{2\pi} d\phi \sqrt{g_{\theta\theta} g_{\phi\phi}} = 4\pi \rho^2(t), \quad (3.19)$$

the Einstein-Hilbert action, including the Gibbons-Hawking-York action, given in equation (3.8) for Lorentzian signature, becomes

$$S_{cl}[V_2] = \pm \frac{\omega}{32\pi G} \int_{t_i}^{t_f} dt \left[ \frac{\dot{V}_2^2(t)}{\omega^2 V_2(t)} \mp 4\Lambda V_2(t) \right] \quad (3.20)$$

after integration by parts. As in equation (3.18), the upper signs correspond to Euclidean signature, and the lower signs correspond to Lorentzian signature. $^7$ $G$ and $\Lambda$ are now the renormalized Newton and cosmological constants. The maximally symmetric extremum of the action (3.20) for Euclidean signature is Euclidean de Sitter space, for which

$$V_2^{(EdS)}(t) = 4\pi \ell_{dS}^2 \cos^2 \left( \frac{\omega t}{\ell_{dS}} \right) \quad (3.21)$$

$^7$Typically, in Euclidean signature the action (3.20) has an overall negative sign, which is surprisingly absent in the large-scale effective action of causal dynamical triangulations [8, 9, 10, 15, 14, 18, 24, 26, 25]. See [17] for a plausible yet tentative explanation.
with \( t \in [-\pi \ell_{ds}/2\omega, +\pi \ell_{ds}/2\omega] \); the maximally symmetric extremum of the action (3.20) for Lorentzian signature is Lorentzian de Sitter spacetime, for which

\[
V_{2}^{(LdS)}(t) = 4\pi \ell_{ds}^2 \cosh^2 \left( \frac{\omega t}{\ell_{ds}} \right)
\]  

(3.22)

with \( t \in (-\infty, +\infty) \). \( \ell_{ds} = \sqrt{1/\Lambda} \) is the de Sitter length.

We first model the ensemble average number \( \langle N_{2}^{SL}(\tau) \rangle \) of spacelike 2-simplices as a function of the discrete time coordinate \( \tau \) on the basis of the spatial 2-volumes \( V_{2}^{(EdS)}(t) \) and \( V_{2}^{(LdS)}(t) \) given in equations (3.21) and (3.22). In particular, we derive a discrete analogue \( N_{2}^{SL}(\tau) \) appropriate to causal triangulations of each of the spatial 2-volumes \( V_{2}^{(EdS)}(t) \) and \( V_{2}^{(LdS)}(t) \), and we subsequently perform a best fit of \( N_{2}^{SL}(\tau) \) to \( \langle N_{2}^{SL}(\tau) \rangle \).

Several authors have previously performed such a derivation in the case of Euclidean de Sitter space [15, 14, 17, 18, 23, 24, 26, 25, 31, 46, 76, 78]; we adapt their techniques to the case of a portion of Lorentzian de Sitter spacetime. We begin by assuming a canonical finite-size scaling Ansatz based on the double scaling limit

\[
V_{3} = \lim_{N_{3} \to \infty} \lim_{a \to 0} C_{3} N_{3} a^{3}
\]  

(3.23)

of the spacetime 3-volume \( V_{3} \); in the infinite-volume \( (N_{3} \to \infty) \) and continuum \( (a \to 0) \) limits, the discrete spacetime 3-volume \( C_{3} N_{3} a^{3} \) approaches the constant value \( V_{3} \). \( C_{3} \) is the effective discrete spacetime 3-volume of a single 3-simplex. Evidence for the applicability of this Ansatz to the scaling of \( \langle N_{2}^{SL}(\tau) \rangle \) is presented in [17, 24, 26, 25, 46]. The motivation
for this Ansatz is as following: \( V_3 \) is the largest-scale physical observable present in our model, so, of all possible discrete observables, we expect the discrete spacetime 3-volume to scale canonically with \( N_3 \) and \( a \). In appendix 3.A we employ the finite-size scaling Ansatz based on equation (3.23) to derive the discrete analogue \( \mathcal{N}_{2}^{\text{SL}}(\tau) \) for each of the spatial 2-volumes \( V_{2}^{(\text{EdS})}(t) \) and \( V_{2}^{(\text{LdS})}(t) \) restricted to the finite global time interval \([t_i, t_f]\).

In the case of Euclidean de Sitter space, we derive that

\[
\mathcal{N}_{2}^{\text{SL}}(\tau) = \frac{\langle N_{3}^{(1,3)} \rangle \cos^2 \left( \frac{\tau}{\bar{s}_0 \langle N_{3}^{(1,3)} \rangle^{1/3}} \right)}{\bar{s}_0 \langle N_{3}^{(1,3)} \rangle^{1/3} + \sin \left( \frac{\tau - \bar{\tau}_1}{\bar{s}_0 \langle N_{3}^{(1,3)} \rangle^{1/3}} \right) \cos \left( \frac{\bar{\tau}_1 + \tau}{\bar{s}_0 \langle N_{3}^{(1,3)} \rangle^{1/3}} \right)},
\]

as previously determined in [78], and, in the case of Lorentzian de Sitter spacetime, we derive that

\[
\mathcal{N}_{2}^{\text{SL}}(\tau) = \frac{\langle N_{3}^{(1,3)} \rangle \cosh^2 \left( \frac{\tau}{\bar{s}_0 \langle N_{3}^{(1,3)} \rangle^{1/3}} \right)}{\bar{s}_0 \langle N_{3}^{(1,3)} \rangle^{1/3} + \sinh \left( \frac{\tau - \bar{\tau}_1}{\bar{s}_0 \langle N_{3}^{(1,3)} \rangle^{1/3}} \right) \cosh \left( \frac{\bar{\tau}_1 + \tau}{\bar{s}_0 \langle N_{3}^{(1,3)} \rangle^{1/3}} \right)},
\]

\( N_{3}^{(1,3)} \) is the number of (1, 3) 3-simplices,

\[
\bar{s}_0 = \frac{2^{1/3} (1 + \xi)^{1/3} \ell_{\text{ds}}}{\omega V_{3}^{1/3}}
\]

is a fit parameter, and \( \xi \) is the ratio of \( \langle N_{3}^{(2,2)} \rangle \) to \( \langle N_{3}^{(1,3)} \rangle + \langle N_{3}^{(3,1)} \rangle \). We now perform best fits of \( \mathcal{N}_{2}^{\text{SL}}(\tau) \) to the measurements of \( \langle \mathcal{N}_{2}^{\text{SL}}(\tau) \rangle \) following the procedure of [78]. Within the caption to each figure depicting such a fit, we report the value \( \chi^2_{\text{red}} \) of the \( \chi^2 \) per degree of freedom for each fit.
To establish a point of comparison, we first consider the ensemble $\mathcal{E}_E$ characterized by $\bar{T} = 21$, $\bar{N}_3 = 30850$, $k_0 = 1.00$, and $N_2^{SL}(S_f^2) = N_2^{SL}(S_i^2) = 4$ (Euclidean-like ensemble $\mathcal{E}_E$) overlain with the best fit discrete analogue $N_2^{SL}(\tau)$ (black line) of the spatial 2-volume $V_2^{(EdS)}(t)$ as a function of the global time coordinate $t$ of Euclidean de Sitter space. $\chi^2_{\text{red}} = 79.91$.

We now test the hypothesis that a portion of Lorentzian de Sitter spacetime accurately describes the ensemble average number $\langle N_2^{SL}(\tau) \rangle$ of spacelike 2-simplices as a function
Figure 3.8: Ensemble average number $\langle N^\text{SL}_2 \rangle$ of spacelike 2-simplices as a function of the discrete time coordinate $\tau$ (blue circles) for $\bar{T} = 29$, $\bar{N}_3 = 30850$, and $k_0 = 1.00$ overlain with the best fit discrete analogue $N^\text{SL}_2(\tau)$ (black line) of the spatial 2-volume $V_2^{(\text{LdS})}(t)$ as a function of the global time coordinate $t$ of Lorentzian de Sitter spacetime. (a) $N^\text{SL}_2(S^2_i) = N^\text{SL}_2(S^2_f) = 500$, $\chi^2_{\text{red}} = 169.86$. (b) $N^\text{SL}_2(S^2_i) = N^\text{SL}_2(S^2_f) = 700$, $\chi^2_{\text{red}} = 143.44$. (c) $N^\text{SL}_2(S^2_i) = N^\text{SL}_2(S^2_f) = 900$, $\chi^2_{\text{red}} = 1435.51$.

of the discrete time coordinate $\tau$ for the Lorentzian-like ensembles represented in figures 3.3 and 3.4. We consider the five ensembles of causal triangulations represented in figures 3.3 and 3.4 including the ensemble $\mathcal{E}_L$. For $\bar{T} = 29$, $\bar{N}_3 = 30850$, $k_0 = 1.00$, we display $\langle N^\text{SL}_2(\tau) \rangle$ overlain with the best form of $N^\text{SL}_2(\tau)$, given in equation (3.25), for $N^\text{SL}_2(S^2_i) = N^\text{SL}_2(S^2_f) = 500$ in figure 3.8(a), for $N^\text{SL}_2(S^2_i) = N^\text{SL}_2(S^2_f) = 700$ in figure 3.8(b), for $N^\text{SL}_2(S^2_i) = N^\text{SL}_2(S^2_f) = 900$ in figure 3.8(c), for $N^\text{SL}_2(S^2_i) = N^\text{SL}_2(S^2_f) = 600$ in figure 3.9(a), and for $N^\text{SL}_2(S^2_i) = N^\text{SL}_2(S^2_f) = 800$ in figure 3.9(b). Visually, $N^\text{SL}_2(\tau)$ again fits $\langle N^\text{SL}_2(\tau) \rangle$ quite satisfactorily. As measured by $\chi^2_{\text{red}}$, the quality of the fits of $N^\text{SL}_2(\tau)$, given in equation (3.25), to $\langle N^\text{SL}_2(\tau) \rangle$ for these five Lorentzian-like ensembles is comparable to the quality of the fit of $N^\text{SL}_2(\tau)$, given in equation (3.24), to $\langle N^\text{SL}_2(\tau) \rangle$ for Euclidean-like ensembles [78]. In particular, these fits for the ensembles $\mathcal{E}_E$ and $\mathcal{E}_L$ have nearly equivalent $\chi^2_{\text{red}}$ values, motivating our choice to compare the ensembles $\mathcal{E}_E$ and $\mathcal{E}_L$. There
We now extend our model to include linear gravitational perturbations \( v_2(t) \) propagating on either Euclidean de Sitter space or Lorentzian de Sitter spacetime. In the path integral formalism one computes the connected 2-point function \( \mathbb{E}_{EdS}[v_2(t) v_2(t')] \) of perturbations
$v_2(t)$ about Euclidean de Sitter space as

$$\mathbb{E}\text{EdS}[v_2(t) v_2(t')] = \frac{\int d\mu(v_2) v_2(t) v_2(t') e^{iS_{cl}[v_2]/\hbar}}{\int d\mu(v_2) e^{iS_{cl}[v_2]/\hbar}} e^{-S_{cl}[v_2]/\hbar}, \quad (3.27)$$

in which $S_{cl}[v_2]$ is the action (3.20) in Euclidean signature for the spatial 2-volume $V_2(t)$ perturbed by $v_2(t)$ about $V_2^{(EdS)}(t)$, and the connected 2-point function $\mathbb{E}\text{LdS}[v_2(t) v_2(t')]$ of perturbations $v_2(t)$ about Lorentzian de Sitter spacetime as

$$\mathbb{E}\text{LdS}[v_2(t) v_2(t')] = \frac{\int d\mu(v_2) v_2(t) v_2(t') e^{iS_{cl}[v_2]/\hbar}}{\int d\mu(v_2) e^{iS_{cl}[v_2]/\hbar}} e^{iS_{cl}[v_2]/\hbar}, \quad (3.28)$$

in which $S_{cl}[v_2]$ is the action (3.20) in Lorentzian signature for the spatial 2-volume $V_2(t)$ perturbed by $v_2(t)$ about $V_2^{(LdS)}(t)$. Expanding the action (3.20) in Euclidean signature to second order in $v_2(t)$, assuming that $V_2^{(EdS)}(t) \gg v_2(t)$, we find that

$$S_{cl}[v_2] = S_{cl}[V_2^{(EdS)}] - \frac{1}{64\pi^2 G \ell_{dS}^3} \int_{\tilde{t}} d\tilde{t} v_2(\tilde{t}) \sec^2 \tilde{t} \left( \frac{d^2}{dt^2} + 2 \tan \tilde{t} \frac{dt}{dt} + 2 \sec^2 \tilde{t} \right) v_2(\tilde{t}) + O[(v_2)^3], \quad (3.29)$$

for $\tilde{t} = \omega t/\ell_{dS}$. The terms of first order in $v_2(t)$ vanish because $V_2^{(EdS)}(t)$ is an extremum of the action (3.20) in Euclidean signature. Expanding the action (3.20) in Lorentzian
signature to second order in \(v_2(t)\), assuming that \(V_2^{(\text{LdS})}(t) \gg v_2(t)\), we find that

\[
S_{\text{cl}}[v_2] = S_{\text{cl}}[v_2^{(\text{LdS})}] + \frac{1}{64\pi^2 G\ell_{\text{ds}}^3} \int_0^\infty \! dt \, v_2(\tilde{t}) \, \text{sech}^2 \tilde{t} \left[ \frac{d^2}{d\tilde{t}^2} - 2 \tanh \tilde{t} \frac{d}{d\tilde{t}} - 2 \text{sech}^2 \tilde{t} \right] v_2(\tilde{t}) + O[(v_2)^3]
\]

(3.30)

for \(\tilde{t} = \omega t / \ell_{\text{ds}}\). The terms of first order in \(v_2(t)\) vanish because \(V_2^{(\text{LdS})}(t)\) is an extremum of the action (3.20) in Lorentzian signature. A standard calculation now gives that

\[
\mathbb{E}[v_2(t) v_2(t')] = \left[ \frac{1}{\hbar} \mathcal{M}(t, t') \right]^{-1},
\]

(3.31)

in which

\[
\mathcal{M}(t, t') = \frac{\delta^2 S_{\text{cl}}[v_2]}{\delta v_2(t) \delta v_2(t')} \bigg|_{v_2(t) = 0, v_2(t') = 0}
\]

(3.32)

is the van Vleck-Morette determinant. For perturbations \(v_2(t)\) about the spatial 2-volume \(V_2^{(\text{EdS})}(t)\) of Euclidean de Sitter space,

\[
\mathcal{M}(t, t') = \frac{1}{64\pi^2 G\ell_{\text{ds}}^3} \sec^2 \tilde{t} \left[ \frac{d^2}{d\tilde{t}^2} + 2 \tan \tilde{t} \frac{d}{d\tilde{t}} + 2 \sec^2 \tilde{t} \right],
\]

(3.33)

and, for perturbations \(v_2(t)\) about the spatial 2-volume \(V_2^{(\text{LdS})}(t)\) of Lorentzian de Sitter spacetime,

\[
\mathcal{M}(t, t') = \frac{1}{64\pi^2 G\ell_{\text{ds}}^3} \text{sech}^2 \tilde{t} \left[ \frac{d^2}{d\tilde{t}^2} - 2 \tanh \tilde{t} \frac{d}{d\tilde{t}} - 2 \text{sech}^2 \tilde{t} \right].
\]

(3.34)
One can show moreover that

\[ M(t, t') = \sum_{j=1}^{\infty} \mu_j \nu_j(t) \nu_j(t') \]  

(3.35)

in which \( \nu_j(t) \) are the eigenfunctions of the operator \( M(t, t') \) with associated eigenvalues \( \mu_j \) satisfying the integral constraint

\[ \int_{t_i}^{t_f} dt \omega \nu_j(t) = 0 \]  

(3.36)

and the boundary conditions \( \nu_j(t_i) = 0 \) and \( \nu_j(t_f) = 0 \). Accordingly,

\[ \mathbb{E}[\nu_2(t) \nu_2(t')] = \sum_{j=1}^{\infty} \frac{\hbar}{\mu_j} \nu_j(t) \nu_j(t') \]  

(3.37)

assuming that \( \mu_j \neq 0 \) for all \( j \), which holds in the cases under consideration.

We next model the ensemble average connected 2-point function \( \langle n_2^{SL}(\tau) n_2^{SL}(\tau') \rangle \) of deviations \( n_2^{SL}(\tau) \) in the number \( N_2^{SL}(\tau) \) of spacelike 2-simplices as a function of the discrete time coordinate \( \tau \) on the basis of the 2-point functions \( \mathbb{E}_{EdS}[\nu_2(t) \nu_2(t')] \) and \( \mathbb{E}_{Lds}[\nu_2(t) \nu_2(t')] \) given in equations (3.27) and (3.28). In particular, we derive a discrete analogue \( n_2^{SL}(\tau) n_2^{SL}(\tau') \) appropriate to causal triangulations of each of the 2-point functions \( \mathbb{E}_{EdS}[\nu_2(t) \nu_2(t')] \) and \( \mathbb{E}_{Lds}[\nu_2(t) \nu_2(t')] \), and we subsequently perform a fit of \( n_2^{SL}(\tau) n_2^{SL}(\tau') \) to \( \langle n_2^{SL}(\tau) n_2^{SL}(\tau') \rangle \). Ambjørn et al [15, 14] and Cooperman [76] have previously performed such a derivation in the case of Euclidean de Sitter space in \( 3 + 1 \) di-
dimensions; we adapt their techniques to the cases of Euclidean de Sitter space in $2+1$ dimensions and a portion of Lorentzian de Sitter spacetime in $2+1$ dimensions. We again assume the finite-size scaling Ansatz based on equation (3.23). In appendix 3.B we employ the Ansatz based on equation (3.23) to derive the discrete analogue $n_{2}^{\text{SL}}(\tau)n_{2}^{\text{SL}}(\tau')$ for each of the 2-point functions $E_{\text{EdS}}[v_{2}(t)\, v_{2}(t')]$ and $E_{\text{LdS}}[v_{2}(t)\, v_{2}(t')]$. Specifically, we derive $n_{2}^{\text{SL}}(\tau)n_{2}^{\text{SL}}(\tau')$ in the form of equation (3.37), determining the eigenvectors $\nu_{j}(\tau)$ and associated eigenvalues $\mu_{j}$ of $n_{2}^{\text{SL}}(\tau)n_{2}^{\text{SL}}(\tau')$. We now perform fits of the eigenvectors $\nu_{j}(\tau)$ to the eigenvectors $\eta_{j}(\tau)$ and of the eigenvalues $\mu_{j}$ to the eigenvalues $\lambda_{j}$. Once the best fit of $N_{2}^{\text{SL}}(\tau)$ to $\langle N_{2}^{\text{SL}}(\tau) \rangle$ fixes the fit parameter $\bar{s}_{0}$, there is in fact no fitting to perform aside from a single overall rescaling of the eigenvalues corresponding to the value of $1/64\pi^{2}\hbar G\ell_{\text{dS}}^{3}$. Employing this value of $\bar{s}_{0}$ accords with our treatment of $v_{2}(t)$ as a perturbation.

To establish a point of comparison, we first consider the ensemble $\mathcal{E}_{E}$ characterized by $\bar{T} = 21$, $\bar{N}_{3} = 30850$, $k_{0} = 1.00$, and $N_{2}^{\text{SL}}(S_{1}^{2}) = N_{2}^{\text{SL}}(S_{1}^{2}) = 4$, for which, as depicted in figure 3.10, $\langle n_{2}^{\text{SL}}(\tau)\, n_{2}^{\text{SL}}(\tau') \rangle$ exhibits the characteristic behavior of phase C. We display the first six eigenvectors $\eta_{j}(\tau)$ of $\langle n_{2}^{\text{SL}}(\tau)\, n_{2}^{\text{SL}}(\tau') \rangle$ overlain with the corresponding eigenvectors $\nu_{j}(\tau)$ of the discrete analogue $n_{2}^{\text{SL}}(\tau)n_{2}^{\text{SL}}(\tau')$ of $E_{\text{EdS}}[v_{2}(t)\, v_{2}(t')]$ in figure 3.10. We display the eigenvalues $\lambda_{j}$ of $\langle n_{2}^{\text{SL}}(\tau)\, n_{2}^{\text{SL}}(\tau') \rangle$ overlain with the corresponding eigenvalues $\mu_{j}$ of $n_{2}^{\text{SL}}(\tau)n_{2}^{\text{SL}}(\tau)$ in figure 3.11. The fits of $\nu_{j}(\tau)$ to $\eta_{j}(t)$ and of $\mu_{j}$ to $\lambda_{j}$ are representative of the application of the above Euclidean model to measurements of
Figure 3.10: First six eigenvectors \( \eta_j(\tau) \) (blue circles) of the ensemble average connected 2-point function \( \langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle \) of deviations \( n_{2}^{SL}(\tau) \) in the number of spacelike 2-simplices as a function of the discrete time coordinate \( \tau \) for \( \bar{T} = 21, \bar{N}_3 = 30850, k_0 = 1.00, \) and \( N_{2}^{SL}(S_2^T) = N_{2}^{SL}(S_2^F) = 4 \) (Euclidean-like ensemble \( \mathcal{E}_E \)) overlain with the eigenvectors \( \nu_j(\tau) \) (black lines) of the discrete analogue \( n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \) of the connected 2-point function \( \mathbb{E}_{EdS}[v_2(t) v_2(t')] \) of perturbations \( v_2(t) \) in the spatial 2-volume \( V_2^{(EdS)}(t) \) as a function of the global time coordinate \( t \) of Euclidean de Sitter space.

\[ \langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle \] [15, 14, 76]. Clearly, this model provides an accurate description of the connected 2-point function \( \langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle \) for the ensemble \( \mathcal{E}_E \).

We now test the hypothesis that the connected 2-point function \( \mathbb{E}_{LdS}[v_2(t) v_2(t')] \) of linear gravitational perturbations \( v_2(t) \) propagating on Lorentzian de Sitter spacetime accurately describes the shape of the ensemble average connected 2-point function \( \langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle \) for the Lorentzian-like ensembles represented in figures 3.3 and 3.4. We consider only the ensemble \( \mathcal{E}_L \) characterized by \( \bar{T} = 29, \bar{N}_3 = 30850, k_0 = 1.00, \) and \( N_{2}^{SL}(S_2^T) = N_{2}^{SL}(S_2^F) = 600 \). We display the first six eigenvectors \( \nu_j(\tau) \) of \( \langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle \) overlain with the cor-
Figure 3.11: Eigenvalues $\lambda_j$ (blue circles) of the ensemble average connected 2-point function $\langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle$ of deviations $n_{2}^{SL}(\tau)$ in the number of spacelike 2-simplices as a function of the discrete time coordinate $\tau$ for $\bar{T} = 21$, $\bar{N}_3 = 30850$, $k_0 = 1.00$, and $N_{2}^{SL}(S^2) = N_{2}^{SL}(S^2) = 4$ (Euclidean-like ensemble $\mathcal{E}_E$) overlain with the eigenvalues $\mu_j$ (black lines) of the discrete analogue $n_{2}^{SL}(\tau) n_{2}^{SL}(\tau')$ of the connected 2-point function $\mathbb{E}_{EdS}[v_2(t) v_2(t')]$ of perturbations $v_2(t)$ in the spatial 2-volume $V_2^{(EdS)}(t)$ as a function of the global time coordinate $t$ of Euclidean de Sitter space.

responding eigenvectors $\nu_j(\tau)$ of the discrete analogue $n_{2}^{SL}(\tau) n_{2}^{SL}(\tau)$ of $\mathbb{E}_{EdS}[v_2(t) v_2(t')]$ in figure 3.12. We display the eigenvalues $\lambda_j$ of $\langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle$ overlain with the corresponding eigenvalues $\mu_j$ of $n_{2}^{SL}(\tau) n_{2}^{SL}(\tau)$ in figure 3.13. Clearly, the above Lorentzian model provides an accurate description of the connected 2-point function $\langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle$ for the ensemble $\mathcal{E}_L$.

These analyses, straightforwardly interpreted, provide evidence supporting the conjecture of Cooperman and Miller: a portion of Lorentzian de Sitter spacetime accurately describes the shape of $\langle N_{2}^{SL}(\tau) \rangle$, and the connected 2-point function of linear perturbations propagating on Lorentzian de Sitter spacetime accurately describes the shape of $\langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle$ for the Lorentzian-like ensembles represented in figures 3.3 and 3.4. Nev-
Figure 3.12: First six eigenvectors $\eta_j(\tau)$ (blue circles) of the ensemble average connected 2-point function $\langle n^{SL}_2(\tau) n^{SL}_2(\tau') \rangle$ of deviations $n^{SL}_2(\tau)$ in the number of spacelike 2-simplices as a function of the discrete time coordinate $\tau$ for $\bar{T} = 29$, $\bar{N}_3 = 30850$, $k_0 = 1.00$, and $N^{SL}_2(S^2) = N^{SL}_2(S^2_f) = 600$ (Lorentzian-like ensemble $E_L$) overlaid with the eigenvectors $\nu_j(\tau)$ (black lines) of the discrete analogue $n^{SL}_2(\tau) n^{SL}_2(\tau')$ of the connected 2-point function $E_{LdS}[v_2(t) v_2(t')]$ of perturbations $v_2(t)$ in the spatial 2-volume $V^{(LdS)}_2(t)$ as a function of the global time coordinate $t$ of Lorentzian de Sitter space.

ertheless, we proffer an even more straightforward explanation of these results in section 3.7, casting serious doubt on the conjecture of Cooperman and Miller.

Ideally, we would extend our analysis of the above modeling of $\langle N^{SL}_2(\tau) \rangle$ and $\langle n^{SL}_2(\tau) n^{SL}_2(\tau') \rangle$ in two directions. First, we would perform a finite-size scaling analysis in which we consider ensembles of causal triangulations characterized by increasing numbers $\bar{N}_3$ of 3-simplices—and commensurately increasing numbers $\bar{T}$ of time slices and $N^{SL}_2(S^2)$ and $N^{SL}_2(S^2_f)$ of initial and final boundary spacelike 2-simplices—to extrapolate the accuracy of our modeling towards the infinite-volume limit. Such a finite-size scaling
Figure 3.13: Eigenvalues $\lambda_j$ (blue circles) of the ensemble average connected 2-point function $\langle n_{2}^{\text{SL}}(\tau) n_{2}^{\text{SL}}(\tau') \rangle$ of deviations $n_{2}^{\text{SL}}(\tau)$ in the number of spacelike 2-simplices as a function of the discrete time coordinate $\tau$ for $\bar{T} = 29$, $\bar{N}_3 = 30850$, $k_0 = 1.00$, and $N_{2}^{\text{SL}}(S_2^i) = N_{2}^{\text{SL}}(S_2^f) = 600$ (Lorentzian-like ensemble $\mathcal{E}_L$) overlain with the eigenvalues $\mu_j$ (black lines) of the discrete analogue $n_{2}^{\text{SL}}(\tau) n_{2}^{\text{SL}}(\tau')$ of the connected 2-point function $\mathbb{E}_{\text{LdS}}[v_2(t) v_2(t')]$ of perturbations $v_2(t)$ in the spatial 2-volume $V_2^{(\text{LdS})}(t)$ as a function of the global time coordinate $t$ of Lorentzian de Sitter space.

analysis is more difficult to perform in the context of transition amplitudes: the manner in which one must commensurately increase $\bar{T}$, $N_{2}^{\text{SL}}(S_2^i)$, and $N_{2}^{\text{SL}}(S_2^f)$ with $\bar{N}_3$ to consider transition amplitudes related by the finite-size scaling Ansatz based on equation (3.23) is nontrivial. For this reason we have not yet performed any scaling analyses of the transition amplitudes; rather, we rely on the similarities of our numerical measurements to those of previous studies as justification for our use of the finite-size scaling Ansatz based on equation (3.23). Second, we would consider models based on departures from Einstein gravity—for instance, Hořava-Lifshitz or higher-order gravity—to assess our model’s accuracy. Cooperman and Houthoff perform such an analysis, though only for Euclidean-like ensembles, in a forthcoming paper [124].
3.7 Argument and refutation

Extraordinary claims require extraordinary evidence. The conjecture of Cooperman and Miller constitutes an extraordinary claim, but we now argue that the analyses presented in section 3.6 of the measurements presented in section 3.5 do not furnish extraordinary evidence. We offer an alternative explanation of these measurements and their analysis, one much more plausible as well as much more mundane.

We based the analyses of section 3.6 on a minisuperspace truncation of (2+1)-dimensional Einstein gravity with either Euclidean de Sitter space or Lorentzian de Sitter spacetime as its ground state. As we presented this model in section 3.6, we did not incorporate with sufficient care the setting of our numerical simulations of causal triangulations. Recall from section 3.4 that we run a given simulation at fixed number $\tilde{N}_3$ of 3-simplices and at fixed numbers $N_{2}^{\text{SL}}(S^2_i)$ and $N_{2}^{\text{SL}}(S^2_f)$ of initial and final boundary spacelike 2-simplices.$^8$ We accounted for these constraints by normalizing $V_2^{(dS)}(t)$ to $V_3$ and $N_{2}^{\text{SL}}(\tau)$ to $\tilde{N}_3$ in the derivation of appendix 3.A and by enforcing boundary conditions on $N_{2}^{\text{SL}}(\tau)$ in the best fit to $\langle N_2^{\text{SL}}(\tau) \rangle$. We did not, however, explicitly include constraints implementing a fixed spacetime 3-volume $V_3$ and fixed initial and final spatial 2-volumes $V_2(t_i)$ and $V_2(t_f)$ in the action (3.20) defining our model. Of course, we did not need to include these constraints in the action (3.20) because we imposed these constraints on our model’s solutions. The analysis of section 3.6 is therefore correct, but, by imposing these constraints on our

$^8$We also fix the number $\tilde{T}$ of time slices; however, our model allows for an arbitrary lapse—the constant $\omega$, which propagates into the fit parameter $\tilde{q}_0$—so we do not impose a constraint associated with fixed $\tilde{T}$. 

model’s solutions instead of including these constraints in our model’s action, we obscured
the interpretation of this analysis. We now augment our model’s action with the relevant
constraints and carefully extract their consequences for the interpretation of our model.

Explicitly imposing these constraints in the action (3.20) for Euclidean signature, we
arrive at the augmented action

\[
S_{cl}[V_2] = \frac{\omega}{32\pi G} \int_{t_1}^{t_f} dt \left[ \frac{V_2^2(t)}{\omega^2 V_2(t)} - 4\Lambda V_2(t) \right] + \lambda_{V_3} \left[ \int_{t_1}^{t_f} dt \omega V_2(t) - V_3 \right] \\
+ \lambda_i \left[ \int_{t_1}^{t_i} dt \omega \delta(t - t_i) V_2(t) - V_2(t_i) \right] \\
+ \lambda_f \left[ \int_{t_1}^{t_i} dt \omega \delta(t - t_i) V_2(t) - V_2(t_f) \right]
\]  
(3.38)

in which \(\lambda_{V_3}\) is the Lagrange multiplier associated with the constraint of fixed spacetime
3-volume \(V_3\), and \(\lambda_i\) and \(\lambda_f\) are the Lagrange multipliers associated with the constraints
of fixed initial and final spatial 2-volumes \(V_2(t_i)\) and \(V_2(t_f)\). The cosmological constant
term also acts to constrain the spacetime 3-volume \(V_3\) with the cosmological constant itself
serving as the associated Lagrange multiplier. We include the additional constraint of fixed
\(V_3\) to make our argument more transparent; in particular, we think of the cosmological
constant \(\Lambda\) as fixed and the Lagrange multiplier \(\lambda_{V_3}\) as variable.

Varying the action (3.38) with respect to \(V_2(t)\), we obtain the equation of motion

\[
2V_2(t)\ddot{V}_2(t) - \dot{V}_2^2(t) \pm 4\omega^2(\Lambda - 8\pi G\lambda_{V_3})V_2^2(t) = 0,
\]  
(3.39)
having the general solution

\[
V_2(t) = \begin{cases} 
  A \cos \left[ \sqrt{\Lambda - 8\pi G \lambda_{V_3}} (t - t_0) \right] & \text{if } \Lambda - 8\pi G \lambda_{V_3} > 0 \\
  A (t - t_0)^2 & \text{if } \Lambda - 8\pi G \lambda_{V_3} = 0 \\
  A \cosh \left[ \sqrt{8\pi G \lambda_{V_3} - \Lambda} (t - t_0) \right] & \text{if } \Lambda - 8\pi G \lambda_{V_3} < 0
\end{cases}
\]

(3.40)

for integration constants \(A\) and \(t_0\). Varying the action (3.38) with respect to \(\lambda_{V_3}\), we obtain the constraint

\[
V_3 = \int_{t_i}^{t_f} dt \omega V_2(t),
\]

(3.41)

and varying the action (3.38) with respect to \(\lambda_i\) and \(\lambda_f\) constrains \(V_2(t)\) to have the initial and final boundary values \(V_2(t_i)\) and \(V_2(t_f)\).

We now focus on the spatial 2-volume \(V_2(t)\) for \(\Lambda - 8\pi G \lambda_{V_3} > 0\) given in the first line of equation (3.40). Let \(\ell_{\text{eff}}^{-2} = \Lambda - 8\pi G \lambda_{V_3}\). Recalling equation (3.21), which gives the spatial 2-volume \(V_2^{(\text{EdS})}(t)\) of Euclidean de Sitter space, we observe that the spatial 2-volume \(V_2(t)\) for \(\ell_{\text{eff}}^{-2} > 0\) is precisely that of Euclidean de Sitter space if \(A = 4\pi \ell_{\text{eff}}^2\) (for de Sitter length \(\ell_{\text{eff}}\)). We recover this solution of section 3.6 because we have merely changed the stage at which we enforce the constraints on our model. Assuming further that \(V_2(t_i) = V_2(t_f)\) dictates that \(t_0 = 0\). The difference \(t_f - t_i\) (and, indeed, the value of \(t_i = -t_f\)) is then determined in terms of \(V_2(t_i) = V_2(t_f)\) and \(\ell_{\text{eff}}\):

\[
\omega(t_f - t_i) = \ell_{\text{eff}} \cos^{-1} \sqrt{\frac{V_2(t_f)}{4\pi \ell_{\text{eff}}^2}}.
\]

(3.42)
Substituting $V_2(t)$ for $\ell_{\text{eff}}^{-2} > 0$, $A = 4\pi \ell_{\text{eff}}^2$, and $t_0 = 0$ into equation (3.41), we obtain

$$V_3 = 4\pi \ell_{\text{eff}}^3 \left[ \cos^{-1} \left( \frac{V_2(t_f)}{4\pi \ell_{\text{eff}}^2} \right) + \sqrt{\frac{V_2(t_f)}{4\pi \ell_{\text{eff}}^2}} \sqrt{1 - \frac{V_2(t_f)}{4\pi \ell_{\text{eff}}^2}} \right], \quad (3.43)$$

which coincides with equation (3.51) upon making use of equation (3.42). Solving equation (3.43) for $4\pi \ell_{\text{eff}}^2$ and replacing $4\pi \ell_{\text{eff}}^2$ in equation (3.40), we obtain

$$V_2(t) = \frac{V_3}{\ell_{\text{eff}}} \left[ \cos^{-1} \left( \frac{V_2(t_f)}{4\pi \ell_{\text{eff}}^2} \right) + \sqrt{\frac{V_2(t_f)}{4\pi \ell_{\text{eff}}^2}} \sqrt{1 - \frac{V_2(t_f)}{4\pi \ell_{\text{eff}}^2}} \right]^{-1} \cos^2 \left( \frac{\omega t}{\ell_{\text{eff}}} \right), \quad (3.44)$$

which coincides with equation (3.53) upon making use of equation (3.42). Equation (3.44) gives the spatial 2-volume as a function of the global time coordinate of a portion of Euclidean de Sitter space constrained to have spacetime 3-volume $V_3$ and initial and final boundary spatial 2-volumes $V_2(t_i) = V_2(t_f)$. The discrete analogue $N_{2}^\text{SL}(\tau)$ of the spatial 2-volume $V_2(t)$ given in equation (3.44) coincides with equation (3.60), making explicit the validity of the analysis of section 3.6.

For given values of $G$ and $\Lambda$, with either the gauge fixing $\omega = \text{constant}$ or the gauge fixing $t_f = \text{constant}$, we may choose values for $V_3$ and $V_2(t_f)$ and determine (if possible) the value of $\lambda V_3$ dictated by the chosen values of $V_3$ and $V_2(t_f)$. If $V_2(t_f)$ is not too large in comparison to $V_3$, then $\Lambda > 8\pi G \lambda V_3$, and the solution is a portion of Euclidean de Sitter space; however, if $V_2(t_f)$ is too large in comparison to $V_3$, then $\Lambda < 8\pi G \lambda V_3$, and the solution is a portion of Lorentzian de Sitter spacetime.

We now give examples of these two cases. Suppose that $G = 1/8\pi$ and $\Lambda = 1$. Choose
Figure 3.14: Spatial 2-volume $V_2(t)$ as a function of the global time coordinate $t$ for $A = 4\pi\ell^2_{\text{eff}}$, $t_0 = 0$, $G = 1/8\pi$, and $\Lambda = 1$ (a) $V_3 = 13500$, $V_2(t_i) = V_2(t_f) = 0$, $\omega = 1.38$, $\lambda_{V_3} = 0.99$, and $\ell^2_{\text{eff}} = 77.63$ (b) $V_3 = 3300$, $V_2(t_i) = V_2(t_f) = 600$, $\omega = 0.31$, $\lambda_{V_3} = 1.04$, and $\ell^2_{\text{eff}} = -22.35$.

First $V_3 = 13500$, $V_2(t_f) = 0$, and $t_f = 10$. Equations (3.42) and (3.43) then yield $\omega = 1.38$ and $\lambda_{V_3} = 0.99$ for which $\ell^2_{\text{eff}} = 77.63$. We display the spatial 2-volume $V_2(t)$ for this case in figure 3.14(a). This first example models the circumstances of the Euclidean-like ensemble $\mathcal{E}_E$: in this case $N_{2}^{\text{SL}}(\tau_f)$ is not too large in comparison to $N_3$, so the discrete analogue $N_{2}^{\text{SL}}(\tau)$ of the spatial 2-volume $V_2^{\text{(EdS)}}(t)$ of Euclidean de Sitter space accurately describes $\langle N_{2}^{\text{SL}}(\tau) \rangle$. Compare figure 3.14(a) to figure 3.7.\(^9\)

Choose second $V_3 = 3300$, $V_2(t_f) = 600$, and $t_f = 14$. Equations (3.42) and (3.43) then yield $\omega = 0.31$ and $\lambda_{V_3} = 1.04$ for which $\ell^2_{\text{eff}} = -22.35$. We display the spatial 2-volume $V_2(t)$ for this case in figure 3.14(b). This second example models the circumstances of the Lorentzian-like ensemble $\mathcal{E}_L$: in this case $N_{2}^{\text{SL}}(\tau_f)$ is too large in comparison to $N_3$, \(^9\)We selected values of $V_3$, $V_2(t_f)$, and $t_f$ so that the plot of $V_2(t)$ in figure 3.14(a) matches the plot of $N_{2}^{\text{SL}}(\tau)$ in figure 3.7. While the values of $V_2(t_f)$ and $t_f$ coincide with the values of $N_{2}^{\text{SL}}(\tau_f)$ and $\tau_f$, the value of $V_3$ does not coincide with the value of $N_3$ because $V_3$ is related to $N_3$ through equation (3.23), which additionally involves $C_3$ and $a$. An analogous comment applies to the comparison just below of figures 3.14(b) and 3.9(a).
so the discrete analogue $N_{2}^{SL}(\tau)$ of the spatial 2-volume $V^{(LdS)}(t)$ of Lorentzian de Sitter spacetime accurately describes $\langle N_{2}^{SL}(\tau) \rangle$. Compare figure 3.14(b) to figure 3.9(a). The discrete analogue $N_{2}^{SL}(\tau)$ of the spatial 2-volume $V^{(LdS)}(t)$ of Lorentzian de Sitter spacetime nevertheless arises from a model based on Euclidean Einstein gravity. Furthermore, the operator $\mathcal{M}(t, t')$ derived from the action (3.38) for linear perturbations $v_{2}(t)$ about the spatial 2-volume $V_{2}(t)$ for $\ell_{\text{eff}}^{2} < 0$ coincides with the operator (3.34), so the discrete analogue $n_{2}^{SL}(\tau) n_{2}^{SL}(\tau')$ of the connected 2-point function $\mathbb{E}_{LdS}[v_{2}(t) v_{2}(t')]$ still serves as the correct model for the connected 2-point function $\langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle$.

The above discussion points towards an explanation of the measurements of $\langle N_{2}^{SL}(\tau) \rangle$ and $\langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle$ presented in section 3.5 and their analysis presented in section 3.6 different from that of the conjecture of Cooperman and Miller. The model based on a minisuperspace truncation of Euclidean Einstein gravity also accurately describes $\langle N_{2}^{SL}(\tau) \rangle$ and $\langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle$ for the Lorentzian-like ensembles: the interaction of the constraints of fixed spacetime 3-volume and fixed initial and final boundary spatial 2-volumes forces $\langle N_{2}^{SL}(\tau) \rangle$ and $\langle n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \rangle$ to be Lorentzian in form. We conclude accordingly that the geometries of causal triangulations comprising Lorentzian-like ensembles are not Lorentzian but Euclidean in nature.

Our argument does not, however, clinch the case against the conjecture of Cooperman and Miller: had we run our reasoning starting from the action (3.20) in Lorentzian signature, Euclidean de Sitter space would have arisen from Lorentzian de Sitter spacetime as the Lagrange multiplier $\lambda_{V_{3}}$ forced $\Lambda - 8\pi G \lambda_{V_{3}}$ to change sign, and we would have
concluded that geometries resembling Lorentzian de Sitter spacetime on sufficiently large scales dominate the ground state of causal dynamical triangulations. We chose to present our argument starting from the action (3.20) in Euclidean signature because we know that the configurations simulated numerically must be Euclidean in nature: the Metropolis algorithm simply cannot handle complex contributions to the partition function (3.7). Still, we would like more definitive evidence for the Euclidean nature of the causal triangulations of Lorentzian-like ensembles represented in figures 3.3 and 3.4.

The two observables that we measured—\( \langle N_{2}^{\text{SL}}(\tau) \rangle \) and \( \langle n_{2}^{\text{SL}}(\tau) n_{2}^{\text{SL}}(\tau') \rangle \)—probe the quantum geometry defined by an ensemble of causal triangulations only on its largest scales. Since we do not consider observables that probe this quantum geometry on small scales, we do not assess the nature—Euclidean or Lorentzian—of the quantum geometry on small scales. To test the conjecture of Cooperman and Miller more definitively, we would like to make a statement regarding the nature of the quantum geometry on smaller scales, in particular, regarding the nature of local interactions, which should naively appear quite different if they are in fact Lorentzian. We should therefore probe the quantum geometry on small scales by measuring appropriate observables.

Accordingly, we consider numerical measurements of the spectral dimension, a scale-dependent measure of the dimensionality of the quantum geometry, which probes the quantum geometry defined by an ensemble of causal triangulations on all scales. In appendix 3.C, following several previous authors [25, 27, 31, 45, 76, 79, 137], we define the spectral dimension \( D_{s}(\sigma) \) as a function of the diffusion time \( \sigma \), and we explain its numerical estima-
Figure 3.15: Ensemble average spectral dimension \( \langle D_s \rangle \) as a function of diffusion time \( \sigma \) for \( \bar{N}_3 = 30850 \) and \( k_0 = 1.00 \) (a) \( \bar{T} = 21 \) and \( N^\text{SL}_2(S^2) = N^\text{SL}_2(S^2_f) = 4 \) (b) \( \bar{T} = 29 \) and \( N^\text{SL}_2(S^2_f) = N^\text{SL}_2(S^2_f) = 600 \).

As in our analysis of the 2-point function \( \langle n^\text{SL}_2(\tau) n^\text{SL}_2(\tau') \rangle \), we compare the spectral dimension \( D_s(\sigma) \) of the ensemble \( \mathcal{E}_E \) characterized by \( \bar{T} = 21, \bar{N}_3 = 30580, k_0 = 1.00, \) and \( N^\text{SL}_2(S^2_f) = N^\text{SL}_2(S^2_f) = 4 \) to the spectral dimension \( D_s(\sigma) \) of the ensemble \( \mathcal{E}_L \) characterized by \( \bar{T} = 29, \bar{N}_3 = 30580, k_0 = 1.00, \) and \( N^\text{SL}_2(S^2_f) = N^\text{SL}_2(S^2_f) = 600 \). We display \( D_s(\sigma) \) for the ensemble \( \mathcal{E}_E \) in figure 3.15(a) and for the ensemble \( \mathcal{E}_L \) in figure 3.15(b).

The plot in figure 3.15(a) shows the behavior of \( D_s(\sigma) \) previously understood as characteristic of phase C [25, 27, 31, 45, 76, 79, 137]. For intermediate diffusion times (\( \sigma \sim 200 \) for \( \mathcal{E}_E, \sigma \sim 150 \) for \( \mathcal{E}_L \)), the spectral dimension peaks at approximately the topological dimension of 3; for smaller diffusion times (\( \sigma \leq 200 \) for \( \mathcal{E}_E, \sigma \leq 150 \) for \( \mathcal{E}_L \)), the spectral dimension dynamically reduces towards a value near 2; and for larger diffusion times (\( \sigma \geq 200 \) for \( \mathcal{E}_E, \sigma \geq 150 \) for \( \mathcal{E}_L \)), the spectral dimension decays exponentially in the presence of positive curvature. The two measurements of \( D_s(\sigma) \) displayed in figure 3.15 exhibit essentially the same qualitative behavior and similar quantitative behavior. The
maximal value of $D_s(\sigma)$ (2.96 for $\mathcal{E}_E$, 2.72 for $\mathcal{E}_L$) is the primary difference. As Benedetti and Henson found for Euclidean-like ensembles [45], the depression of $D_s(\sigma)$ below the topological value of 3 is a finite-size effect. We have verified that this depression is also a finite-size effect for Lorentzian-like ensembles. Although ensembles $\mathcal{E}_E$ and $\mathcal{E}_L$ are both characterized by $\bar{N}_3 = 30850$, we suspect that the ensemble $\mathcal{E}_L$ exhibits stronger finite-size effects because the random walker can only probe a small portion of a quantum geometry resembling Lorentzian de Sitter spacetime on sufficiently large scales. Since $D_s(\sigma)$ for the ensemble $\mathcal{E}_L$ behaves so similarly to $D_s(\sigma)$ for the ensemble $\mathcal{E}_E$, we take these measurements of $D_s(\sigma)$ as evidence that the geometries of causal triangulations comprising the ensemble $\mathcal{E}_L$ are Euclidean in nature, supporting our above conclusion.

### 3.8 Lorentzian from Euclidean

Studying the causal dynamical triangulations of $(2 + 1)$-dimensional Einstein gravity in the presence of initial and final spacelike boundaries, Cooperman and Miller identified several ensembles of causal triangulations the quantum geometry of which on sufficiently large scales appears to resemble closely that of Lorentzian de Sitter spacetime [78]. On the basis of these findings, they conjectured that the partition function (3.6) is dominated by causal triangulations the quantum geometry of which is nearly that of Lorentzian de Sitter spacetime on sufficiently large scales, possibly via a mechanism akin to that of the Hartle-Hawking no-boundary proposal. The conjecture of Cooperman and Miller presented an ex-
citing possibility: the definition of a Lorentzian quantum theory of gravity via a Euclidean path integral, alleviating the necessity of reversing the Wick rotation of causal dynamical triangulations. We have argued for a much more plausible and mundane explanation of their findings: the implementation and interaction of multiple constraints may result in the partition function (3.7) being dominated by (Euclidean) causal triangulations that closely resemble Lorentzian de Sitter spacetime on large scales. While not particularly exciting, our explanation adds one further piece of evidence for the proper behavior of the partition function defined via causal dynamical triangulations. Our explanation also serves as a cautionary tale: beware hastily drawing conclusions regarding signs of signature change within the partition function (3.6) of causal dynamical triangulations.

The issue of reversing the Wick rotation of causal dynamical triangulations thus remains. The results of modeling the large-scale quantum geometry within phase C on the basis of a minisuperspace truncation of Euclidean Einstein gravity, as exemplified by our modeling of the ensemble $E_E$ (and, indeed, also the ensemble $E_L$), suggest a straightforward possibility: since (Euclidean) causal triangulations resembling Euclidean de Sitter space on sufficiently large scales dominate the partition function (3.6), obtained by Wick rotation from the path sum (3.5), (Lorentzian) causal triangulations resembling Lorentzian de Sitter spacetime on sufficiently large scales dominate the path sum (3.5). For this interpretation to have force, one must establish a rigorous path from the Euclidean theory to the Lorentzian theory by demonstrating an Osterwalder-Schrader-type theorem for causal dynamical triangulations. Although technically challenging, achieving such a theorem is likely within
reach since the action $S^{(E)}_{\text{cl}}(\mathcal{J}_c)$ for Einstein gravity is reflection-positive, a key axiom of the Osterwalder-Schrader reconstruction theorem. We maintain that the promising results of causal dynamical triangulations warrant such an effort.

**Acknowledgments**

We thank Christian Anderson, David Kamensky, and especially Rajesh Kommu for allowing us to employ parts of their computer codes. We also thank Joe Henson, Ian Morrison, Erik Schnetter, and especially Steve Carlip and Renate Loll for useful discussions. JHC acknowledges support from the Department of Energy under grant DE-FG02-91ER40674 at the University of California, Davis and from Stichting voor Fundamenteel Onderzoek der Materie itself supported by Nederlandse Organisatie voor Wetenschappelijk Onderzoek. KL acknowledges support from the SURF program at Chapman University and the hospitality of the Department of Physics at the University of California, Davis. JMM acknowledge support from the National Science Foundation under REU grant PHY-1004848 at the University of California, Davis. This work utilized the Janus supercomputer, which is supported by the National Science Foundation (award number CNS-0821794) and the University of Colorado, Boulder. The Janus supercomputer is a joint effort of the University of Colorado, Boulder, the University of Colorado, Denver, and the National Center for Atmospheric Research. This research was supported in part by the Perimeter Institute for Theoretical Physics. Research at the Perimeter Institute is supported by the Government of
Canada through Industry Canada and by the Province of Ontario through the Ministry of Economic Development and Innovation.
Appendix

3.A Derivation of $\mathcal{N}^{\text{SL}}_2(\tau)$

We derive the discrete analogue $\mathcal{N}^{\text{SL}}_2(\tau)$ of the spatial 2-volume $V_2^{(\text{EdS})}(t)$ as a function of the global time coordinate $t$ of Euclidean de Sitter space, given in equation (3.21), and of the spatial 2-volume $V_2^{(\text{LdS})}(t)$ as a function of the global time coordinate $t$ of Lorentzian de Sitter spacetime, given in equation (3.22).

We start from the doubling scaling limit of the discrete spacetime 3-volume given in equation (3.23):

$$V_3 = \lim_{N_3 \to \infty} \lim_{a \to 0} C_3 N_3 a^3.$$  \hspace{2cm} (3.45)

Assuming that equation (3.45) holds for finite number $N_3$ of 3-simplices and lattice spacing $a$ without significant corrections, we express equation (3.45) with its left hand side as an integral over the global time coordinate $t$ and its right hand side as a sum over the discrete
time coordinate $\tau$:

$$\int_{t_i}^{t_f} dt \omega V_2(t) = 2C_3a^3(1 + \xi) \sum_{\tau=1}^{\tilde{T}} N_{2}^{\text{SL}}(\tau).$$  \hfill (3.46)

According to the finite-size scaling Ansatz based on equation (3.45), in the combination of the infinite volume and continuum limits, we expect the relation

$$\frac{\tau}{N_3^{1/3}} = \frac{t}{V_3^{1/3}}$$  \hfill (3.47)

between $\tau$ and $t$ and the relation

$$\frac{N_2^{\text{SL}}}{N_3^{2/3}} = \frac{V_2}{V_3^{2/3}}$$  \hfill (3.48)

between $N_2^{\text{SL}}$ and $V_2$. Constants of proportionality in the relations (3.47) and (3.48) are redundant for the following derivation. In these limits we identify the integral $\int dt V_3^{-1/3}$ with the sum $\sum_\tau \Delta \tau N_3^{-1/3}$ in equation (3.46), yielding

$$\omega V_3^{1/3} V_2(t) = 2C_3a^3(1 + \xi) N_3^{1/3} N_{2}^{\text{SL}}(\tau).$$  \hfill (3.49)

Solving for $N_2^{\text{SL}}(\tau)$, we obtain

$$N_2^{\text{SL}}(\tau) = \frac{\omega V_3^{1/3}}{2C_3N_3^{1/3} a^3(1 + \xi)} V_2(t).$$  \hfill (3.50)

We next need to substitute appropriate expressions for the spatial 2-volume $V_2(t)$. Consid-
ering the measurements of ensemble average $\langle N^{SL}_2(\tau) \rangle$ presented in section 3.5, we consider the finite portions of Euclidean de Sitter space and of Lorentzian de Sitter spacetime for which $t \in [t_i, t_f]$. This portion of Euclidean de Sitter space has spacetime 3-volume

$$V_3 = \int_{t_i}^{t_f} dt \omega V_2^{(EdS)}(t) = 2\pi \ell_{dS}^3 \left\{ \frac{\omega(t_f - t_i)}{\ell_{dS}} + \sin \left[ \frac{\omega(t_f - t_i)}{\ell_{dS}} \right] \cos \left[ \frac{\omega(t_f + t_i)}{\ell_{dS}} \right] \right\},$$

(3.51)

which is equivalent to equation (3.43), and this portion of Lorentzian de Sitter spacetime has spacetime 3-volume

$$V_3 = \int_{t_i}^{t_f} dt \omega V_2^{(LdS)}(t) = 2\pi \ell_{dS}^3 \left\{ \frac{\omega(t_f - t_i)}{\ell_{dS}} + \sinh \left[ \frac{\omega(t_f - t_i)}{\ell_{dS}} \right] \cosh \left[ \frac{\omega(t_f + t_i)}{\ell_{dS}} \right] \right\}.$$

(3.52)

Solving equation (3.51) for $4\pi \ell_{dS}^2$ in terms of $V_3$ and substituting into equation (3.21) yields

$$V_2^{(EdS)}(t) = \frac{2V_3}{\ell_{dS}} \left\{ \frac{\omega(t_f - t_i)}{\ell_{dS}} + \sin \left[ \frac{\omega(t_f - t_i)}{\ell_{dS}} \right] \cos \left[ \frac{\omega(t_f + t_i)}{\ell_{dS}} \right] \right\}^{-1} \cos^2 \left( \frac{\omega t}{\ell_{dS}} \right),$$

(3.53)

while solving equation (3.52) for $4\pi \ell_{dS}^2$ in terms of $V_3$ and substituting into equation (3.22) yields

$$V_2^{(LdS)}(t) = \frac{2V_3}{\ell_{dS}} \left\{ \frac{\omega(t_f - t_i)}{\ell_{dS}} + \sinh \left[ \frac{\omega(t_f - t_i)}{\ell_{dS}} \right] \cosh \left[ \frac{\omega(t_f + t_i)}{\ell_{dS}} \right] \right\}^{-1} \cosh^2 \left( \frac{\omega t}{\ell_{dS}} \right).$$

(3.54)
We substitute equation (3.53) into equation (3.50), obtaining

\[
N_{2SL}(\tau) = \frac{\omega V_3^{1/3}}{2C_3 N_3^{1/3} \alpha^3 (1 + \xi)} \frac{2V_3}{\ell_{ds}} \times \left\{ \frac{\omega(t_f - t_i)}{\ell_{ds}} + \sin \left[ \frac{\omega(t_f - t_i)}{\ell_{ds}} \right] \cos \left[ \frac{\omega(t_f + t_i)}{\ell_{ds}} \right] \right\}^{-1} \cos^2 \left( \frac{\omega t}{\ell_{ds}} \right),
\]

(3.55)

and we substitute equation (3.54) into equation (3.50), obtaining

\[
N_{2SL}(\tau) = \frac{\omega V_3^{1/3}}{2C_3 N_3^{1/3} \alpha^3 (1 + \xi)} \frac{2V_3}{\ell_{ds}} \times \left\{ \frac{\omega(t_f - t_i)}{\ell_{ds}} + \sinh \left[ \frac{\omega(t_f - t_i)}{\ell_{ds}} \right] \cosh \left[ \frac{\omega(t_f + t_i)}{\ell_{ds}} \right] \right\}^{-1} \cosh^2 \left( \frac{\omega t}{\ell_{ds}} \right).
\]

(3.56)

Using equation (3.45) and replacing \( t \) with \( V_3^{1/3} \tau/N_3^{1/3} \) according to relation (3.47), equation (3.55) becomes

\[
N_{2SL}(\tau) = \frac{\omega V_3^{1/3} N_3 \cos^2 \left( \frac{\omega V_3^{1/3} \tau}{\ell_{ds} N_3^{1/3}} \right)}{N_3^{1/3} (1 + \xi) \ell_{ds}} \left\{ \frac{\omega V_3^{1/3} \tau}{\ell_{ds} N_3^{1/3}} + \sin \left[ \frac{\omega V_3^{1/3} \tau}{\ell_{ds} N_3^{1/3}} \right] \cos \left[ \frac{\omega V_3^{1/3} \tau}{\ell_{ds} N_3^{1/3}} \right] \right\}.
\]

(3.57)

and equation (3.56) becomes

\[
N_{2SL}(\tau) = \frac{\omega V_3^{1/3} N_3 \cosh^2 \left( \frac{\omega V_3^{1/3} \tau}{\ell_{ds} N_3^{1/3}} \right)}{N_3^{1/3} (1 + \xi) \ell_{ds}} \left\{ \frac{\omega V_3^{1/3} \tau}{\ell_{ds} N_3^{1/3}} + \sinh \left[ \frac{\omega V_3^{1/3} \tau}{\ell_{ds} N_3^{1/3}} \right] \cosh \left[ \frac{\omega V_3^{1/3} \tau}{\ell_{ds} N_3^{1/3}} \right] \right\}.
\]

(3.58)

Substituting \( N_3^{(1,3)} \) for \( N_3 \) according to the identity \( N_3 = 2(1 + \xi)N_3^{(1,3)} \) and defining the
parameter

\[ \tilde{s}_0 = \frac{2^{1/3}(1 + \xi)^{1/3}}{\omega V_3^{1/3}}, \]  

(3.59)

we finally arrive at the discrete analogue \( \mathcal{N}^\text{SL}_2(\tau) \) of the spatial 2-volume \( V_2^{(EdS)}(t) \),

\[ \mathcal{N}^\text{SL}_2(\tau) = \frac{\langle N_3^{(1,3)} \rangle}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \left[ \frac{\tau_{f} + \tau_{i}}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} + \sin \left( \frac{\tau_{f} - \tau_{i}}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) \cos \left( \frac{\tau_{f} + \tau_{i}}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) \right]. \]  

(3.60)

and the discrete analogue \( \mathcal{N}^\text{SL}_2(\tau) \) of the spatial 2-volume \( V_2^{(LdS)}(t) \),

\[ \mathcal{N}^\text{SL}_2(\tau) = \frac{\langle N_3^{(1,3)} \rangle}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \left[ \frac{\tau_{f} + \tau_{i}}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} + \sinh \left( \frac{\tau_{f} - \tau_{i}}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) \cosh \left( \frac{\tau_{f} + \tau_{i}}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) \right]. \]  

(3.61)

For the case in which \( \tau_i = -\bar{T}/2 \) and \( \tau_f = \bar{T}/2 \), equation (3.60) simplifies to

\[ \mathcal{N}^\text{SL}_2(\tau) = \frac{\langle N_3^{(1,3)} \rangle}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \left[ \frac{\bar{T}}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} + \sin \left( \frac{\bar{T}}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) \right]^{-1} \cos^2 \left( \frac{\tau}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right), \]  

(3.62)

and equation (3.61) simplifies to

\[ \mathcal{N}^\text{SL}_2(\tau) = \frac{\langle N_3^{(1,3)} \rangle}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \left[ \frac{\bar{T}}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} + \sinh \left( \frac{\bar{T}}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) \right]^{-1} \cosh^2 \left( \frac{\tau}{\tilde{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right). \]  

(3.63)
3.B Derivation of \( n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \)

We derive the discrete analogue \( n_{2}^{SL}(\tau) n_{2}^{SL}(\tau') \) of the connected 2-point function \( \mathbb{E}_{EdS}[v_2(t) v_2(t)] \) of gravitational perturbations \( v_2(t) \) of the spatial 2-volume \( V_2^{(EdS)}(t) \) of Euclidean de Sitter space and of the connected 2-point function \( \mathbb{E}_{LdS}[v_2(t) v_2(t)] \) of gravitational perturbations \( v_2(t) \) of the spatial 2-volume \( V_2^{(LdS)}(t) \) of Lorentzian de Sitter spacetime.

We start from equations (3.33) and (3.34), the expressions for the van Vleck-Morette determinants \( \mathcal{M}(t, t') \). We first discretize the operator \( \mathcal{M}(t, t') \) on a 1-dimensional lattice of \( \bar{T} \) sites, transforming the differential operators of equations (3.33) and (3.34) into finite-difference operators. Specifically, replacing \( t \) with \( V_1^{1/3} \tau / N_3^{1/3} \) according to relation (3.47), substituting \( N_3^{(1,3)} \) for \( N_3 \) according to the identity \( N_3 = 2(1 + \xi)N_3^{(1,3)} \), and employing the definition (3.59) of the fit parameter \( \bar{s}_0 \), equation (3.33) becomes

\[
M(\tau, \tau') = \frac{1}{64\pi^2 G_{EdS}^3} \text{sec}^2 \left( \frac{\tau}{\bar{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) \left[ \frac{\Delta^2}{\Delta \tau^2} \right] \Delta^2 \Delta \tau + 2 \tan \left( \frac{\tau}{\bar{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) \text{sec}^2 \left( \frac{\tau}{\bar{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) \left[ \frac{\Delta}{\Delta \tau} \right],
\]

and equation (3.34) becomes

\[
M(\tau, \tau') = \frac{1}{64\pi^2 G_{EdS}^3} \text{sech}^2 \left( \frac{\tau}{\bar{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) \left[ \frac{\Delta^2}{\Delta \tau^2} \right] \Delta^2 \Delta \tau - 2 \tanh \left( \frac{\tau}{\bar{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) \text{sech}^2 \left( \frac{\tau}{\bar{s}_0 \langle N_3^{(1,3)} \rangle^{1/3}} \right) \left[ \frac{\Delta}{\Delta \tau} \right].
\]
$\Delta/\Delta \tau$ and $\Delta^2/\Delta \tau^2$ denote appropriate finite-difference operators. $M(\tau, \tau')$ is now just a $\bar{\mathcal{T}} \times \bar{\mathcal{T}}$ symmetric matrix. We next add to $M(\tau, \tau')$ two $\bar{\mathcal{T}} \times \bar{\mathcal{T}}$ matrices: one implementing the constraint (3.36) and one enforcing the boundary conditions $\nu_j(t_i) = 0$ and $\nu_j(t_f) = 0$. We finally numerically diagonalize the constrained operator $M(\tau, \tau')$ to obtain its eigenvectors $\nu_j(\tau)$ and associated eigenvalues $\mu_j$. We input the value of $\bar{s}_0$ obtained from the best fit of $N_2^{SL}(\tau)$ to $\langle N_2^{SL}(\tau) \rangle$, and we scale each $\mu_j$ by an overall constant, corresponding to the value of the coefficient $1/64\pi^2\hbar G \ell_d^3$, obtained by exactly matching the values of $\mu_1$ and $\lambda_1$.

### 3.C Definition and measurement of the spectral dimension

The spectral dimension, a measure of the dimensionality of a space as experienced by a diffusing random walker, is defined via the heat equation governing this walker’s diffusion. On a Wick-rotated causal triangulation the integrated heat equation takes the form

$$
\mathcal{K}_{\mathcal{T}_c}(s, s', \sigma + 1) = (1 - \varrho)\mathcal{K}_{\mathcal{T}_c}(s, s', \sigma) + \frac{\varrho}{N(\mathcal{N}(s))} \sum_{s'' \in \mathcal{N}(s)} \mathcal{K}_{\mathcal{T}_c}(s'', s', \sigma). \quad (3.66)
$$

The heat kernel $\mathcal{K}_{\mathcal{T}_c}(s, s', \sigma)$ gives the probability of diffusion from $D$-simplex $s$ to $D$-simplex $s'$ (or vice versa) in $\sigma$ diffusion time steps; the diffusion constant $\varrho$ characterizes the dwell probability in a given time step; and $\mathcal{N}(s)$ is the set of $N(\mathcal{N}(s))$ nearest neighbors.
of the $D$-simplex $s$. We set $\varrho = 4/5$. The heat trace or return probability, defined as

$$P_{T_c}(\sigma) = \frac{1}{N_D} \sum_{s \in T_c} K_{T_c}(s, s, \sigma),$$

(3.67)

gives the probability for a random walker to return to its starting $D$-simplex in $\sigma$ diffusion time steps. The spectral dimension $\mathcal{D}_s^{(T_c)}(\sigma)$ quantifies the scaling of the return probability $P_{T_c}(\sigma)$ with diffusion time $\sigma$:

$$\mathcal{D}_s^{(T_c)}(\sigma) = -2 \frac{d \ln P_{T_c}(\sigma)}{d \ln \sigma}.$$  

(3.68)

The definition (3.68) is primarily motivated by the fact that, for diffusion of a random walker on a continuous Riemannian manifold, the spectral dimension at zero diffusion time coincides with this manifold’s topological dimension.

Given an ensemble of causal triangulations representative of those contributing to the partition function (3.7), we numerically estimate the spectral dimension $\mathcal{D}_s(\sigma)$ as follows. The number $N_D$ of $D$-simplices comprising a typical causal triangulation $T_c$ is of order $10^5$, so we estimate the return probability $P_{T_c}(\sigma)$ by considering only a subset of $K$ randomly selected $D$-simplices $s_k$:

$$P_{T_c}^{(K)}(\sigma) = \frac{1}{K} \sum_{s_k \in T_c} K_{T_c}(s_k, s_k, \sigma).$$

(3.69)
One clearly recovers the return probability $\mathcal{P}_{T_c}(\sigma)$ in the limit as $K$ approaches $N_D$:

$$\mathcal{P}_{T_c}(\sigma) = \lim_{K \to N_D} \mathcal{P}_{T_c}^{(K)}(\sigma). \quad (3.70)$$

Since the number $N(T_c)$ of causal triangulations comprising an ensemble is necessarily finite, we estimate the expectation value $\mathbb{E}[\mathcal{P}(\sigma)]$ of the return probability $\mathcal{P}(\sigma)$ by its average over an ensemble:

$$\langle \mathcal{P}(\sigma) \rangle = \frac{1}{N(T_c)} \sum_{j=1}^{N(T_c)} \mathcal{P}_{T_c}^{(j)}(\sigma). \quad (3.71)$$

One clearly recovers the expectation value $\mathbb{E}[\mathcal{P}(\sigma)]$ in the limit as $N(T_c)$ diverges without bound:

$$\mathbb{E}[\mathcal{P}(\sigma)] = \lim_{N(T_c) \to \infty} \langle \mathcal{P}(\sigma) \rangle. \quad (3.72)$$

Taking both of the above estimations into account, we then estimate the return probability $\mathbb{E}[\mathcal{P}(\sigma)]$ as

$$\langle \mathcal{P}^{(K)}(\sigma) \rangle = \frac{1}{N(T_c)} \sum_{j=1}^{N(T_c)} \mathcal{P}_{T_c}^{(K)}(\sigma). \quad (3.73)$$

One clearly recovers the expectation value $\mathbb{E}[\mathcal{P}(\sigma)]$ in the double limit:

$$\mathbb{E}[\mathcal{P}(\sigma)] = \lim_{K \to N_D \atop N(T) \to \infty} \langle \mathcal{P}^{(K)}(\sigma) \rangle. \quad (3.74)$$
We estimate the spectral dimension as

\[ D_s^{(K)}(\sigma) = -2 \frac{d \ln \langle P^{(K)}(\sigma) \rangle}{d \ln \sigma} \]  \hspace{1cm} (3.75)

for an appropriate discretization of the derivative with respect to \( \sigma \).
Part II

Relativistic Astrophysics
Chapters 4 and 5 approach relativistic astrophysics from two different perspectives. Chapter 4 is motivated by the need to numerically solve a broad class of problems in an efficient manner. It therefore develops a novel tool—a discretization method—which has the potential to improve the efficiency of calculations in the future. Chapter 5 contains the first steps necessary for solving one specific problem in relativistic astrophysics. That said, both problems are united by the need to wrestle with the gauge freedom of general relativity.

Although they completely describe the relevant physics, the Einstein equations form an underdetermined system when translated into the language of partial differential equations. To complete them, one must choose a coordinate system—a gauge. Decades of hard experience has taught the relativity community that some gauge choices are better than others and that a good gauge is adapted to the problem one wishes to solve. Without a good choice of coordinates—and a good way of describing them in the language of partial differential equations—there is not even a guarantee that GR forms a well-posed initial value problem. (For a review of this topic, see [192]).

In [160] (chapter 4) I developed, with advice from my supervisor, a novel numerical method which substantially reduces memory and communication overheads for a given accuracy. We took special care to make sure our method is compatible with the formulations of the Einstein equations and the gauge choices popular in numerical relativity.

\[^{10}\]One may also choose a coordinate-free representation, such as the Regge calculus used in the previous chapters or a Tetrad-type formalism. The latter can be thought of as choosing a new gauge at each point on the manifold; the former, as approximating the Lie group of diffeomorphisms as a finite-sized discrete symmetry group.

\[^{11}\]Not all problems are best formulated as initial value problems, of course.
Recall from chapters 2 and 3 that in lattice quantum field theory, one starts with a discrete theory and asks what continuum (low-energy) theory it is compatible with. In [160], we took the opposite stance. We started with a continuum model and rigorously built a discrete approximation which is guaranteed to be compatible with a particular formulation of the Einstein equations. One way we ensure this compatibility is by demanding that certain discrete operators retain the algebraic properties of their continuum counterparts.

Sometimes a problem can be most easily solved by utilizing several different coordinate systems. Such is the case with the problem discussed in chapter 5. In [115], Herdeiro, Radu, and Rúnarsson presented a set of stationary spacetimes representing black holes with massive vector hair. They arrived at these solutions by solving a time-independent problem, which is formulated as an elliptic system of partial differential equations. To do so, they chose a particularly simple, manifestly time-independent coordinate system.

By utilizing numerical tools developed by William East, I am now studying the time-dependent behaviour of perturbations to these hairy black hole solutions. Unfortunately, Herdeiro, Radu, and Rúnarsson’s coordinate system is ill adapted to dynamical simulations. As a necessary first step, I therefore present a new coordinate system better adapted to this problem.
Chapter 4

An Operator-Based Discontinuous
Galerkin Method Compatible with the
BSSN Formulation of the Einstein
Equations

4.1 Context

This work (first presented in [160]) is a collaboration between myself and Erik Schnetter. The work was done between September 2013 and March 2016. The paper was submitted for peer review April 2016 and published November 2016. The original idea and the first prototype code are due to E. Schnetter. As corresponding and first author, I performed
the bulk of the work on this project, including the development of the refined formulation, the write-up, the literature review, the mathematical proofs, and the numerical studies.

4.2 Abstract

Discontinuous Galerkin Finite Element (DGFE) methods offer a mathematically beautiful, computationally efficient, and efficiently parallelizable way to solve partial differential equations (PDEs). These properties make them highly desirable for numerical calculations in relativistic astrophysics and many other fields. The BSSN formulation of the Einstein equations has repeatedly demonstrated its robustness. The formulation is not only stable but allows for puncture-type evolutions of black hole systems. To-date no one has been able to solve the full (3+1)-dimensional BSSN equations using DGFE methods. This is partly because DGFE discretization often occurs at the level of the equations, not the derivative operator, and partly because DGFE methods are traditionally formulated for manifestly flux-conservative systems. By discretizing the derivative operator, we generalize a particular flavor of DGFE methods, Local DG methods, to solve arbitrary second-order hyperbolic equations. Because we discretize at the level of the derivative operator, our method can be interpreted as either a DGFE method or as a finite differences stencil with non-constant coefficients.

4.3 Introduction

In numerical relativity, Einstein’s equations are typically decomposed in one of several ways. Drawing on the constraint damping proposed by Gundlach et al. [110], the generalized harmonic (GH) formulation was originally developed in second-order form by Pretorious [180]. Pretorious used it with finite differences to provide the first successful evolution and merger of a binary black hole system [179]. Lindblom et al. [149] rewrote the generalized harmonic formulation in first-order form. Using pseudospectral methods,
this formulation has successfully been used to accurately describe a wide variety of astrophysical situations. The literature is very extensive, but the interested reader can find much of the relevant work in [51, 193, 136, 134, 121] and references therein.

The Baumgarte-Shapiro-Shibata-Nakamura (BSSN) formulation of the Einstein equations [199, 42, 53, 3, 7] is a second-order formulation. Some key ingredients of the BSSN formulation are the conformal re-scaling of geometric quantities, treating the trace of the connection coefficients as independent variables, and the separation of the trace of the extrinsic curvature tensor from its other components. These ingredients not only make the formulation well-posed [191] but allow for so-called “moving puncture” evolutions, where the singularity within a black hole is not resolved on the computational grid, and where the thus non-physical interior of the black hole can be safely evolved thanks to the characteristic structure of the system [59, 39].\(^1\) In three dimensions, puncture solutions are very desirable because they are significantly easier to implement than the other techniques for avoiding singularities.

For these reasons, the BSSN formulation of the Einstein equations is used by many relativity groups and a great deal of expertise has been acquired. This is strong motivation for the development of efficient numerical methods for evolving the BSSN equations. For smooth problems, such as the Einstein equations, pseudospectral methods converge exponentially. They are therefore a very appealing approach to solving the Ein-

\(^1\)In the BSSN system, there can be gauge modes which travel superluminally and escape the black hole interior. Since they are gauge, this is not unphysical or problematic. However, the high-frequency parts of these modes must be resolved outside the black hole to ensure good convergence. See [55] and [90] for detailed analyses of this difficulty.
stein equations. They have been successfully used with the GH formulation in a number of contexts, especially for compact binary mergers of all flavors [51, 193, 136]. They have also been enormously successful in generating initial data for numerical relativity [103, 102, 172, 35, 202, 201, 2, 138, 60].\(^2\) If one imposes appropriately flux-conservative penalty-type boundary conditions and uses many small spectral domains, these techniques become nodal discontinuous Galerkin finite element (DGFE) methods [100].

DGFE methods combine the high-order accuracy of spectral methods with the flexibility and parallelizability of finite volume type methods [119]. In smooth regions they provide spectral accuracy and in non-smooth regions they can be combined with high-resolution shock capturing (HRSC) techniques to accurately resolve discontinuities. (See [85, 184, 243, 241, 57, 135] for some recent applications of DGFE combined with HRSC for relativistic hydrodynamics.) Importantly, DGFE methods allow for a domain decomposition which requires only a single layer of ghost points.

There are several extensions of DGFE methods for second-order and non-flux-conservative systems. Using distributional theory, Vol’pert [223], LeFloch and collaborators [80, 146, 147, 123], and Colombeau and coworkers [72, 73] have all developed different techniques to define shock wave solutions for hyperbolic systems of the form

\[ \partial_t \psi + g(\psi) \partial_i \psi = 0 \quad \forall \ i \in \{1, 2, 3\}, \]

\(^2\)We note that the calculation of initial data involves solving an elliptic differential system, not a hyperbolic one. This requires different considerations than those we discuss here.
where $\psi$ is a collection of variables, each of which may be discontinuous. These techniques have been applied numerically first in a finite volume context [64, 217, 218, 63, 167] and later in a discontinuous Galerkin setting [187, 211].

In [213], Teukolsky develops a formalism for DGFE methods in arbitrary curved space-times for both conservative and non-conservative first-order systems. In [212], Taylor et al. derive a penalty method for the wave equation based on energy methods. Interior Penalty discontinuous Galerkin (IPDG) methods [36, 197, 65, 108, 109, 118, 107] discretize a second-order system by imposing additional penalty boundary terms. Local discontinuous Galerkin (LDG) methods, developed by Shu and collaborators [69, 237, 238] and based on the early work by Bassi et al. [40, 41], introduce auxiliary variables to facilitate second-differentiation. These variables are evaluated at each time step but not evolved and allow penalty boundaries to be imposed as for a first-order system. For a review, see [239].

DGFE methods exist within a rich ecosystem of penalty methods, including but not limited to: spectral penalty methods [94, 95, 120], spectral finite volume methods [225, 226, 227, 229, 151, 207], and spectral difference methods [150, 228]. Indeed, nodal DGFE methods (and likely pseudospectral penalty methods in general) can be cast as multi-domain summation-by-parts finite differences methods [96, 91]. In this formalism, the penalty boundary terms are called simultaneous approximation terms [61, 208, 92].

In [215] and [216], Tichy evolved a static black hole on a single spectral domain using the BSSN system and a pseudospectral scheme. Using a variational principle, Zumbusch developed a DGFE discretization in both space and time for the second-order GH formula-
tion of the Einstein equations [245]. Field et al. developed a DGFE method for the second-order BSSN equations in spherical symmetry [93]. In [184], Radice and Rezzola developed a DGFE formulation for fluids in a general relativistic setting. In the process, they use a DGFE method to solve the Einstein equations in spherical symmetry with maximal slicing and areal coordinates. In [57], Bugner et al. build on this work to combine DGFE methods with an HRSC scheme based on weighted essentially nonoscillatory (WENO) algorithms for fluids on a fixed, curved, spacetime background. Motivated by the first-order-in-space nature of DGFE methods, Brown et al. developed a fully first-order version of the BSSN system and evolved a binary black hole in-spiral using finite differences. They also evolved a reduction of the system to spherical symmetry using a DGFE scheme [54].

It is desirable to evolve the second-order BSSN equations in full 3+1 dimensions via DGFE methods. Unfortunately, to-date this has not been possible.

DGFE methods are typically formulated for manifestly flux-conservative systems of the form

\[ \partial_t \psi + \partial_j f^j(\psi) = 0, \]  

(4.1)

where \( \psi \) is a collection of variables and \( f \) is a nonlinear flux, which is a function of \( \psi \). In these systems, differentiation of the flux is the most natural operation. Therefore, one may discretize \( \partial_j f^j \) all at once.

\footnote{In this gauge, the reduction of the Einstein equations to spherical symmetry is an elliptic, rather than hyperbolic, system.}
In contrast, the BSSN system is roughly of the form

\[ \partial_t \psi = \mathcal{L} [\psi, \partial_j \psi, \partial_j \partial_k \psi], \tag{4.2} \]

where \( \mathcal{L} \) is a nonlinear operator that acts on \( \psi \) and its derivatives. In this case, the natural operation is to differentiate \( \psi \) directly. Therefore, there is no reason to discretize the entire operator \( \mathcal{L} \), which may be very cumbersome. Instead, it may be cleaner to discretize the differential operators \( \partial_i \).

There are then two related difficulties in evolving the BSSN equation using DGFE methods. First DGFE methods are usually formulated for manifestly flux-conservative systems, which are first-order in space and time. The penalty-type boundary conditions imposed in DGFE methods must therefore be generalized to second-order systems. Second, DGFE methods are usually formulated at the level of the equations, not the level of the derivative operator. Given the complexity of the BSSN system, this is a serious impediment to the development of a working scheme.

In this work we develop a new generalization of DGFE methods, operator-based local discontinuous Galerkin (OLDG) methods, which address these problems and allows us to evolve the BSSN equations. We then subject our approach to a battery of community-developed tests for numerical relativity: the Apples-with-Apples tests [6, 38]. So that we can handle second-order systems, we base our scheme off of LDG methods. We draw particular inspiration from the work of Xing et al. [236], where they develop a superconvergent
energy-conserving LDG method for the wave equation.

To avoid the complications of discretizing the BSSN equations, we perform our discretization at the level of the derivative operator, rather than at the level of the equations. This requires a different formalism for describing the piecewise polynomial space. Our formalism uses distributional theory and is inspired by a pedagogical exercise in [119], which we make rigorous. Because we focus our discretization at the level of the differential operator and not the equations, our method provides a drop-in solution for working relativity codes. All that is necessary to convert a finite differences code to a DGFE code is to replace the derivative operator with ours.

Discretizing at the level of the derivative requires special care with respect to the stability of our scheme. Integration by parts is often an integral step in proofs of the well-posedness of a continuum system of initial-value problems [4]. Essentially, one finds an energy norm and shows that it is non-increasing. Hyperbolic systems for which an energy norm exists are called symmetrizable hyperbolic. In their pioneering work, Sarbach et al. showed that the second-order BSSN system is a second-order version of a first-order strongly hyperbolic system [191]. Strong hyperbolicity is weaker than symmetrizable hyperbolicity, but both imply stability.

The discrete analog of integration by parts is summation by parts, developed by Kreiss and Scherer [142, 143], and it tremendously simplifies proofs of numerical stability. Indeed, a summation-by-parts operator, combined with strong or symmetrizable hyperbolicity and appropriate conditions on initial and boundary data, is often enough to demonstrate
stability.

Given the complicated nature of the BSSN equations and their discretizations, we do not seek to prove the stability of our scheme. Rather we insist that our discretized derivative operator satisfy summation by parts. Given the strongly hyperbolic nature of the BSSN system, we expect this restriction to provide linear stability. Nonlinear stability is enforced both by a truncation scheme we develop and by more traditional filtering techniques as needed.

Our paper is organized as follows. In section 4.4, we develop the formalism for OLDG methods and define the OLDG operator. For brevity, we skip the details of our results regarding summation-by-parts, stability, and convergence. The interested reader can find these in appendices 4.B, 4.C and 4.D respectively. In section 4.5, we describe some of the computational properties of interest, such as computation, communication, and memory access costs. In section 4.6, we describe the numerical tests we perform and their results. Finally, in section 4.7, we offer some concluding remarks.

4.4 Methods

In the usual formulation of DGFE methods one replaces the conserved flux through the boundary of a fixed volume with a numerical flux, which takes information from within the volume and from the boundaries of neighbouring volumes [119]. Here we take a different approach. We use distributional theory to replace the derivative of a smooth function with
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Omega = [X_L, X_R]$</td>
<td>Our domain of interest</td>
</tr>
<tr>
<td>$\tilde{\Omega} = [X_L - \epsilon, X_R + \epsilon]$</td>
<td>Our extended domain</td>
</tr>
<tr>
<td>$\Omega^k = [x^k_L, x^k_R]$</td>
<td>An element within $\Omega$</td>
</tr>
<tr>
<td>$K$</td>
<td>The number of elements in domain $\Omega$</td>
</tr>
<tr>
<td>$P^k$</td>
<td>The maximum polynomial order of the test functions ${\Phi^k}$ used to represent a function within element $\Omega^k$</td>
</tr>
<tr>
<td>$h^k$</td>
<td>The width of an element, $x^k_R - x^k_L$</td>
</tr>
<tr>
<td>$\psi, \phi$</td>
<td>Piecewise smooth functions living on our broken domain</td>
</tr>
<tr>
<td>$X^k$</td>
<td>The characteristic function for $\Omega^k$</td>
</tr>
<tr>
<td>$\Theta^k_L, \Theta^k_R$</td>
<td>The Heaviside function centred on $x^k_L$ and $x^k_R$ respectively</td>
</tr>
<tr>
<td>$\delta^k_L, \delta^k_R$</td>
<td>The Dirac delta function centred on $x^k_L$ and $x^k_R$ respectively</td>
</tr>
<tr>
<td>$\psi^k$</td>
<td>The restriction of $\psi$ onto $\Omega^k$</td>
</tr>
<tr>
<td>$\psi^k_i, \psi^k_r$</td>
<td>$\psi^k$ evaluated at $x^k_L$ and $x^k_R$ respectively</td>
</tr>
<tr>
<td>$\Phi^k_i$</td>
<td>The $i^{th}$ test function in $\Omega^k$</td>
</tr>
<tr>
<td>$(\psi^<em>)^k_i, (\psi^</em>)^k_r$</td>
<td>Weak boundary terms</td>
</tr>
<tr>
<td>$\tilde{\psi}_s^k, \tilde{\psi}_d^k$</td>
<td>The smooth and discontinuous extensions of $\psi^k$ outside its element respectively</td>
</tr>
<tr>
<td>$x^k_i$</td>
<td>The $i^{th}$ collocation point ($0 \leq i \leq P^k$) on discretized element $\Omega^k$</td>
</tr>
<tr>
<td>$w^k_i$</td>
<td>Weights for the discrete inner product within an element</td>
</tr>
<tr>
<td>$\hat{\psi}^k_i$</td>
<td>$\psi^k$ evaluated at $x^k_i$</td>
</tr>
<tr>
<td>$\hat{\psi}^k_i$</td>
<td>The projection of $\psi^k$ onto $\Phi^k_i$</td>
</tr>
<tr>
<td>$\hat{\varphi}^k$</td>
<td>The change of basis matrix that maps the spectral coefficients for $\psi^k$ to those for its derivative</td>
</tr>
<tr>
<td>$V^k$</td>
<td>The Vandermonde matrix, which transforms between the modal and nodal bases</td>
</tr>
<tr>
<td>$d^k$</td>
<td>The narrow nodal derivative operator</td>
</tr>
<tr>
<td>$w^k$</td>
<td>The discrete weights as a matrix</td>
</tr>
<tr>
<td>$b^k$</td>
<td>The boundary operator, the discretization of a Dirac delta function over the boundary of an element $\Omega^k$</td>
</tr>
<tr>
<td>$F^k$</td>
<td>The “fetch” operator, which pulls information from elements neighbouring element $\Omega^k$</td>
</tr>
<tr>
<td>$D^k$</td>
<td>The wide derivative operator, which takes neighbouring elements into account</td>
</tr>
</tbody>
</table>

Table 4.1: Notation used in the construction of our discontinuous Galerkin scheme. All symbols are explained in more detail in the main text.
the weak derivative of a piecewise smooth function, appropriately chosen to recover the small communication overhead characteristic of these types of methods.

This approach was first proposed by Hesthaven and Warburton [119], who use it pedagogically to argue that the strong form of the canonical DGFE operator is just an encoding of the notion of a weak derivative. We make this assertion rigorous and argue that this weak derivative formulation provides a generic way to place arbitrary nonlinear hyperbolic equations into the DGFE framework, even if they are not manifestly flux-conservative.

This focus on the derivative operator has very practical consequences: It allows computer programmes that currently employ finite differences methods to replace the finite differences stencil with our OLDG stencil, converting a finite differences code to a DGFE code. This transition requires: implementing loop tiling (for efficiency) so that the band-diagonal finite differences operator can be replaced by our block-diagonal discontinuous Galerkin operator, implementing the truncation scheme described in section 4.4.9, and replacing the Kreiss-Oliger dissipation operator with the right-hand-side filter operator discussed in section 4.4.12. We discuss those issues in this section below. This change from a finite difference to a DGFE method should improve the parallel efficiency significantly, as we discuss in the next section 4.5.

In a DGFE method, there are two levels of discretization. At the top level, one breaks the domain of interest $\Omega$ into many subdomains, or elements, $\{\Omega_k\}_{k=1}^K$ which overlap on a set of measure zero (see figure 4.1). One must choose how to approximate the derivatives of a function that lives on this broken domain. At a lower level, one must choose how to
approximate the piece of the global function that lives on each subdomain. From the perspective of a single element, one can think of the former as a choice of boundary conditions for the piece of the function living on the element and the latter as an ansatz for the types of functions that can live on the element. These two levels of discretization can be lumped into a single discretization step. However, making them explicit allows us to develop our discretization approach formally.

For the reader’s convenience we provide a reference for the notation used in our construction in table 4.1.

### 4.4.1 The Main Idea

Before we proceed with our construction, we present a toy problem which encapsulates some of the core ideas of our method. Consider two smooth functions:

1. \( \phi : [0, 1] \rightarrow \mathbb{R} \) \hspace{1cm} (4.3)
2. \( \pi : [0, 1] \rightarrow \mathbb{R} \) \hspace{1cm} (4.4)

each of which is defined on the interval \([0, 1]\). From these two functions, we can construct a third function \( \psi : [0, 1] \rightarrow \mathbb{R} \) defined by

\[
\psi(x) = \begin{cases} 
\phi(x) & \text{if } 0 \leq x \leq x_0 \\
\pi(x) & \text{if } x_0 < x \leq 1 
\end{cases}
\] \hspace{1cm} (4.5)
for some $0 < x_0 < 1$.

We wish to differentiate $\psi$. However, generically, $\psi$ has a discontinuity at $x_0$ and is not a differentiable function. If we treat $\psi$ as a distribution then its derivative can be defined weakly. To make this manifest, we write $\psi$ as the sum of two distributions:

$$
\psi(x) = \phi(x)\Theta(x_0 - x) + \pi(x)\Theta(x - x_0),
$$

(4.6)

where

$$
\Theta(x) = \begin{cases} 
0 & \text{if } x < 0 \\
1 & \text{if } x \geq 0 
\end{cases}
$$

(4.7)

is the Heaviside function. Then the weak, or distributional derivative of $\psi$ is given by

$$
\partial_x \psi(x) = [\partial_x \phi(x)] \Theta(x_0 - x) - \phi(x)\delta(x - x_0)
$$

$$
+ [\partial_x \pi(x)] \Theta(x - x_0) + \pi(x)\delta(x - x_0)
$$

$$
= [\partial_x \phi(x)] \Theta(x_0 - x) + [\partial_x \pi(x)] \Theta(x - x_0)
$$

$$
+ [\pi(x) - \phi(x)] \delta(x - x_0),
$$

(4.8)

where we have used the fact that the distributional derivative of the Heaviside function $\Theta(x)$ is the Dirac delta function $\delta(x)$ [84].

Equation (4.8) is well defined under integration with any smooth test function that has
compact support over the interval $[0, 1]$. In other words,

$$
\int_0^1 \Phi \partial_x \psi dx = \int_0^{x_0} \Phi \partial_x \phi dx + \int_{x_0}^1 \Phi \partial_x \pi dx + [\pi - \phi] \Phi \bigg|_{x_0}
$$

for all smooth functions $\Phi$ such that $\Phi(0) = \Phi(1) = 0$.

In the following sections, we will use a decomposition much like that given by equation (4.6). In this toy example, our decomposition relied on the existence of functions $\phi$ and $\pi$ which were defined on the entire interval $[0, 1]$. More generally, these functions may not be given to us a priori, but they can be constructed.

### 4.4.2 The Broken Domain

We now proceed with the main construction. For simplicity suppose that the domain of interest $\Omega$ is the real interval $[X_l, X_r], X_l < X_r \in \mathbb{R}$. For higher dimensions, we simply assume a Cartesian product topology. We break our domain $\Omega$ into $K$ interior elements

$$
\Omega^k := [x_l^k, x_r^k], \quad x_l^k < x_r^k \in \mathbb{R}
$$

(4.9)

for all $k = 1, \ldots, K$ and two boundary elements

$$
\Omega^0 := \{x_r^0\}, \quad \Omega^{K+1} := \{x_l^{K+1}\}
$$

(4.10)
Figure 4.1: A broken domain with $K = 3$.

with $x_r^0 = X_l$ and $x_l^{K+1} = X_r$ such that,

$$x_r^{k-1} = x_l^k \text{ and } x_r^k = x_l^{k+1}$$  \hspace{1cm} (4.11)

for all $k = 1, \ldots, K$. (Note that these boundary elements are singleton sets.) We also demand that the union of all $K + 2$ elements comprises the whole domain:

$$\bigcup_{k=0}^{K+1} \Omega^k = \Omega.$$  \hspace{1cm} (4.12)

For convenience, we define the element width

$$h^k = x_r^k - x_l^k.$$  \hspace{1cm} (4.13)

Figure 4.1 shows the structure of $\Omega$ for three elements, e.g., when $K = 3$.

On our domain $\Omega$, we wish to represent the arbitrary function $\psi(x)$. $\psi$ may be either a scalar or vector quantity. For simplicity, we will assume here that it is a scalar. We call the
restriction of \( \psi \) onto a given element \( \Omega^k \), \( \psi^k \). We demand that each \( \psi^k \) be smooth but we allow \( \psi \) to have jump discontinuities at the domain boundaries \( \{ x^k_r = x^{k+1}_l \} \). Notice that in this description, each \( \Omega^k \) overlaps with its neighbours \( \Omega^{k-1} \) and \( \Omega^{k+1} \) at exactly one point. These points of overlap are exactly the points where \( \psi^k \) is allowed to be discontinuous and

\[
\psi^k_r := \psi^k(x^k_r) \quad \text{and} \quad \psi^{k+1}_l := \psi^{k+1}(x^{k+1}_l) \tag{4.14}
\]

can be thought of as the left- and right-hand limits of \( \psi \) respectively. These conditions are equivalent to the standard choices one makes for a typical one-dimensional DGFE method.

For clarity, we define the following convention. Functions living on the domain \( \Omega \) will be represented by lower-case Greek letters. Elements will be indexed by a superscript, and positions within an element will be denoted by a subscript.

We can formalize the restriction of \( \psi \) to \( \psi^k \) by defining the characteristic function

\[
\chi^k(x) := \begin{cases} 
1 & \text{if } x \in \Omega^k \\
0 & \text{else}
\end{cases} \tag{4.15}
\]

such that

\[
\psi(x) \chi^k(x) = \begin{cases} 
\psi^k(x) & \text{if } x \in \Omega^k \\
0 & \text{else}
\end{cases} \tag{4.16}
\]

Note that product (4.16) is defined pointwise as a function. However, since the product of two distributions is in general not well-defined, it is not a proper distribution. The tech-
niques Colombeau and coworkers [72] can be used to define such a distribution. However, we will not need to do so.

We also introduce two inner products, one local to a subdomain $\Omega^k$ and one for the entire domain. If $\psi^k$ and $\phi^k$ are functions on $\Omega^k$, then the subdomain inner product is

$$\langle \psi^k, \phi^k \rangle_{\Omega^k} = \int_{x_l^k}^{x_r^k} \psi^k(x) w^k(x) \phi^k(x) dx,$$  \hspace{1cm} (4.17)

where $w^k(x)$ is an as-of-yet undecided weight function. If $\psi$ and $\phi$ are instead functions on the whole domain $\Omega$, then the overall inner product is:

$$\langle \psi, \phi \rangle_{\Omega} = \int_{X_l}^{X_r} \psi(x) w(x) \phi(x) dx = \sum_{k=1}^{K} \langle \psi^k, \phi^k \rangle_{\Omega^k},$$  \hspace{1cm} (4.18)

where $w(x)$ is the weight function for the whole domain and $\psi^k$ and $\phi^k$ are the appropriate restrictions of $\psi$ and $\phi$.

We will sometimes be interested in a slightly extended integral. Let

$$\varepsilon > 0, \varepsilon \in \mathbb{R}.$$  \hspace{1cm} (4.19)

Then we define the extended domain

$$\tilde{\Omega} := [X_l - \varepsilon, X_r + \varepsilon]$$  \hspace{1cm} (4.20)
and extended inner product

\[ \langle \psi, \phi \rangle_{\tilde{\Omega}} := \int_{X_l-\varepsilon}^{X_r+\varepsilon} \psi(x)w(x)\phi(x)dx. \] (4.21)

This extension is useful for handling the discontinuities at \( \Omega^0 \) and \( \Omega^{K+1} \).

Since we allow functions on \( \Omega \) to be piecewise smooth and therefore not everywhere differentiable, we are interested in their properties in a weak or distributional sense. Therefore for each \( k = 1, \ldots, K \) and each element \( \Omega^k \), we define a set of smooth test functions, \( \{ \Phi^k_i \}_{i=0}^{\infty} \), each defined on the whole extended interval \( \tilde{\Omega} = [X_l - \varepsilon, X_r + \varepsilon] \).
We demand that, for all \( i \in \mathbb{N} \) and \( k = 1, \ldots, K \),

\[
\Phi^k_i(X_l - \varepsilon) = \Phi^k_i(X_r + \varepsilon) = 0
\] (4.22)

and that, for all \( \phi^k \in C^\infty(\Omega^k) \), there exists a set of spectral coefficients \( \{ \hat{\phi}^k_i \}_{i=0}^\infty \), \( \hat{\phi}^k_i \in \mathbb{R} \), such that

\[
\phi^k(x) = \sum_{i=0}^{\infty} \hat{\phi}^k_i \chi^k \Phi^k_i(x).
\] (4.23)

In other words, every \( \Phi^k_i \) has compact support on the interval \([X_l - \varepsilon, X_r + \varepsilon]\) and, for each \( k \), the set of restrictions \( \{ \chi^k \Phi^k_i \}_{i=0}^\infty \) forms an orthonormal basis for \( C^\infty(\Omega^k) \). Other than insisting on smoothness and compact support, we do not need to constrain our test functions outside of their respective elements. (From now on, we will represent spectral coefficients with hats.)

The choice of weight function \( w^k \) is tied to the choice of basis functions \( \Phi^k_i \). We may use it to ensure that our test functions are orthogonal. The weight function for the whole domain, \( w(x) \) must be chosen for compatibility with \( w^k \).

Note that our basis functions are quite different from those defined in a standard discontinuous Galerkin method. In the usual DGFE approach one can define the test functions as a set of piecewise smooth functions on \( \Omega \) with discontinuities at element boundaries. However, if the test functions themselves are discontinuous, we cannot rigorously apply distribution theory.

In our case we define a set of test functions on each element. However, each test
function is defined outside the element, with compact support on an appropriately extended domain, as shown in figure 4.2. This construction ensures that all our test functions are smooth and that the standard results from distributional theory hold.

In particular, if $\Theta(x)$ and $\delta(x)$ are the Heaviside and Dirac delta functions respectively, then

$$\left\langle \partial_x \Theta, \Phi^k_i \right\rangle_\Omega = \left\langle \delta, \Phi^k_i \right\rangle_\Omega$$

(4.24)

for all $i \in \mathbb{N}$ and $k = 1, 2, \ldots, K$ [84].

For convenience, we define the *shifted* Heaviside and Dirac delta functions

$$\Theta^k_i(x) := \Theta(x_i^k - x), \quad \Theta^k_r(x) := \Theta(x - x_r^k), \quad (4.25)$$

$$\delta^k_i(x) := \delta(x - x_i^k), \quad \text{and} \quad \delta^k_r(x) := \delta(x - x_r^k), \quad (4.26)$$

which are centred on the element boundaries $\{x_i^k, x_r^k\}_{k=1}^K$. (Note that $\Theta^k_i$ is inverted in $x$ so it is nonzero for $x < x_i^k$ and zero for $x > x_i^k$.) In this language, we can write the distributional derivative of the characteristic function as

$$\left\langle \partial_x \chi^j, \Phi^k_i \right\rangle_\Omega = \left\langle \delta^j_i, \Phi^k_i \right\rangle_\Omega - \left\langle \delta^j_r, \Phi^k_i \right\rangle_\Omega$$

(4.27)

for all $i \in \mathbb{N}$ and all $j, k = 1, 2, \ldots, K$.

For the remainder of this paper, the restrictions $\chi^k \Phi^k_i$ of our test functions $\Phi^k_i$ to the interior of an element $\Omega^k$ are assumed to be the Legendre polynomials, defined in appendix
4.A. (It doesn’t matter what they are outside $\Omega^k$, as long as they are smooth and have compact support.) This means our element-wise weight function is $w^k = 1$.

We make this assumption partly to make contact with traditional nodal DGFE methods, where this is the norm, and partly for simplicity. Although we did not explore them, other choices such as Chebyshev polynomials are certainly possible and potentially desirable. Choice of (orthonormal) basis function and associated collocation points affect both the timestep and the conditioning of the linear operator that transforms between nodal and modal bases [119] (see section 4.4.4). For more details on some of the effects varying the basis or collocation points can have, see [119] or [177].

As long as the collocation points include the boundaries of the element, as described in appendix 4.A, our main results are unchanged by the choice of basis or collocation points.

### 4.4.3 The Modified Derivative Operator

To take the derivative of an arbitrary distribution $\phi$ defined on $\Omega$, we will decompose it into the sum of several distributions, just as in section 4.4.1. However, to follow this procedure, each $\phi^k$ must be smoothly defined on the whole extended domain $\tilde{\Omega}$. Therefore, for each $\psi^k$, we define a smooth continuation

\[
\tilde{\psi}^k \in C^\infty(\tilde{\Omega}) : \tilde{\Omega} \rightarrow \mathbb{R}
\] (4.28)
such that
\[ \chi^k \tilde{\psi}^k_s = \chi^k \psi. \] (4.29)

In other words, within \( \Omega^k \), \( \psi \) and \( \tilde{\psi}^k_s \) must agree. However, outside of \( \Omega^k \) they generically do not.

Armed with this machinery, we can perform a procedure analogous to that in section 4.4.1 to compute the derivative of a piecewise smooth function \( \phi \) on \( \Omega \). To this end, consider the following discontinuous continuation of the restriction of \( \psi^k \): \[ \tilde{\psi}^k_d(x) := \tilde{\psi}^k_s(x) \chi^k(x) + (\psi^*)_l^k \Theta^k_l(x) + (\psi^*)_r^k \Theta^k_r(x), \] (4.30)

where \((\psi^*)_l^k\) and \((\psi^*)_r^k\) are chosen to incorporate information about \( \psi^k_{r-1} \) and \( \psi^k_{r+1} \) respectively. Note that since \( \tilde{\psi}^k_s \) is smooth, \( \tilde{\psi}^k_d \) is a proper distribution with no ambiguities other than the choice of \( \psi^* \), which is analogous to the choice of numerical flux in a traditional DGFE scheme.

We now approximate the derivative of \( \psi \) as the weak derivative of \( \tilde{\psi}^k_d(x) \) for \( x \in \Omega^k \). To calculate this weak derivative, we differentiate and take the inner product with an arbitrary
test function $\Phi^k_i$:

\[
\langle \Phi^k_i, \partial_x \tilde{\psi}_d^k \rangle_{\Omega} = \langle \Phi^k_i, \partial_x \left[ \chi^k(x) \tilde{\psi}_d^k(x) \right] \rangle_{\Omega} \\
+ \langle \psi^* \rangle^k_l \langle \Phi^k_i, \partial_x \Theta^k_i(x) \rangle_{\Omega} + \langle \psi^* \rangle^k_r \langle \Phi^k_i, \partial_x \Theta^k_r(x) \rangle_{\Omega} \\
= \langle \Phi^k_i, \chi^k \partial_x \tilde{\psi}_d^k \rangle_{\Omega} + \langle \Phi^k_i, \tilde{\psi}_d^k \partial_x \chi^k \rangle_{\Omega} \\
+ \langle \psi^* \rangle^k_l \langle \Phi^k_i, \partial_x \Theta^k_i(x) \rangle_{\Omega} + \langle \psi^* \rangle^k_r \langle \Phi^k_i, \partial_x \Theta^k_r(x) \rangle_{\Omega} \\
= \langle \Phi^k_i, \partial_x \psi^k \rangle_{\Omega} - \left[ \langle \psi^* \rangle^k_l - \psi^k_l \right] \langle \Phi^k_i, \delta^k_l \rangle_{\Omega} + \left[ \langle \psi^* \rangle^k_r - \psi^k_r \right] \langle \Phi^k_i, \delta^k_r \rangle_{\Omega} \\
= \langle \Phi^k_i, \partial_x \psi^k \rangle_{\Omega} + \Phi^k_i(x) \left[ \langle \psi^* \rangle^k - \psi^k \right] \bigg|_{x_d^k} \\
= \langle \Phi^k_i, \partial_x \psi^k \rangle_{\Omega} + \langle \Phi^k_i, (\psi^*)^k - \psi^k \rangle_{\partial \Omega^k},
\]

(4.31)

where in the first step we employ the product rule; in the second, we utilize equations (4.15), (4.27), and (4.24); in the third, we utilize the definition of the Dirac delta function; and in the final step, we recognize $\langle \phi, \psi \rangle_{\partial \Omega^k}$ as the integral over the boundary of $\Omega^k$, defined in the usual way.

Note that, although $\partial_x \tilde{\psi}_d^k$ is well-defined as distribution, the product of distributions $g(\psi)\partial_x \tilde{\psi}_d^k$, for an arbitrary nonlinear function $g$, may not be. Therefore, although we have a distributional derivative, we may not be able to use it to weakly define a system of equations. This difficulty can be overcome on a system-by-system basis via the work of, e.g., LeFloch et al. [80, 146, 147, 123] or Colombeau et al. [72, 73]. We are more interested in a general framework which may be used with any sufficiently well-behaved hyperbolic system. Therefore we do not address this issue here.
4.4.4 Moving to the Discrete

Our description of the derivative so far assumes that the restriction $\psi^k$ is an arbitrary smooth function. Therefore, to obtain a discrete scheme, we must make an ansatz about $\psi^k$. We choose a pseudospectral ansatz. We briefly discuss some of the details of this ansatz in appendix 4.A. For a review of pseudospectral methods, see [104].

To make our method a Galerkin method, for each element $\Omega^k$, we choose some $P^k \in \mathbb{N}$ and some subset $\{\Phi_i^k\}_{i=0}^{P^k}$ of the test functions $\Phi_i^k$ and demand that any function $\psi^k$ is a linear combination of the restriction of those test functions onto $\Omega^k$:

$$\psi^k(x) = \sum_{i=0}^{P^k} \hat{\psi}_i^k \Phi_i^k(x) \quad \forall x \in \Omega^k,$$

where $\hat{\psi}_i^k \in \mathbb{R}$ are spectral coefficients. Note that since we have chosen our test functions to be the Legendre polynomials, $P^k$ is also the highest-order polynomial which can be represented within an element.

Then, we demand that the weak derivative relations (4.24) and (4.31) only hold for all $0 \leq i \leq P^k$, rather than for all $k \in \mathbb{N}$. In this modal representation, we can relate $\psi^k(x)$ to its derivative $\partial_x \psi^k(x)$ in the standard way,

$$\partial_x \psi^k(x) = \sum_{i=0}^{P^k} \left( \partial_x \hat{\psi}_i^k \right) \Phi_i^k(x),$$
where the spectral coefficients are given by

$$\left( \partial_x \psi \right)_i^k = \sum_{j=0}^{P_k} \left\langle \partial_x \Phi_i^k(x), \Phi_j^k(x) \right\rangle_{\Omega_k} \hat{\psi}^k_j,$$

which can be thought of as a change-of-basis operation. For convenience, we define the matrix $\hat{d}^k$, whose components are given by

$$\hat{d}^k_{ij} := \left\langle \partial_x \Phi_i^k(x), \Phi_j^k(x) \right\rangle_{\Omega_k}.$$  

We also construct a nodal representation. Within each element, we assume a discrete set of points $\{x_{i}^{k}\}_{i=0}^{P_k}$ such that $x_{0}^{k} = x_{l}^{k}$ and $x_{P_k}^{k} = x_{r}^{k}$. One good choice for $\{x_{i}^{k}\}_{i=0}^{P_k}$ is the Gauss-Lobatto points of $\{\Phi_{i}^{k}\}_{k=0}^{P_k}$. Given a set of values $\psi_{i}^{k}$, $\psi^{k}(x)$ is assumed to interpolate the values such that $\hat{\psi}^{k}(x_{i}^{k}) = \psi_{i}^{k}$. In this nodal representation, the element-wise inner product is approximated by the Gauss-Lobatto quadrature rule

$$\left\langle \phi_i^{k}, \psi_i^{k} \right\rangle_{\Omega_k} \approx \sum_{i=0}^{P_k} \phi_i^{k} \psi_i^{k} w_i^{k},$$

where $w_i^{k}$ are the discrete weights of the inner product.

We emphasize that even when the continuum weights are trivial, as they are for the Legendre polynomials, the discrete weights $w_i^{k}$ will not be. Given a set of test functions and continuum weights, they are determined by the locations of collocation points, as described in appendix 4.A. Here we show the approximate values of the weights and locations of the
collocation points for an element of order $P^k = 4$ and weight $h^k$.

\[
\begin{pmatrix}
0.1 \\
0.54 \\
0.71 \\
0.54 \\
0.1 \\
\end{pmatrix}
\approx \frac{h^k}{2}
\text{ and }
\begin{pmatrix}
-x^k_l + x^k_r \\
\frac{h^k}{2} + \frac{h^k}{2} \\
0 \\
0.65 \\
1 \\
\end{pmatrix}
\approx x^k_l + x^k_r + \frac{h^k}{2}
\] (4.37)

If we collect the $\phi^k_i$ and $\psi^k_i$ into vectors $\phi^k$ and $\psi^k$ respectively, we can define the quadrature rule as a matrix operation:

\[
\langle \phi^k, \psi^k \rangle_{\Omega^k} = (\Phi^k)^T w^k \psi^k,
\] (4.38)

where $w^k_{ij} = w^k_i \delta_{ij}$. The inner product over the whole domain remains the sum over the element-wise inner products as in equation (4.18).

To move between the modal and nodal representations, we introduce the generalized Vandermonde matrix,

\[
\nabla^k_{ij} := \Phi^k_j (x^k_i),
\] (4.39)

such that the vector of nodal coefficients is obtained by applying the Vandermonde matrix to the vector of modal coefficients:

\[
\psi^k = \nabla^k \hat{\psi}^k.
\] (4.40)
This also gives us a nodal representation of element-wise differentiation. We define the element-wise derivative operator as

\[ d^k := \mathcal{V}^k \hat{d} (\mathcal{V}^k)^{-1}, \]  

which is nothing more than the standard pseudospectral derivative operator for the domain \( \Omega^k \). For reasons that will soon become clear, we call \( d^k \) the \textit{narrow} derivative operator.

At this point, we must make a choice about how to represent an inner product that integrates over the boundary of a domain. As a guiding principle, we will use the discrete analogue of integration by parts, summation by parts, for the inner product within a single element. In other words, we want the following to hold:

\[ \langle \phi_k, \partial_x \psi_k \rangle_{\Omega^k} + \langle \partial_x \phi_k, \psi_k \rangle_{\Omega^k} = \langle \phi_k, \psi_k \rangle_{\partial \Omega^k}. \]  

In the discrete case, we can write an inner product over the boundary of an element as a standard inner product where one of the operands is multiplied by a special matrix, \( b^k \), which we call the \textit{boundary operator}:

\[ \langle \phi_k^k, \psi_k \rangle_{\partial \Omega^k} = (\phi_k^k)^T w^k b^k \psi_k = \left( b^k \phi_k^k \right)^T w^k \psi_k. \]
The boundary operator is determined by equations (4.38) and (4.42) to be given by

\[
w^k b^k = w^k d^k + \left(w^k d^k\right)^T, \tag{4.44}
\]

where \(w^k\) and \(d^k\) are the element-wise weight and differentiation matrices respectively.

The matrix \(b^k\) is essentially a discretization of the Dirac delta function centred on the boundary of \(\Omega^k\). It has components given by

\[
b^k_{ij} = \frac{2b^k_0}{h^k} \left(\delta_{ij,P^k} \delta_{j,P^k} - \delta_{i,0} \delta_{j,0}\right), \tag{4.45}
\]

where \(b^k_0 > 0 \in \mathbb{R}\) depends on \(P^k\). The product \(w^k b^k\) has the particularly simple form

\[
\left(w^k b^k\right)_{ij} = \delta_{ij,P^k} \delta_{j,P^k} - \delta_{i,0} \delta_{j,0} \tag{4.46}
\]

so that the integral over the boundaries of an element matches the continuum result,

\[
\langle \phi^k, \psi^k \rangle_{\partial \Omega^k} = \phi^k_r \psi^k_r - \phi^k_l \psi^k_l.
\]
This matrix

\[
\begin{pmatrix}
-20 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 20 \\
\end{pmatrix}
\]

provides an example of \(b^k\) in matrix form for an element of width \(h^k\) and order \(P^k = 4\).

We now set about the task of defining the discrete analog of equation (4.31) which, in our scheme, will replace the pseudospectral derivative (4.41). We must first choose a definition of \(\psi^*\), the quantity in equation (4.30) that represents the information we pull from a neighbouring element. There are a number of choices one can make. However, we make the following, relatively simple, choice. Let the suggestively named \((b^k)^{-1}F^k\) be any operator that maps

\[
\psi^k_l \rightarrow \psi^{k-1}_r \quad (4.48)
\]

and \(\psi^k_r \rightarrow \psi^{k+1}_l\). \( (4.49)\)

Note that the name \((b^k)^{-1}F^k\) is an abuse of notation, as \(b^k\) has no inverse. We then define the \textit{fetch operator}

\[
F^k := b^k \left[ (b^k)^{-1}F^k \right], \quad (4.50)
\]

which performs the same role as \((b^k)^{-1}F^k\), but selects only boundary terms for integration,
making any non-boundary properties of \((b^k)^{-1}F^k\) irrelevant.

Like the boundary operator \(b^k\), the fetch operator \(F^k\) is essentially a discretization of a delta function centred at the boundary of \(\Omega^k\). However, since functions are allowed to be discontinuous at element boundaries, there is an ambiguity. The boundary operator selects for the values of a function within an element, the “inside limit,” while the fetch operator selects for values of a function outside an element, the “outside limit.” If we introduce the shorthand

\[
x_r^{k-1} = x_{-1}^k \text{ and } x_i^{k+1} = x_{P^k+1},
\]

then the fetch operator has components

\[
F_{ij}^k = \frac{2b_0^k}{h_k} \left( \delta_{i,P^k} \delta_{j,P^k+1} - \delta_{i,0} \delta_{j,-1} \right)
\]

(4.52)

for all \(i = 0, 1, \ldots, P^k\) and \(j = -1, 0, \ldots, P^k, P^k + 1\).

Now let

\[
(\psi^*)^k = \frac{1}{2} \xi \left[ (b^k)^{-1}F^k \right] \psi^k - \frac{1}{2} \psi^k
\]

(4.53)

so that

\[
b^k(\psi^*)^k = \frac{1}{2} \left[ \xi F^k \psi^k - b^k \psi^k \right],
\]

(4.54)

where \(\xi \in \mathbb{R}\). Combining this choice of \((\psi^*)^k\) with equation (4.31) results in the following
discrete element-wise, *weak* derivative operator:

$$\partial_x \tilde{\psi}^k = D^k \psi^k := \left[ d^k - \frac{1}{2} b^k + \frac{1}{2} \xi F^k \right] \psi^k,$$

(4.55)

where we call $D^k$ the element-wise *wide derivative* operator because it takes information from neighbouring elements. $D^k$ is the differential operator for OLDG methods.

In general we define the following convention. Any operator that takes information from a *single* element we call *narrow*. Any operator that takes information from an element and its nearest neighbours, we call *wide*. Narrow operators will be represented by lower-case Latin symbols while wide operators will be represented by upper-case Latin symbols. In the functional notation of equations (4.48) and (4.49), the addition of wide and narrow operators is unambiguous. In matrix notation, we simply pad the narrow operator with columns of zeros so that the matrices are the same shape.

### 4.4.5 Summation By Parts

We insist that our operator $D^k$ from equation (4.55) satisfy summation by parts. Given the strongly hyperbolic nature of the BSSN system [191], we expect this restriction to provide linear stability.

By construction, summation-by-parts is satisfied within each element $\Omega^k$. However, stability proofs require integration over the *whole domain* $\Omega$. We define the wide derivative
operator over the whole domain $D$ such that, for all $\psi$ on $\Omega$,

\[ D\psi(x) = (D^k\psi^k)(x) \quad \forall \ x \in \Omega^k \ \forall \ 1 \leq k \leq K. \tag{4.56} \]

We then seek a value of $\xi$, defined in equation (4.55), such that, for all $\psi, \phi$,

\[ \langle \psi, D\phi \rangle_{\Omega} + \langle D\psi, \phi \rangle_{\Omega} = \langle \psi, \phi \rangle_{\partial\Omega}, \tag{4.57} \]

where

\[ \langle \psi, \phi \rangle_{\partial\Omega} = \psi\phi \bigg|_{x=X_r} - \psi\phi \bigg|_{x=X_l}. \tag{4.58} \]

Note that there are two collocation points at the physical position $X_l$: $x^0_r$ and $x^1_l$, and similarly for $X_r$. Therefore, we can reasonably expect summation by parts to average over these two values in some way. For example, we might accept the relationship

\[ \langle \psi, \phi \rangle_{\partial\Omega} = \frac{1}{2} \left[ (\psi^K_r \phi^{K+1}_l + \psi^{K+1}_r \phi^K_l) - \left( \psi^1_r \phi^0_l + \psi^0_r \phi^1_l \right) \right], \]

which is the mixed average of left- and right-hand limits of $\phi$ and $\psi$ at the boundary. (This combination is arbitrary. Other relationships might also be acceptable.)

We find that the only value of $\xi$ that satisfies summation by parts is $\xi = 1$. (See appendix 4.B for a proof.) Therefore, the final version of the derivative operator is

\[ D^k = d^k - \frac{1}{2} \left[ b^k - F^k \right]. \tag{4.59} \]
Note that $\xi = 1$ may not be the only stable choice. Depending on the continuum differential system, other values may result in a scheme that is dissipative at element boundaries so that the energy norm is non-increasing. Indeed, dissipation at element boundaries is typical of DGFE schemes [67]. However, we did not explore this possibility.

In appendix 4.C, we provide an example of how summation-by-parts can be used to demonstrate the stability of an OLDG discretization of the wave equation in second-order form.

### 4.4.6 Properties

Given a set of collocation points $x^k_i$, $D^k$ maps the $P^k + 3$ points $\{x^{k-1}_i\} \cup \{x^k_i\}_{i=0}^{P^k} \cup \{x^{k+1}_i\}$ to the $P^k + 1$ points $\{x^k_i\}_{i=0}^{P^k}$. Physically, it maps a function on the exterior faces and interior of an element $\Omega_k$ to a function defined only on the interior. In this picture the OLDG stencil can be thought of as a finite differences stencil with non-constant coefficients and a special, weak, boundary operator. If we use our wide derivative to discretize the linear wave equation in first-order form, we recover a standard DGFE method in the strong formulation with a simple central flux. We demonstrate this in appendix 4.E. Therefore OLDG methods are truly a generalization of current DGFE methods.

Figures 4.3, 4.4, and 4.5 show examples of $D^k$ for elements of width $h^k$ and order $P^k = 2, 3,$ and $4$ respectively. Here we use the shorthand described in equation (4.51) such that $i$ ranges from $0$ to $P^k$ and $j$ ranges from $-1$ to $P^k + 1$. The first and last columns pull information from the neighbouring elements. Because of space constraints, with the excep-
\begin{equation}
D^k = \frac{1}{h^k} \left(\begin{array}{cccc}
-3 & 0 & 4 & -1 \\
0 & -1 & 0 & 1 \\
0 & 1 & -4 & 0 \\
\end{array}\right)
\end{equation}

Figure 4.3: The wide derivative operator $D^k$ for elements with a Legendre basis, polynomial order $P^k = 2$, and element width $h^k$.

In the first column, only the first row is nonzero. And likewise in the last column, only the last row is nonzero. This indicates that collocation points in the neighbouring elements affect only the face of the element $\Omega^k$. The internal diagonal of the matrix vanishes, which indicates that the derivative of a function $\phi$ at a collocation point $x$ is independent of the value of $\phi$ at $x$. Finally, because $D^k$ represents a first derivative with respect to $x$, $D^k \phi \rightarrow -D^k \phi$ as $x \rightarrow -x$. This is reflected in the symmetry properties of the matrix, which obeys the relationship:

\begin{equation}
D_{ij}^k = -D_{(P^k-i)(P^k-j)}^k 
\end{equation}

for all $i = 0, 1, 2, \ldots, P^k$ and $j = -1, 1, 2, \ldots, P^k + 1$. These properties are generic in our scheme.
Figure 4.4: The wide derivative operator $D^k$ for elements with a Legendre basis, polynomial order $P^k = 3$, and element width $h^k$.

\[
D^k \approx \frac{1}{h^k} \begin{pmatrix}
-6 & 0 & 8.09 & -3.09 & 1 & 0 \\
0 & -1.62 & 0 & 2.24 & -0.62 & 0 \\
0 & 0.62 & -2.24 & 0 & 1.62 & 0 \\
0 & -1 & 3.09 & -8.09 & 0 & 6
\end{pmatrix}
\]

Figure 4.5: The wide derivative operator $D^k$ for elements with a Legendre basis, polynomial order $P^k = 4$, and element width $h^k$.

\[
D^k \approx \frac{1}{h^k} \begin{pmatrix}
-10 & 0 & 13.51 & -5.33 & 2.82 & -1 & 0 \\
0 & -2.48 & 0 & 3.49 & -1.53 & 0.52 & 0 \\
0 & 0.75 & -2.67 & 0 & 2.67 & -0.75 & 0 \\
0 & -0.52 & 1.53 & -3.49 & 0 & 2.48 & 0 \\
0 & 1 & -2.82 & 5.33 & -13.51 & 0 & 10
\end{pmatrix}
\]

4.4.7 Consistency

It is important to check that, if we differentiate a smooth function, our OLDG derivative converges to the continuum derivative in the separate limits of $h^k \to 0$ and $P^k \to \infty$. We call this property consistency.

To check consistency, suppose we seek to approximately differentiate a smooth function $\phi$. Then

\[
D^k \phi^k = d^k \phi^k
\]

and our OLDG derivative reduces to a pseudospectral derivative within the domain $\Omega^k$. Therefore the standard results for spectral methods apply and, as long as the physical solution is smooth, equation (4.59) provides a consistent approximation of a derivative.
[48, 195, 119]. More generally, if the continuum function $\phi$ is continuous at element boundaries, the spectral consistency results hold. If the continuum solution is discontinuous at element boundaries, then the wide component of our operator contributes and the issue is more delicate. We do not address it here.

4.4.8 Higher Derivatives

So far, we have only described how to approximate the continuum operator $\partial_x$. However, we’d also like to approximate higher derivatives. To approximate the second derivative $\partial_x^2 \psi$ of a function $\psi$, we introduce an auxiliary variable

$$\nabla \psi := D\psi,$$  \hspace{1cm} (4.61)

where $D$ is the wide derivative operator defined in equation (4.56), which is not evolved, but calculated globally whenever a second derivative is required. We then calculate the second derivative as

$$\frac{\partial^2}{\partial x^2} \psi \approx D\nabla \psi.$$  \hspace{1cm} (4.62)

This may not be the most efficient way of calculating higher derivatives. However, it was the only generic approach we could find that produced second derivatives compatible with the first derivatives defined in section 4.4.5 in the summation by parts sense:

$$\langle \psi, D^2 \phi \rangle_\Omega + \langle D\psi, D\phi \rangle_\Omega = \langle \psi, D\phi \rangle_{\partial \Omega},$$  \hspace{1cm} (4.63)
where $D^2$ refers to a wide second derivative operator.

We note that this approach to higher derivatives does come at a price. Since we must know $\nabla \phi$ on neighbouring elements before we can calculate $\partial_x^2 \phi$, we must wait for this calculation to finish before proceeding with our calculation of the right-hand-side. This reduces the parallelism of the scheme.

### 4.4.9 Truncation

In this section, we make a connection to a recent development in the LDG methods of Shu and collaborators. Consider the linear first-order in time, second-order in space wave equation defined on the interval $[X_l, X_r]$,

$$
\frac{\partial \phi}{\partial t} = \psi
$$

$$
\frac{\partial \psi}{\partial t} = c^2 \partial_x^2 \phi,
$$

(4.64)

where $\phi, \psi \in L_2(\Omega)$ are subject to appropriate initial and boundary conditions. We seek to discretize equations (4.64) using our DGFE scheme. This means replacing $\phi$ and $\psi$ by approximate representations $\phi^k$ and $\psi^k$ made up of the basis functions $\Phi_i^k$. We abuse notation here and denote by $\psi^k$ not only the restriction to element $\Omega^k$ but also the discretization.

Naively, one would use the same number of basis functions to represent $\phi$ and $\psi$ within each element. And indeed this is precisely the procedure Cockburn and Shu propose in [69]. In [236], Xing et al. show that if $\psi$ is represented as a piecewise polynomial of order $P^k = p \in \mathbb{N}$ in each element $\Omega^k$, but $\phi$ is represented only as a piecewise polynomial of order
\( P^k = p - 1 \), it is possible to obtain an energy conserving method that is superconvergent such that the error in the solution is of order \( O \left( (h^k)^{P_k + 3/2} \right) \). Whether or not a DGFE method is superconvergent depends strongly on the choice of numerical flux. For linear first-order systems, superconvergence is expected [119] but it is not always achieved in LDG methods [239].

Motivated by this choice, we represent \( \phi \) as a polynomial of order \( P^k = p \) but we represent \( \psi \) as a polynomial of order \( P^k = p - 1 \). This matches with the observation due to Richardson [188] that in the numerical solution of a hyperbolic PDE, the \( (h^k)^{-n} \) term in the solution introduced by differentiating \( n \) times is exactly cancelled by multiplications by \( h^k \) due to time integration (with a CFL factor). \( \phi \) is \( \psi \) integrated, so it is multiplied by \( h^k \), reducing the error. We do not obtain superconvergence, but we have found this procedure to significantly improve convergence and stability.

More generally, given a system of equations,

\[
\partial_t \psi = \mathcal{L} [\psi, \partial_j \psi, \partial_j \partial_k \psi],
\]

we call all variables with no spatial derivatives in their right-hand-sides primary variables and all other variables conjugate variables. Within an element, we represent all variables as linear combinations of the \( P^k \) basis functions. However, at each integrator substep and for each conjugate variable, we set the spectral coefficient associated with the \( (P^k) \)th mode of a conjugate variable to zero. We call this procedure truncation. The process of
truncation acts as artificial dissipation, improving the nonlinear stability of our scheme. (OLDG methods are linearly stable, so the stability cannot be improved.) Truncation also eliminates terms of order $O\left((h^k)^{P_k+1}\right)$ in our pointwise error estimates, providing cleaner convergence results.

### 4.4.10 Time Integration

With our derivatives defined, we integrate our scheme via the method of lines and an explicit time integrator such as the fourth-order total variation diminishing Runge-Kutta scheme proposed by Gottlieb and Shu [101] or the 8(7) scheme due to Dormand and Prince [83, 181].

Unfortunately, the timesteps for a DGFE method must be smaller than for a finite differences method. The time step size for DGFE methods for global timestepping is believed to obey

$$\Delta t \leq \min_{0<k<K} \Delta t^k \tag{4.65}$$

where

$$\Delta t^k := C_{\text{CFL}} \frac{h^k}{(P^k + 1)^2} \tag{4.66}$$

where $C_{\text{CFL}} > 0 \in \mathbb{R}$ is a constant that depends on the system being solved [119]. In our simulations we choose

$$C_{\text{CFL}} \leq 0.45 \tag{4.67}$$

however larger timesteps may be possible.
Although there is no proof for high-order elements [66],\(^4\) we present here a heuristic argument first developed in [219].\(^5\) Consider the linear system

\[
\partial_t \psi + a \partial_x \psi = 0
\] (4.68)

with semi-discrete OLDG approximation

\[
\partial_t \psi^k + a D^k \psi^k = 0.
\] (4.69)

We further approximate equation (4.69) as

\[
\partial_t \psi^k = \lambda^k \psi^k, \text{Real}(\lambda^k) \leq 0,
\] (4.70)

where \(\lambda^k\) plays the role of the eigenvalues of the operator \(a D^k\). The standard theory of explicit ordinary differential equation solvers [177] tells us that equation (4.70) is only stable for \(\lambda^k \Delta t^k \in \mathcal{R}\), where \(\mathcal{R}\) is some finite region in the complex plane.

Generically, \(\mathcal{R}\) could have a very strange shape. However, a circle of radius \(C_{\text{CFL}}\) can be contained within it. Therefore, our goal is not to understand \(\mathcal{R}\) in full generality but only to find how the norm of \(\lambda^k\) scales with \(h^k\) and \(P^k\) so that we can demand \(\Delta t^k\) satisfy

\[
\Delta t^k |\lambda^k| \leq C_{\text{CFL}}.
\] (4.71)

\(^4\)The bound was proven for polynomial order \(P^k = 1\) [68].

\(^5\)Our argument most closely follows the presentation that can be found in [66].
In [219] Trefethen and Trummer found that the 2-norm of a nodal derivative operator like $aD^k$ scales as
\[
\|aD^k\|_2 \leq C_B \frac{a}{h^k} (P^k + 1)^2
\] 
for some constant $C_B$.\(^6\) This in turn implies that the smallest norm of $\lambda$ should be similarly constrained. Thus we have,
\[
\Delta_t^k (P^k + 1)^2 \leq C_{CFL},
\] 
where we have absorbed the additional constants $a$ and $C_B$ into $C_{CFL}$. This implies equation (4.66).

There exist several techniques for increasing the CFL factor beyond that implied by equations (4.65) and (4.66), which we do not explore but which may be of use to the interested reader. For more details on these techniques, see [139, 230, 66] and references therein. Local time-stepping in the context of DGFE methods is discussed in [183]. We do not implement local time-stepping, however the implementation for OLDG methods should be the same as for standard DGFE methods.

### 4.4.11 Convergence

We do not prove convergence for the BSSN system. However, we know from section 4.4.7 and appendix 4.C that an OLDG discretization of the linear wave equation is both consistent and stable. Therefore by the Lax-Richtmyer theorem [145], it is convergent,\(^6\) In traditional DGFE methods, $C_B$ depends on the numerical flux. Since we fix our penalties to be the same in all cases, $C_B$ does not vary for OLDG methods.
with error bounded by the standard pseudospectral consistency bounds [48, 195, 119].

We can also obtain stronger, pointwise bounds. In appendix 4.D we show that, for the linear, second-order wave equation (4.64) with arbitrary initial conditions, we have

\[
\phi^k_i = \phi(t, x^k_i) + \mathcal{E}^k(t, x^k_i) \left( h^k \right)^{P^k},
\]

(4.74)

where \( \phi^k_i \) is the numerical solution in element \( \Omega^k \) with element width \( h^k \) and element order \( P^k \) at collocation point \( x^k_i \). \( \phi \) is the true solution, and \( \mathcal{E}^k \) is a function, which may depend on the location of the collocation points \( x^k_i \) but is independent of \( h^k \) and \( P^k \), that determines the error. In this proof we show that convergence does not depend on whether or not we truncate as described in section 4.4.9.

We argue that these convergence results provide analytic evidence that discretizations of more general systems are convergent to the appropriate order when discretized with our OLDG stencil. The numerical experiments discussed in section 4.6 support our claim.

### 4.4.12 Filtering

It is well known that a numerical scheme that is linearly stable may become unstable when used to solve a nonlinear system of equations, even when those equations are well-posed. In spectral and DGFE methods, we can interpret the loss of stability as emerging from the fact that the interpolating polynomial representing the derivative of a function is not the same as the derivative of an interpolating polynomial. This is usually called *aliasing*.
error. Stability can often be restored by filtering the spectral coefficients to remove energy from the short-wavelength modes [144].

These filtering techniques were originally motivated by the need to capture shocks in nonlinear flux-conservative systems and, more generally, to efficiently represent discontinuous functions spectrally [222]. However, they are often required even when the solution is smooth. If the system is nonlinear, aliasing error can drive an instability. The BSSN equations are no exception [215, 216].

One can think of the truncation scheme described in section 4.4.9 as a type of filter and we have found that, in many situations it provides all the required dissipation. However, since truncation is a projection-type operation, it is not tunable. We therefore develop a more traditional, tunable, filtering scheme which can be utilized if necessary.

We choose a modification of the spectral viscosity technique developed by Tadmor and collaborators for pseudospectral methods [209, 154, 210, 194]. Consider the semi-discrete system of the form

$$\frac{\partial}{\partial t} \phi = \mathcal{L}[\phi, D\phi, D^2\phi], \quad x \in [X_l, X_r],$$  \hspace{1cm} (4.75)

where \( \mathcal{L} \) is some nonlinear operator that acts on \( \phi \) and its first two wide OLDG derivatives, as defined in equation (4.59), subject to appropriate initial and boundary data. To make contact with traditional pseudospectral methods, we initially assume that \( K = 1 \) such that there is a single element, of width \( h^k = h \) and order \( P^k = p \). In this limit, our discontinuous Galerkin scheme becomes a pseudospectral method with weak boundary conditions. Once we have developed filtering in this setting, we will generalize to the full case.
Tadmor modifies equation (4.75) by including an artificial dissipation term which vanishes in the continuum limit:

\[
\frac{\partial}{\partial t}\phi = \mathcal{L}[\phi, D\phi, D^2\phi] + \epsilon_p (-1)^{s+1} \left[ \partial_x (1 - x^2) \partial_x \right]^s \phi,
\]

where \( s \) is the so-called order of the dissipation and strength of dissipation, and \( \epsilon_p \) varies as

\[
\epsilon_p \sim \frac{C_s h}{p^{2s-1}}.
\]

The constant \( C_s \) depends on the regularity of the solution \( \phi \). Roughly, it should be

\[
C_s \sim \max_{0 \leq k \leq s} \| \phi \|^k_{\infty},
\]

where \( \| \phi \|_{\infty} \) is the infinity-norm of \( \phi \). Under these conditions, Tadmor shows that equation (4.76) converges spectrally to the true solution. By inspection, we can see that Tadmor’s spectral viscosity technique is the spectral analog to the artificial dissipation proposed by Kreiss and Oliger [141].

Tadmor’s spectral viscosity technique is roughly equivalent to filtering the modes \( \hat{\phi} \) of \( \phi \) via the exponential filter first proposed by Vandeven [222]. In this case, the spectral representation of \( \phi \)

\[
\phi(x) = \sum_{i=0}^p \hat{\phi}_i \Phi_i(x),
\]
where $\Phi_i$ are the test functions, becomes

$$\phi^\sigma(x) = \sum_{i=0}^{p} \sigma \left( \frac{i}{p} \right) \hat{\phi}_i \Phi_i(x), \quad (4.78)$$

where

$$\sigma(\eta) := \exp \left[ -(C_s p \Delta t)^{\eta^s} \right], \quad (4.79)$$

where $\Delta t$ is the discrete time step used in evolution.

On the other hand, applying an exponential filter to the modes $\hat{\phi}$ is equivalent to solving equation (4.75) but also solving the ordinary differential equation

$$\frac{d}{dt} \hat{\phi}^i = -\epsilon_p \hat{\phi}^i, \quad t \in [0, \Delta t] \quad (4.80)$$

for all $0 \leq i \leq p$ at each time step. We claim that equations (4.75) and (4.80) need not be solved in separate steps and that the artificial viscosity formulation of Tadmor (4.76) can be well approximated by “filtering the right-hand-side” as

$$\frac{\partial}{\partial t} \phi = \mathcal{L} [\phi, D\phi, D^2\phi] - C_s p V \mathcal{F} V^{-1} \phi, \quad (4.81)$$

where $V$ is the Vandermonde matrix defined in equation (4.39) and

$$\mathcal{F}_{ij} := \left( \frac{i}{p} \right)^{2s} \delta_{ij}, \quad (4.82)$$
is a diagonal matrix defining the decay coefficients of the modal representation of $\phi$.

Alternatively, we can use a modified version of $\mathcal{F}$:

$$
\mathcal{F}_{ij} := \begin{cases} 
0 & \text{if } i/p \leq \eta_{\text{crit}}, \\
\delta_{ij} \left[ \frac{i/p - \eta_{\text{crit}}}{1 - \eta_{\text{crit}}} \right]^{2s} & \text{else}
\end{cases}
\quad (4.83)
$$

where $0 \leq \eta_{\text{crit}} < 1$ [119]. Equation (4.83) does not precisely correspond to the viscosity term provided in equation (4.76). Rather, it filters only the higher-order modes. Ideally, this is less destructive to the accuracy of the solution. We are currently using equation (4.83) in our implementation, but further investigation is necessary to determine what approach is best.

Of course, a DGFE method usually has more than one element, and we would like the dissipation term in equation (4.81) to scale appropriately with the number of elements. We therefore modify our dissipation term to the final form

$$
\frac{\partial}{\partial t} \phi^k = \mathcal{L} \left[ \phi^k, D^k \phi^k, (D^k)^2 \phi^k \right] \\
- C_s \frac{D^k}{h^k} \nu^k \mathcal{F}^k \left( \nu^k \right)^{-1} \phi^k,
\quad (4.84)
$$

where we have re-introduced the indexing for the element $\Omega^k$. In the full DGFE scheme, $C_s$ remains a global quantity, independent of the grid and $\mathcal{F}^k$ generalizes to multiple elements in the obvious way. Since the average distance between nodes is approximately $\Delta x^k := h^k/P^k$, the right-hand-side of equation (4.84) manifestly has the appropriate units of $1/\Delta x^k$.
for a hyperbolic problem.

4.5 Asymptotic Properties

In this section we analyze the computational costs of OLDG methods and compare them to finite differences.

4.5.1 Floating Point Operations for First Derivatives

Here we ask how many floating point operations are required to approximate $\frac{\partial}{\partial x}\phi$ for some function $\phi$. For simplicity, suppose a three-dimensional domain with $K$ elements on a side, for $K^3$ elements total. To make contact with finite differences, each element is of the same order $P^k = p$, $p$ even, such that the number of collocation points on a side is $n = (p + 1)K$. Further suppose that we are evolving only one variable.

Our in-element wide derivative operator is dense with vanishing diagonal. Therefore, in one dimension, a first derivative requires

$$2 \left[ (p + 1)^2 - (p + 1) + 2 \right] = 2 [(p + 1)p + 2]$$

floating point operations per element for the multiplication of the length $p + 3$ in-element state vector by our differentiation matrix. (The overall factor of 2 comes from the fact that add and multiply are separate operations.) In three dimensions and over the whole domain,
This translates to

\[
NF_{DG}^{(1)} := 2K^3(p + 1)^2 [(p + 1)p + 2]
= 2 \frac{n^3}{p + 1} [(p + 1)p + 2]
= 2n^3 \left[ p + \frac{2}{1 + p} \right]
\]  

(4.85)

floating point operations. In contrast, a \(p^{th}\)-order finite differences stencil requires

\[
NF_{FD}^{(1)} := 2n^3 p
\]  

(4.86)

floating point operations for a first derivative.

Figure 4.6 plots \(NF_{DG}^{(1)}/NF_{FD}^{(1)}\), which tells us how much more a DGFE derivative costs compared to a finite differences derivative. To leading order, both DGFE and finite differences stencils require a number of operations equal to \(O(n^3 p)\). However, the DGFE method has sub-leading terms which will contribute significantly when \(p\) is small and which become negligible when \(p\) is large.

### 4.5.2 Floating Point Operations for First and Second Derivatives

For most wave-like systems such as the BSSN system, we need both the first and second derivatives of variables in the state vector. We therefore ask how many floating point
operations are required to approximate both

\[ \partial_i \phi \text{ and } \partial_i \partial_j \phi, \ i, j = 1, 2, 3 \tag{4.87} \]

for some continuum variable \( \phi \). We make the same assumptions here as in section 4.5.1.

In the OLDG case, we take a first derivative, store it, and calculate a second derivative. Therefore the cost to approximate quantity (4.87) is just the cost of calculating the three first derivatives of \( \phi \) and then the cost of differentiating each of those quantities for a total of

\[
\begin{align*}
\text{NF}_{DG}^{(1,2)} &:= 3\text{NF}_{DG}^{(1)} + 6\text{NF}_{DG}^{(1)} \\
&= 18n^3 \left( p + \frac{2}{1 + p} \right) = \mathcal{O} \left( n^3 p \right)
\end{align*}
\tag{4.88}
\]

floating point operations.

Finite differences differentiation could be performed the same way, but it is typically not done. Usually \( \partial_i^2 \phi \) and \( \partial_i \partial_j \phi, \ i \neq j \) are calculated independently as full stencils without any intermediate steps or storage. With this approach, approximating \( \partial_i \partial_j \phi, \ i \neq j \) costs \( 2n^3 p^2 \) operations for each combination of \( i \) and \( j \). Approximating \( \partial_i^2 \phi \) to the same order of accuracy requires an extra two operations for the additional non-zero stencil point.
So, for all three directions, the cost of approximating quantity (4.87) is

\[
\begin{align*}
NF_{\text{FD}}^{(1,2)} &:= 3NF_{\text{FD}}^{(1)} + 6n^3(p + 1) + 6n^3p^2 \\
&= 6n^3[p + (p + 1) + p^2] \\
&= 6n^3(p + 1)^2 = \mathcal{O}(n^3p^2)
\end{align*}
\] (4.89)

operations.

Figure 4.6 plots \(NF_{\text{DG}}^{(1,2)}/NF_{\text{FD}}^{(1,2)}\), which tells us how much more (or less) it costs to calculate both first and second derivatives using our OLDG stencil compared to finite differences. For \(p \geq 3\), the cost of calculating all first and second derivatives for a function is larger for finite differences as usually implemented than for OLDG methods.

The standard finite differences implementation trades computational cost for memory and communication overhead. We note that finite differences implementations could calculate first derivatives and store them, just as we do in OLDG methods. See e.g. [125] for much more advanced FD stencil algorithms that greatly reduce memory access cost. We are unable to make this trade-off in our OLDG scheme because we need to respect both the weak boundary conditions between elements and summation-by-parts over the whole domain.
Figure 4.6: The ratio of the computation cost in floating point operations of our DGFE stencil compared to finite differences for the approximation of a first derivative (blue) and both first and second derivatives (green). The black line shows equal cost.

4.5.3 Communication Cost

We now investigate the communication costs in a distributed memory environment. We are interested in strong scaling: given a three-dimensional problem of fixed size $S^3$, where $S$ is the number of collocation points in a single dimension, across how many separate memory domains can we efficiently distribute the calculation?

To differentiate at a collocation point $g$, a finite differences stencil of order $p$, $p$ even, needs $p/2$ collocation points on each side. For distributed memory, this translates to a layer of unevolved “ghost cells,” $p$ cells deep around the border of the memory domain, as shown in figure 4.7. These cells are synchronized whenever differentiation is required and cause
In contrast, a DGFE method requires a layer of ghost cells only one cell deep, no matter the order of the method, as shown in figure 4.7. This stems from the fact that, to differentiate within an element, one only needs data from the boundaries of neighbouring elements. The distributional derivative of the DGFE approximation couples the elements only weakly.

We quantify communication overhead by calculating the ratio

\[ \text{OH} := \frac{N_{\text{ghost cells}}}{N_{\text{interior cells}}}, \quad (4.90) \]

or the number of ghost cells divided by the number of interior cells. This depends on the
total problem size $S^3$ and the number of memory domains $D$ across which we want to
distribute our problem. Additionally, in the finite differences case, it depends on the order
of the method.

The number of interior cells is always $\lfloor S^3/D \rfloor$. (It is impossible to have a fractional
number of interior cells. However, for brevity of notation, we will suppress the floor term
from now on.) This translates to an overhead of

$$ \text{OH}_{\text{FD}} = \frac{D}{S^3} \left( \frac{S}{D^{1/3}} + p \right)^3 - 1 $$

for finite differences. At scale, $S^3$ and $D$ are of the same order, so even a moderate $p$ such
as $p = 4$ can produce very large overheads. In contrast, DGFE methods have an overhead
of

$$ \text{OH}_{\text{DG}} = \frac{D}{S^3} \left( \frac{S}{D^{1/3}} + 2 \right)^3 - 1 $$

independent of the order of the element.

Since our method is an LDG method, we introduce extra communication steps. We
communicate when we calculate both first and second derivatives and when we perform
the truncation operation described in section 4.4.9. This does not change the ratio of ghost
to interior cells, but is an additional communication cost.

As a concrete example we plot in figure 4.8 the overhead associated with a fixed prob-
lem size of $S^3 = 1000^3$ collocation points and for different values of $D$. We compare
fourth- and eighth-order finite differences stencils with discontinuous Galerkin methods of
Figure 4.8: The communication overhead for computing the solution to a 3D problem of fixed size $1000^3$ cells compared between DGFE methods and fourth- and eighth-order finite differences methods. For comparison, we also include the unrealistic cost for an approach that scales perfectly.

any order. Perfect strong scaling has a constant overhead of 0 (blue line). If one arbitrarily assumes a maximum acceptable overhead of 1.0, meaning we have as many ghost cells as interior cells, then the eighth-order stencil scales to about $D = 3.5 \times 10^4$ domains, fourth-order to $D = 2.7 \times 10^5$, and DGFE to $D = 2.2 \times 10^6$. In this particular situation, DGFE stencils of any order scale about ten times further than fourth-order finite differences.
4.5.4 Memory Access Cost

Loading values from memory is a costly operation; accessing memory has on today’s systems a latency more than a hundred times larger than a floating point operation. It is thus important that as many memory load operations as possible can be served from a cache. To allow this, one simple optimization method arranges loops in such a way that one first loads a block of collocation points into the cache, and then performs as many operations as possible on this block without requiring additional memory accesses. This is called loop blocking.

Here we assume ideal loop blocking, and then calculate how many memory accesses are necessary to calculate a derivative. The small number of ghost zones required by DGFE methods also serves to improve performance.

To calculate a derivative at a collocation point \( g \), its respective neighbours must also be present. This means that when we calculate a derivative within a block of collocation points, we once again have interior and ghost cells. To quantify the additional memory access cost due to ghost cells, we define a memory access overhead

\[
MO := \frac{N_{\text{ghost cells}}}{N_{\text{interior cells}}}. \tag{4.93}
\]

Here we calculate the memory access overhead for an idealized L3 cache of fixed size \( C \) per process for OLDG methods and for finite differences.
A double-precision number requires 8 Bytes, so a cache of size $C$ can contain

$$N_{\text{total}} := N_{\text{ghost cells}} + N_{\text{interior cells}} = \frac{C}{V (8 B)}$$

(4.94)

total points, where $V$ is the total number of variables required. If we define $l$ such that

$$N_{\text{interior cells}} = : l^3,$$

(4.95)

then the total number of cells is

$$N_{\text{total}}^{\text{FD}} = (l + p)^3$$

(4.96)

for a finite differences stencil of order $p$ and

$$N_{\text{total}}^{\text{DG}} = (l + 2)^3$$

(4.97)

for DGFE methods at any order. If we solve for $l$ we find that

$$N_{\text{interior cells}}^{\text{FD}} = \left( N_{\text{total}}^{1/3} - p \right)^3$$

(4.98)

for finite differences and

$$N_{\text{interior cells}}^{\text{DG}} = \left( N_{\text{total}}^{1/3} - 2 \right)^3$$

(4.99)
for DGFE methods. This gives us a memory access overhead of

$$MO_{FD} = \frac{N_{total}^{FD}}{\left(\frac{N_{total}^{FD}}{3} - p\right)^3} - 1$$ (4.100)

for finite differences and

$$MO_{DG} = \frac{N_{total}^{FD}}{\left(\frac{N_{total}^{FD}}{3} - 2\right)^3} - 1$$ (4.101)

for DGFE methods.

As a concrete example, we calculate the memory access overhead for the BSSN system and a realistic cache size of $C = 1.5$ MByte per core. Here we ignore details of a realistic cache and assume a simple 1.5 MB “container.” The Einstein Toolkit [152, 89, 71] implementation of the BSSN system has 24 evolved variables [3, 7]. The cache must also contain the right-hand-side and everything we need to calculate it. In the case of finite differences, this is just the state vector and the right-hand-side. In the case of our discontinuous Galerkin scheme, this includes both first derivatives and temporary variables for the truncation operation described in section 4.4.9. Therefore we find that the BSSN system requires

$$V_{BSSN}^{FD} = 48 \text{ and } V_{BSSN}^{DG} = 92$$ (4.102)

for finite differences and OLDG stencils respectively.

Given these assumptions, we plot the memory access overhead points as a function of
Figure 4.9: The memory access overhead for the BSSN system with a fixed cache size of $C = 1.5$ MB for both finite differences and DGFE stencils as a function of the stencil order. We assume the BSSN system requires 48 variables per collocation point for finite differences and 92 variables per collocation point for DGFE.

The stencil order for both finite differences and DGFE stencils in figure 4.9. At second-order, and for this cache size, finite differences utilizes the cache slightly better. But at higher order, our discontinuous Galerkin scheme becomes significantly more efficient.

### 4.6 Numerical Tests

We have implemented our OLDG scheme as a thorn in the Einstein Toolkit [152, 89, 71], using Kranc for code generation [127, 140]. We provide our implementation, which is
based on the McLachlan thorn [55, 156], in a public repository [159]. We emphasize that our implementation is a proof-of-concept implementation and has not yet been optimized for performance. We are currently testing a more efficient implementation.

To establish the basic numerical properties of our method, we first investigate its applicability in the context of the second-order wave equation (4.64) of section 4.6.1. We then investigate in the context of the BSSN equations by performing some of the standard Apples-With-Apples tests [6, 38], which we discuss in sections 4.6.3, 4.6.4, and 4.6.5.

The original Apples With Apples tests were intended to test not only the stability and convergence of a code, but also the formulation of the Einstein equations on which that code is based. In this work, we are interested only in establishing the numerical properties of our scheme. Thus we only perform those Apples with Apples tests which probe the numerical scheme rather than the formulation.

We discuss both stability and convergence. All our numerical tests are performed with truncation as described in section 4.4.9 but without the filtering described in section 4.4.12.

### 4.6.1 The Second-Order Equation

Recall the second-order wave equation (4.64):

\[
\frac{\partial \phi}{\partial t} = \psi
\]
\[
\frac{\partial \psi}{\partial t} = c^2 \nabla^2 \phi.
\]
A DGFE method has two types of resolution: the order of the polynomial interpolant within each element and the total number of elements in the domain $K$. In appendix 4.D, we show that the pointwise error obeys equation (4.74):

$$\phi^k_i = \phi(t_i, x_i^k) + \mathcal{E}^k(t_i, x_i^k) \left( h^k \right)^{p_k},$$

however, because DGFE methods are defined only in a weak sense, and because the positions of collocation points change with resolution, equation (4.74) best translates to the following statement over the whole domain

$$\|E[\phi]\|_2 := \|\phi^k - \phi\|_2 \leq \|\mathcal{E}^k\|_2 \left( h^k \right)^{p_k}, \quad (4.103)$$

where $\|\phi\|_2$ is the 2-norm of $\phi$ over the domain. (Here we abuse notation and allow $\phi^k$ to not only represent the restriction of $\phi$ onto the element $\Omega^k$ but the numerical solution.)

To investigate convergence of OLDG methods, we numerically solve equation (4.64) with different numbers of elements and different in-element orders (the order stays fixed to $P^k = p$ over the whole grid.) The former is called $h$-refinement. The latter is called $p$-refinement. We solve equation (4.64) in 1D over the domain $x \in \left[-\frac{1}{2}, \frac{1}{2}\right]$ with periodic boundary conditions and initial conditions of the form

$$\phi(t = 0, x) = A \sin(2\pi k x) \quad (4.104)$$

$$\psi(t = 0, x) = -A \omega \cos(k x), \quad (4.105)$$
where \( \omega = \sqrt{k^2 c^2} \) for some amplitude \( A \) and wavenumber \( k \). For simplicity, we fix \( c = k = A = 1 \). We plot our error at \( t = 0.75 \).

Figure 4.10 shows the convergence under \( h \)-refinement using fourth-order elements. When we double the number of elements, we halve the element width \( h \). Equation (4.103) tells us that we should see the error scale as \( h^{-p} \), or \( h^{-4} \) in this case. And indeed measurements confirm this prediction.

Figure 4.11 shows the convergence under \( p \)-refinement with a fixed element width \( h = 1/8 \). The convergence given in equation (4.103) now translates to exponential decay with \( p \). This rapid convergence rate is often called “spectral” or “evanescent” convergence. Our
measured convergence agrees with this expectation. We measure the convergence rate to be

$$\|E[\phi]\|_2 = be^{-ap}$$

for $b \approx e^{5.66}$ and $a \approx 3.34$.

In figures 4.12 and 4.13, we compare the pointwise error of our OLDG approach with the pointwise error of a finite differences scheme. For figure 4.12, we use a fourth-order stencil and 40 collocation points (or 8 elements). For figure 4.13, we use an eighth-order stencil and 36 collocation points (or 4 elements). The curves are generated by using fourth-
Figure 4.12: The pointwise errors for the second-order wave equation (4.64) with OLDG (top) and finite differences (bottom) stencils respectively after 0.75 wave periods. These simulations were run with fourth-order stencils and 40 collocation points (or 8 elements).

order and eighth-order interpolation respectively. In the OLDG case, this interpolation corresponds to the modal representation within an element. The dots are measured values at the collocation points. For the OLDG stencil, the vertical lines show element boundaries. For consistency, all simulations were run with the same CFL factor and time integrator.

For the OLDG stencil, element boundaries are visible by eye as locations where the function is no longer smooth. We find the pointwise error for the OLDG stencil is worse than for the finite differences stencil of the same order and resolution. For fourth-order stencils, the OLDG error is about ten times worse than the finite differences error. For
eighth-order stencils it is almost fifty times worse at element boundaries. This error can be mitigated somewhat by the post-processing technique discussed in section 4.6.2 below.

We also find that the error for the OLDG stencil is significantly higher frequency than the error for the finite differences stencil, even in the linear case. The high-frequency nature of the error indicates that artificial dissipation may reduce error, even in the linear case when it is not required for stability. However, we did not investigate this possibility. Our experiments indicate that these traits are roughly generic, although the factor by which

Figure 4.13: The pointwise errors for the second-order wave equation (4.64) with OLDG (top) and finite differences (bottom) stencils respectively after 0.75 wave periods. (Note the differences in scale.) These simulations were run with eighth-order stencils and 36 collocation points (or 4 elements).
the finite differences error is smaller may depend on the order of the method.

This is a weakness of our OLDG method compared to both traditional DGFE methods and finite differences methods. Traditional DGFE methods have significantly more freedom with their numerical flux and they can, for example, employ upwinding to reduce their error.

However, we emphasize that we have been “fair” to finite differences methods by comparing stencils of the same order. At first glance, one might assume that a spectral method would have less error than a finite differences method. However this is only true if the spectral method is allowed to utilize arbitrarily high-order polynomials. In the comparisons shown in figures 4.12 and 4.13, we have restricted our discontinuous Galerkin elements to use polynomials of order fixed to that of the finite differences stencil.

In the fixed order case, we need about 14 fourth-order elements (or 70 collocation points) to do as well as the fourth-order finite differences stencil with 40 points. (See the top panel in figure 4.14.) However, if we vary the polynomial order within elements, we can get comparable accuracy to the finite differences stencil at similar computational cost and, as discussed in section 4.5, significantly improved communication and cache properties.

The bottom panel of figure 4.14 shows the pointwise error for the wave equation with 7 fifth-order elements, or 42 collocation points. The error is comparable to the fourth-order finite differences case shown in figure 4.12. And if our code were optimized, the computational cost would be similarly comparable. This example highlights how, even though the pointwise error for OLDG methods may seem inferior to finite differences in a
Figure 4.14: The pointwise errors for the wave equation with fourth- (top) and fifth-order (bottom) OLDG stencils respectively after 0.75 wave periods. The simulation for the fourth-order stencil was run with 14 elements and the simulation for the fifth-order stencil was run with seven. Note that the errors are comparable to the finite differences calculation in figure 4.12.

“fair” comparison, they will perform better in realistic situations.

4.6.2 Post-processing Element Boundaries

The pointwise error shown in figure 4.13 highlights a conceptual difficulty with DGFE methods. The “continuum” function recovered by modal representation is not continuous. Rather, it is piecewise smooth. However, we often know from physical considerations that the function that we solve for should be smooth over the whole domain. How then do we
recover a continuous function?

Given the collocation points $\phi^k_i$, we solve for the modal representation, and therefore the polynomial interpolant, within an element by solving equation (4.40)

$$\phi^k = \mathcal{V}^k \hat{\phi}^k$$

for the modal coefficients $\hat{\phi}^k_i$. However, we can replace equation (4.40) by the following

$$A^k \phi^{k\text{\_wide}} = \mathcal{V}^k \hat{\phi}^k, \quad (4.106)$$

where $\phi^{k\text{\_wide}}$ is a length $P^k + 3$ vector containing the points

$$\{\phi^{k-1}_r\} \cup \{\phi^k_i\}_{i=0}^{P^k} \cup \{\phi^{k+1}_l\}$$

and $A^k$ is a wide operator that maps $\phi^{k\text{\_wide}}$ to a length $P^k + 1$ vector which represents $\phi^k$ but with the left- and right-hand limits of $\phi^k$ at element boundaries mapped to their average. $A^k_{ij}$ has components

$$A^k_{ij} := \delta_{ij} + \frac{1}{2} (\delta_{i,0} \delta_{j,-1} - \delta_{i,0} \delta_{j,0}$$

$$+ \delta_{i,P^k} \delta_{j,P^k+1} - \delta_{i,P^k} \delta_{j,P^k}) \quad (4.107)$$

for all $i = 0, 1, 2, \ldots, P^k$ and $j = -1, 0, 1, \ldots, P^k + 1$. Here we use the shorthand
introduced in equation (4.51). The matrix

\[
A_{ij}^k = \begin{pmatrix}
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\
\end{pmatrix}
\]  \hspace{1cm} (4.108)

provides an example of \(A^k\) for polynomial order \(P^k = 4\).

If we solve equation (4.106) to generate our interpolation, we produce a function that is continuous everywhere, though it may not be smooth everywhere. Figure 4.15 shows the effects of this averaging scheme on pointwise errors for a simulation run with four eighth-order elements. As usual, the curves are the interpolated solution and the points are the collocation points. This procedure reduces pointwise error near element boundaries, but otherwise does not significantly change the pointwise error.

### 4.6.3 The Robust Stability Test

Our goal, of course, is not to solve the wave equation but to solve equations as complex as the BSSN system. To test our scheme’s properties in this more realistic and demanding setting, we perform the Apples-with-Apples tests developed by the community [6, 38].

The most basic of the Apples-with-Apples tests is the robust stability test. The robust
Figure 4.15: The error in the interpolating polynomial for the second-order wave equation (4.64) calculated using equation (4.40) (top) and equation (4.106) (bottom) respectively. This simulation was run with four eighth-order elements. The error is plotted at time $t = 0.75$. In this case, the maximum pointwise error is reduced by a factor of two.

stability test is an experimental and numerical analog to the stability condition for hyperbolic systems. It discerns whether the numerical approximation of our formulation of the continuum equations is stable to linear perturbations.

We take a constant-time slice of vacuum Minkowski space and introduce random noise. We choose the amplitude of the noise such that the Hamiltonian constraint is linearly satisfied. We then evolve the spacetime and watch the deviation from Minkowski space for various resolutions. A method passes the test if the deviation grows at most exponentially
Figure 4.16: The robust stability test under $h$-refinement. We use $4^{th}$-order elements for this test. We find the growth rate to be at most linear and to decrease with resolution.

in time, such that the maximum rate is independent of the resolution. Practically this means that the rate of growth in the deviation must not increase with increased resolution.

Figures 4.16 and 4.17 show the results of the robust stability test under $h$- and $p$-refinement respectively. We plot the $xy$-component of the spatial metric $\gamma$, which should vanish in the continuum. Therefore we measure error in the linear regime as a function of time. We find our discretization of the Einstein equations to be stable. By eye, the growth rate is at most linear and decreases with resolution under both kinds of refinement. For this test, we use a three-dimensional domain $\Omega = [-0.5, 0.5]^3$. 
Figure 4.17: The robust stability test under \( p \)-refinement. We use \( 4^3 \) elements for this test. We find the growth rate to be at most linear and to decrease with resolution.

### 4.6.4 Gauge Wave Test

As a test of the accuracy and convergence of our scheme for the BSSN system, we perform the *gauge wave* test \([6, 38]\). The gauge wave test is a periodic coordinate transformation of Minkowski space, which provides a metric with a known analytic solution against which we can compare. The metric for the one-dimensional gauge wave test is

\[
ds^2 = (1 - H)(-dt^2 + dx^2) + dy^2 + dz^2
\]  

(4.109)
Figure 4.18: A two-dimensional slice of the three-dimensional gauge wave at $t = 0$, generated with 8th-order elements. We plot the $xx$-component of the metric as a function of space in the $z = 0$ plane. The amplitude is commensurably smaller since $\gamma_{xx}$ is a projection of the rotated $\gamma$ onto the $x$-axis.

with

$$H = A \sin \left( \frac{2\pi(x - t)}{d} \right),$$

(4.110)

for some amplitude $A$ and wavelength $d$. This can be converted to a three-dimensional wave by rotating about the $y$- and $z$-axes by $\pi/4$ each. Figure 4.18 shows a two-dimensional slice of $xx$-component of the spatial metric for the 3D gauge wave. The boundaries of the cells are the collocation points and the value in the cell is the average value in that region.

We have chosen a deliberately low resolution to highlight the structure of the grid. The
non-uniform position of the collocation points can be seen in the varying cell sizes.

In all of our gauge wave simulations, we use an amplitude of $A = 0.01$ and a period of $d = 1$. In one dimension our domain is the interval $x \in [-0.5, 0.5]$. In three dimensions, our domain is the box $(x, y, z) \in [-0.5, 0.5]^3$. For the one-dimensional gauge wave, we use fourth-order elements. For the three-dimensional gauge wave, we use eighth-order elements. For time integration we use an explicit fourth-order or eighth-order Runge-Kutta integrator, as appropriate.

Figure 4.19 shows the error in the $x:x$-component of the metric for the one-dimensional gauge wave at $t = 3.75$ light crossing times, rescaled by $h^{-4}$, where $h$ is the element width for each element. As in figure 4.12, we generate the curves by interpolation using the modal representation within an element. The curves do not line up perfectly, but they are all contained within an envelope function, which converges at fourth-order, as expected.

Figure 4.20 shows the $L_2$-norm of the error for the one dimensional gauge wave as a function of time. We once again re-scale the error by $h^{-4}$. The fact that the curves overlap demonstrates fourth-order convergence for the system as it evolves in time.

Figure 4.21 shows the $L_2$-norm of the error for the three-dimensional gauge wave, using eighth-order elements. We now rescale by $h^{-8}$ and, again, the fact that the curves overlap demonstrates convergence of the appropriate order under $h$-refinement. At eighth order, we need very few elements before we see good convergence.

We note that since the y-axis is rescaled in figures 4.20 and 4.21, it does not repre-
Figure 4.19: “Pointwise” convergence of the one-dimensional gauge wave under $h$-refinement, with a fixed order of $p = 4$, at $t = 3.75$. We plot the error in the $xx$-component of the metric, rescaled by the element width to the $4^{th}$ power. The left pane shows most of the domain, while the right panel stretches out the axes so that the error is more visible. Since the collocation points do not align, pointwise convergence can’t be expected. However, the fact that the envelopes of the errors align indicates good convergence. For the 32 element simulation, the absolute error is approximately $O(10^{-8})$. For the 64 element simulation, it is approximately $O(10^{-10})$.

sent the true error. In particular, although the rescaled error is large, the absolute error is comparable to or better than that in [6, 38].

4.6.5 Gamma Driver Gauge Wave Test

The gauge wave prescription given in equation (4.109) has a harmonic lapse and vanishing shift. We would like to test more realistic lapse and shift conditions in this simplified
context, so we seek a generalization of the gauge wave. In \[38\], Babiuc et al. propose the \textit{shifted gauge wave}, which generalizes the original gauge wave to include a nonzero shift.

This nonzero shift is harmonic however and, as discussed in \[38\], the BSSN system performs poorly in this setting. Physical evolutions of the BSSN system typically use the \textit{Gamma driver} shift condition of the form \[7\]

\[
(\partial_t - \mathcal{L}_\beta) \beta^i = B^i \tag{4.111}
\]

\[
(\partial_t - \mathcal{L}_\beta) B^i = \alpha^2 \zeta (\partial_t - \mathcal{L}_\beta) \bar{\Gamma}^i - \eta B^i, \tag{4.112}
\]
Figure 4.21: The $L_2$-norm error over space of the error of the $xx$-component of the metric for the 3D gauge wave, rescaled by $1/h^8$. The curves overlap, showing that the system is converging at $8^{th}$-order. This simulation was run with $8^{th}$-order elements. For the $2^3$ element run, the error is approximately $O(10^{-8})$. For the $4^3$ element run, the error is approximately $O(10^{-10})$. For the $6^3$ element run, the error is approximately $O(10^{-12})$.

where $\beta^i$ are the components of the shift, $L_\beta$ is the Lie derivative in the $\beta$-direction, $\zeta, \eta \in \mathbb{R}$ are constants, and

$$\tilde{\Gamma}^i = -\partial_j \left( \psi^4 \gamma^{ij} \right)$$

(4.113)

is the conformally rescaled connection for spatial metric $\gamma$ and conformal factor $\psi$. The Gamma driver shift is combined with the $1+\log$ slicing condition, first developed by Bernstein [49] and Anninos et al. [34]. This is of the form [3, 7]:

$$\partial_t \alpha = -2\alpha K,$$

(4.114)
where $\alpha$ is the lapse and $K$ is the trace of the extrinsic curvature tensor.

Motivated by these observations, we propose a new version of the shifted gauge wave test, the *Gamma driver gauge wave*, which tests our method using the gauge conditions typically used with the BSSN system. We use the same domains and initial conditions as the gauge wave test, but we impose $1+\log$ slicing and a simplified version of the Gamma driver shift condition:

$$\partial_t \beta^i = \zeta \tilde{\Gamma}^i - \eta \beta^i,$$

(4.115)

where we choose $\zeta = \eta = 3/4$ [3, 7].

We do not know the analytic solution to this system of gauge conditions, but we can study convergence in this setting by comparing several coarse resolutions to a fine resolution instead of an analytic solution. This type of convergence test, which is weaker than convergence to a known solution, is called a *self-convergence test*.

In self-convergence, one must take care to rescale the error by the correct amount. A system is self-convergent to order $P$ if

$$\frac{(\gamma_{xx})_1 - (\gamma_{xx})_3}{h_1^P - h_3^P} = \frac{(\gamma_{xx})_2 - (\gamma_{xx})_3}{h_2^P - h_3^P},$$

(4.116)

where $i$ indexes three resolutions, such that $i = 1$ is the coarsest and $i = 3$ is the finest. For notational simplicity, we assume all elements have the same width and order and we therefore suppress the element index $k$. We also suppress dependence on $x$ and $t$. For details of where equation (4.116) comes from, see appendix 4.F.
Figure 4.22: The $L_2$-norm of the difference between two coarse resolutions and one fine resolution for the Gamma driver gauge wave as a function of time, normalized by an appropriate factor based on the number and order of the elements. The fact that the curves overlap demonstrates 4th-order convergence. For this test, we use 64, 128, and 192 elements. These differences are small, approximately $O(10^{-12})$ for the comparison between 64 elements and 192 elements and approximately $O(10^{-13})$ for the comparison between 128 elements and 192 elements.

Figure 4.22 shows the self-convergence for the one-dimensional Gamma driver gauge wave with fourth-order elements. We plot the $L_2$-norm over space of equation (4.116). Because our implementation is not yet optimized for performance, the 3D self-convergence test was too expensive. Therefore we do not perform the Gamma driver gauge wave test in 3D.

As an additional check, we can treat $P$ in equation (4.116) as a free variable, the convergence order, and solve for it numerically. If convergence is as we expect, we will recover
that $P$ is the same as the order of the element. Indeed, we can perform this calculation globally by taking the $L_2$-norm of $(\gamma_{xx})_i - (\gamma_{xx})_3$ for $i = 1, 2$ at each timestep and solving for $P$. We then obtain a measure of convergence as a function of time. Figure 4.23 shows the result of this procedure for the Gamma driver gauge wave using both 4th- and 8th-order elements. The measured convergence order agrees very well with our expectations.

### 4.7 Concluding Remarks

By performing an LDG discretization at the level of the differential operator, rather than at the level of the equations, we have developed a novel DGFE scheme that can be
used to discretize arbitrary second-order hyperbolic equations, in particular also the BSSN formulation of Einstein’s equations. In the process, we have made the formalism proposed by Hesthaven and Warburton [119] rigorous and combined it with summation by parts.

We analyzed and tested our scheme and its stability and accuracy for a series of standard test problems in numerical relativity, and find the expected polynomial (for $h$-refinement) and exponential (for $p$-refinement) convergence. Compared to finite differencing methods, the solution error is larger when using the same number of collocation points, but as for other DGFE methods, our OLDG scheme requires significantly fewer memory accesses and has a significantly lower communication overhead, and is thus more scalable on current high performance computing architectures.

Moreover, our focus on the derivative operator allows codes that currently employ finite differences methods to straightforwardly replace the finite differences stencil with our OLDG stencil, converting a finite differences code to a DGFE code. This should improve the parallel efficiency of such codes.

### 4.8 Acknowledgments

The authors thank David Radice for many helpful discussions. The authors also thank: Federico Guercilena for help with developing the testing infrastructure; Saul Teukolsky and the SpECTRE collaboration for discussions on DGFE methods; and Scott Field, Wolfgang Tichy and Jan Hesthaven for their help with literature search.
The authors acknowledge support from the Natural Sciences and Engineering Research Council of Canada and from the National Science Foundation of the USA (OCI 0905046, PHY 1212401). J. Miller acknowledges the support of the ACM/SIGHPC SC15 Student Travel Grant. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Research and Innovation.

We are grateful to the countless developers contributing to open source projects on which we relied in this work, including Cactus [99, 58] and the Einstein Toolkit [152, 89, 71], Kranc [127, 140], HDF5 [214], Python [189], numpy and scipy [221, 132], Matplotlib [126], and the yt-project [220].

This research used computing facilities at the Perimeter Institute, the Center for Computation & Technology at LSU, the Shared Hierarchical Academic Research Computing Network (SHARCNET) and Compute Canada, the Extreme Science and Engineering Discovery Environment (XSEDE) (TG-PHY100033), and the Blue Waters sustained-petascale computing project (PRAC_jr6).
Appendix

4.A Legendre Polynomials, Collocation Points, and Gauss Lobatto Quadrature

The following treatment can be found in standard textbooks such as [177] or [119]. Consider the domain $\Omega = [-1, 1]$ and a function $\psi \in L_2(\Omega)$. We wish to approximate the continuum function $\psi$ and its derivatives in a reasonable way. One such approximation is that

$$\psi(x) \approx \psi_{\text{modal}} := \sum_{i=0}^{P} \hat{\psi}_i \Phi_i(x), \quad (4.117)$$

where $P \in \mathbb{N}$, each $\Phi_i$ is some polynomial basis function such that $\{\Phi_i\}_{i=0}^{\infty}$ forms a complete orthonormal basis of $L_2(\Omega)$, and each $\hat{\psi}_i$ is a constant in $x$ defined by the projection

$$\hat{\psi}_i = \int_{-1}^{1} \psi(x) \Phi_i(x) dx = \langle \psi, \Phi_i \rangle_{\Omega}. \quad (4.118)$$
214

<table>
<thead>
<tr>
<th>Order $P$</th>
<th>Collocation Points $x_i$</th>
<th>Weights $w_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
<td>1.33</td>
</tr>
<tr>
<td></td>
<td>$\pm 1$</td>
<td>0.33</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>$\pm 0.65$</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>$\pm 1$</td>
<td>0.1</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>$\pm 0.47$</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>$\pm 0.83$</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>$\pm 1$</td>
<td>0.05</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>$\pm 0.36$</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>$\pm 0.68$</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>$\pm 0.90$</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>$\pm 1$</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Table 4.2: The approximate locations of the abscissas and values of associated weights for Gauss-Lobatto quadrature for several values of $P$.

Since the $\Phi_i$’s form a complete basis, equation (4.117) becomes exact in the limit $P \to \infty$. The demand that each $\Phi_i$ is a polynomial is crucial, as we will soon see.

We can also approximate $\psi$ in a so-called nodal basis, where we treat it as the polynomial that interpolates between known points $\psi(x_i)$ for some set of points $x_i \in \Omega$, $i = 0, 1, \ldots, N$, with $N \in \mathbb{N}$. We can write this as

$$
\psi(x) \approx \psi_{\text{nodal}}(x) := \sum_{i=0}^{N} \psi_i l_i(x),
$$

(4.119)
where \( \psi_i = \psi(x_i) \) and \( l_i(x) \) are the \( P^{th} \)-order Lagrange polynomials,

\[
l_i(x) = \prod_{0 \leq j \leq N, j \neq i} \frac{x - x_j}{x_i - x_j}.
\] (4.120)

Interpolating polynomials are unique. So in the special case \( N = P \), the two representations (4.117) and (4.119) in fact describe the same polynomial.

Suppose \( N = P \). Crucially, if equations (4.117) and (4.119) are different representations of the same function, there should be a way to transform between them. This is the Vandermonde matrix defined by its action as a transformation operator

\[
\psi_i = \sum_{j=0}^{P} \mathcal{V}_{ij} \hat{\psi}_j.
\] (4.121)

From equation (4.121), it is easy to show that

\[
\mathcal{V}_{ij} = \Phi_j(x_i).
\] (4.122)

In a realistic calculation, we will need both the Vandermonde matrix and its inverse \( \mathcal{V}^{-1} \). Therefore, the matrix should be well-conditioned.

One basis for which \( \mathcal{V} \) is well conditioned is the basis of Legendre polynomials. The Legendre polynomials are solutions to Legendre’s differential equation

\[
\frac{d}{dx} \left[(1 - x^2) \frac{d}{dx} \Phi_i(x)\right] + i(i + 1) \Phi_i(x) = 0 \forall i \in \mathbb{N},
\] (4.123)
but they are most easily defined recursively as:

\begin{align*}
\Phi_0(x) &= 1 \\
\Phi_1(x) &= x \\
(i + 1)\Phi_i(x) &= (2i + 1)x\Phi_i(x) - i\Phi_{i-1}(x)
\end{align*}

for all $i \in \mathbb{N}$.

In general, we may not be able to compute the integral in equation (4.117) with perfect accuracy or efficiency. Therefore, we would like to find a quadrature rule that allows us to efficiently calculate approximate integrals:

\[ \int_{-1}^{1} f(x) dx \approx \sum_{i=0}^{P} w_i f(x_i), \]

for some set of weights $w_i$ and collocation points $x_i$. Choosing a quadrature rule involves a choice both of weights $w_i$ and abscissas $x_i$. There are several good choices for quadrature rules. Because we are interested in using this discretization for a DGFE method, we want a rule where the abscissas include the endpoints of the domain $\Omega$. We therefore use Gauss-Lobatto quadrature.

\[ x_i \in \{-1, 1\} \cup \text{roots } (\Phi'_{p-1}), \]
where roots \( \Phi'_{p-1} \) is the set of solutions to the equation

\[
\Phi'_{p-1}(x) = 0.
\] (4.129)

The weights are then defined as the solutions to the linear system

\[
\int_{-1}^{1} \Phi_j(x) dx = \sum_{i=0}^{P} w_i \Phi_j(x_i)
\] (4.130)

for all \( j = 0, 1, \ldots, P \). In table 4.2, we provide approximate values for some of the Gauss-Lobatto quadrature points and their associated weights for different values of \( P \). A procedure for calculating these points precisely can be found in any standard numerical text such as \cite{177}.

By setting the collocation points in equation (4.119) equal to the abscissas defined by equation (4.128), we obtain the discrete approximation of \( L_2(\Omega) \). This representation is exact for all polynomials of order no greater than \( P \).

We are interested in performing this discretization within a DGFE element

\[
\Omega^k = [x^k_l, x^k_r],
\]

which is related to the interval \([-1, 1]\) by a simple linear coordinate transformation. This
transformation introduces a factor of $h^k/2$ into each of the weights:

$$w_i^k = \frac{h^k}{2} w_i$$ \hspace{1cm} (4.131)

where the weights must be computed for the element order $P^k$. This coordinate transformation also introduces a factor of $2/h^k$ into the modal derivative operator defined in equation (4.34).

4. B Proof That $\xi = 1$ Satisfies Summation By Parts

In this section, we prove that the scheme described in section 4.4 satisfies summation by parts if and only if $\xi = 1$, as asserted in section 4.4.5. We begin with the element-wise derivative operator defined in equation (4.55):

$$\partial_x \tilde{\psi}^k = D^k \psi^k = \left[ d^k - \frac{1}{2} b^k + \frac{1}{2} \xi F^k \right] \psi^k,$$

where $d^k$ and $b^k$ are the element-wise differentiation and boundary operators respectively and the fetch operator $F^k$ produces information about the boundary of neighbouring elements as defined in equation (4.50). We seek the conditions on it such that equation (4.57) holds:

$$\langle \psi, D\phi \rangle_\Omega + \langle D\psi, \phi \rangle_\Omega = \langle \psi, \phi \rangle_{\partial\Omega},$$
where $D$ is the derivative operator defined on the whole domain as given by equation (4.56):

$$D \psi(x) = (D^k \psi^k)(x) \quad \forall \ x \in \Omega^k \ \forall \ 1 \leq k \leq K.$$ 

In terms of the element-wise inner product, we demand that

$$\sum_{k=1}^{K} \left[ \langle D^k \psi^k, \phi^k \rangle_{\Omega^k} + \langle \psi^k, D^k \phi^k \rangle_{\Omega^k} \right] = \phi \psi \bigg|_{\partial \Omega}$$

(4.132)

(or some equivalent relation) for all test functions $\phi$ and $\psi$ in $L^2(\Omega)$. For convenience, we define

$$\phi_L = \phi(X_l), \quad \phi_R = \phi(X_r),$$

and

$$\psi_L = \psi(X_l), \quad \psi_R = \psi(X_r),$$

(4.133)

(4.134)

for the boundary elements $\Omega_0 = \{X_l\}$ and $\Omega_{K+1} = \{X_r\}$. Now, if we plug definition (4.55) into condition (4.132), we find that

$$\phi \psi \bigg|_{\partial \Omega} = S_1 + S_2$$

(4.135)

where we have split the sum over elements into two sums

$$S_1 = \sum_{k=1}^{K} \left[ \langle \left( d^k - \frac{1}{2} b^k \right) \phi^k, \psi^k \rangle_{\Omega^k} + \langle \phi^k, \left( d^k - \frac{1}{2} b^k \right) \psi^k \rangle_{\Omega^k} \right]$$

(4.136)
and
\[ S_2 = \frac{1}{2} \sum_{k=1}^{K} \left[ \langle F^k \phi^k, \psi^k \rangle_{\Omega_k} + \langle \phi^k, F^k \psi^k \rangle_{\Omega_k} \right], \tag{4.137} \]

which we will handle separately.

Let us examine \( S_1 \) first. Recall that \( b^k \) is defined in equation (4.44) by the relation:
\[ w^k b^k = w^k d^k + (d^k)^T w^k, \]
where \( w^k \) is the element-wise weight operator. Therefore,
\[ w^k \left( d^k - \frac{1}{2} b^k \right) = w^k d^k - \frac{1}{2} w^k b^k \]
\[ = w^k d^k - \frac{1}{2} w^k (w^k)^{-1} \left( w^k d^k + (d^k)^T w^k \right) \]
\[ = \frac{1}{2} \left( w^k d^k - (d^k)^T w^k \right), \tag{4.138} \]

and similarly
\[ \left( d^k - \frac{1}{2} b^k \right)^T w^k = (d^k)^T w^k - \frac{1}{2} \left( (d^k)^T w^k + w^k d^k \right) \left[ (w^k)^{-1} \right]^T w^k \]
\[ = -\frac{1}{2} \left( w^k d^k - (d^k)^T w^k \right), \tag{4.139} \]

where we have used the fact that a diagonal matrix is its own transpose. Therefore,
\[ w^k \left( d^k - \frac{1}{2} b^k \right) + \left( d^k - \frac{1}{2} b^k \right)^T w^k = 0 \tag{4.140} \]
and

\[
\left\langle \left( d^k - \frac{1}{2} b^k \right) \phi^k, \psi^k \right\rangle_{\Omega^k} + \left\langle \phi^k, \left( d^k - \frac{1}{2} b^k \right) \psi^k \right\rangle_{\Omega^k}
\]
\[
= (\phi^k)^T \left( d^k - \frac{1}{2} b^k \right)^T w^k \psi^k + (\phi^k)^T w^k \left( d^k - \frac{1}{2} b^k \right) \psi^k
\]
\[
= (\phi^k)^T \left[ w^k \left( d^k - \frac{1}{2} b^k \right) + \left( d^k - \frac{1}{2} b^k \right)^T w^k \right] \psi^k
\]
\[
= 0
\]  

(4.141)

for all \( \phi \) and \( \psi \) and for all elements \( \Omega^k \). Therefore, every term in the sum \( S_1 \) vanishes and

\[
S_1 = 0.
\]  

(4.142)

We now focus our attention on \( S_2 \). Recall from equation (4.50) that the fetch operator is

\[
F^k = b^k (b^{-1} F)^k,
\]

where \( (b^{-1} F)^k \) is defined in equations (4.48) and (4.49) such that

\[
(b^{-1} F)^k \psi^k_l = \psi^{k-1}_r
\]
\[
and (b^{-1} F)^k \psi^k_r = \psi^{k+1}_l.
\]

The fetch operator does the same thing, but selects only the boundary term, thus making anything else \( (b^{-1} F)^k \) does irrelevant. Furthermore, recall from equation (4.46) that the
product $w^k b^k$ is symmetric and has unit absolute value. So the sum $S_2$ becomes

$$S_2 = \frac{1}{2} \xi \sum_{k=1}^{K} \left[ \left( \psi_r^k \phi_{l+1}^k - \psi_l^k \phi_{r-1}^k \right) + \left( \phi_{l+1}^k \phi_r^k - \psi_{r-1}^k \phi_l^k \right) \right],$$

which is a telescoping sum. The $k + 1$ terms cancel with the $k - 1$ terms, leaving

$$S^2 = \frac{1}{2} \xi \left[ \left( \psi_R^K \phi_R + \psi_R \phi_R^K \right) - \left( \psi_L^1 \phi_L + \psi_L \phi_L^1 \right) \right].$$

We can interpret

$$\left( \psi_R^K \phi_R + \psi_R \phi_R^K \right)$$

and

$$\left( \psi_L^1 \phi_L + \psi_L \phi_L^1 \right)$$

as twice the average of $\phi$ and $\psi$ evaluated on the right and left boundaries of the domain respectively. Therefore their difference is a reasonable definition of the product $\phi \psi$ evaluated at the boundary of $\Omega$. And so

$$S^2 = (\xi) (\psi \phi)_{\partial \Omega}$$

and equation (4.135) is satisfied if and only if $\xi = 1$. □
4.C  Stability for the Wave Equation

Here we use summation-by-parts to demonstrate the stability an OLDG discretization of the linear first-order-in-time, second-order-in-space wave equation. This calculation provides an example of how one demonstrates stability with summation-by-parts.

Consider the second-order-in-space wave equation (4.64)

\[
\frac{\partial}{\partial t} \phi = \psi \\
\frac{\partial}{\partial t} \psi = c^2 \frac{\partial^2}{\partial x^2} \phi
\]

on the domain Ω subject to appropriate initial and boundary conditions. Using the OLDG approach, this translates to the semi-discrete system

\[
\frac{\partial}{\partial t} \phi^k = \psi^k \\
\frac{\partial}{\partial t} \psi^k = c^2 D^k \pi^k \\
\pi^k = D^k \phi^k,
\]

where \( D^k \) is the element-wise wide derivative operator. The continuum operator \( \frac{\partial}{\partial t} \) commutes with the discrete linear operator \( D^k \) so that

\[
\frac{\partial}{\partial t} \pi^k = D^k \frac{\partial}{\partial t} \phi^k = D^k \psi^k,
\]

where we have used the equations of motion to remove the time derivative. We do not use
equation (4.147) for evolution. Rather, we treat it as a constraint which is automatically satisfied.

This system admits the energy norm

\[ H = \frac{1}{2} \left[ \langle \psi, \psi \rangle_\Omega + c^2 \langle \pi, \pi \rangle_\Omega \right], \tag{4.148} \]

which is manifestly positive-definite. To show that our discretization is stable, we show that \( H \) is non-increasing in time. We differentiate equation (4.148) to find

\[
\partial_t H = \langle \psi, \partial_t \psi \rangle_\Omega + c^2 \langle \pi, \partial_t \pi \rangle_\Omega \\
= c^2 \langle \psi, D\pi \rangle_\Omega + c^2 \langle \pi, D\psi \rangle_\Omega, \tag{4.149}
\]

where we have used the discrete equations of motion (4.146) and the constraint (4.147).

Finally, we integrate by parts to obtain

\[
\partial_t H = c^2 \left[ \langle D\psi, \pi \rangle_\Omega - \langle D\psi, \pi \rangle_{\partial \Omega} \right] + c^2 \langle \psi, \pi \rangle_{\partial \Omega} \\
= c^2 \langle \psi, \pi \rangle_{\partial \Omega}, \tag{4.150}
\]

where we have used equation (4.57). The value of this expression depends on the boundary condition. For a large class of boundary conditions, including periodicity, homogeneous Dirichlet (\( \psi = 0 \)) or von Neumann (\( \pi = 0 \)), or maximally dissipative boundary conditions, this term is either zero or negative.
In this case, we have

$$\partial_t \mathcal{H} \leq 0.$$ \hfill (4.151)

Then, since $\mathcal{H} \geq 0$ and $\mathcal{H}$ is non-increasing, equation (4.146) provides a stable scheme. We note that, although we perform our calculation for a second-order system, it proceeds almost identically for a fully first-order system.

### 4.D Convergence for the Wave Equation

Here we present a calculation showing that the scheme described in section 4.4 is convergent for the linear wave equation.

#### 4.D.1 Strategy

The strategy of our proof is as follows. We use our discontinuous Galerkin scheme to solve the linear wave equation given arbitrary initial conditions and compare to the analytic solution. We write both the analytic solution and the “numerical” solution in terms of element width $h$ and element order $p$ so that we can write the error as a function of these two quantities. (As usual, we assume that all elements are the same width and order.)

We use Wolfram Mathematica \cite{128} to symbolically carry out the OLDG differentiation and Runge-Kutta integration, as described in section 4.4. In this way, our initial conditions can be truly arbitrary, and we only need to provide it in terms of a finite number of arbitrary constants. We have made our Mathematica code public and placed it in an online repository,
where it can be examined [158].

### 4.D.2 The Continuum Problem

Consider the one-dimensional domain $\Omega = \mathbb{R}$ and the interval

$$
T = [0, T] \text{ for some } T > 0 \in \mathbb{R}. \quad (4.152)
$$

We seek functions $\psi(t, x)$ and $\pi(t, x)$ which satisfy the linear wave equation in its first order in time reduction (4.64),

$$
\begin{align*}
\frac{\partial \phi}{\partial t} &= \psi \\
\frac{\partial \psi}{\partial t} &= c^2 \frac{\partial^2 \phi}{\partial x^2}
\end{align*}
$$

for all $x \in \Omega$ and all $t \in T$. Without loss of generality, we assume that $c = 1$.

If $\phi$ is analytic in $x$, then at any time $t$ it can be well-approximated by a power series

$$
\phi(t = 0, x) = a_0 + \sum_{i=1}^{N} (a_i + b_i)x^i, \quad (4.153)
$$

where $a_i$ and $b_i$, $0 \leq i \leq N$ are arbitrary constants that determine the initial profile. Of course, the solution to the wave equation given this initial condition is known. On the real line, $\psi$ is a superposition of right- and left-travelling waves that advect in each direction
with speed $c$:

$$
\phi(t, x) = a_0 + \sum_{i=1}^{N} a_i (x + t)^i + \sum_{i=1}^{N} b_i (x - t)^i
$$

(4.154)

with time-derivative

$$
\psi(t, x) = \sum_{i=1}^{N} i a_i (x + t)^{i-1} + \sum_{i=1}^{N} i b_i (x - t)^{i-1}.
$$

(4.155)

Therefore our initial condition for $\psi$ is given by

$$
\psi(t = 0, x) = \sum_{i=1}^{N} i (a_i + b_i) x^{i-1}.
$$

(4.156)

Note that the initial condition for $\psi$ contains one fewer modes than the initial condition for $\phi$. Enforcing this at each time step is equivalent to applying the truncation procedure discussed in section 4.4.9.

### 4.D.3 The Discrete Setup

To test the OLDG method developed in section 4.4, we use it to calculate a numerical approximation to the solution given by equations (4.154) and (4.155) with equations (4.153) and (4.156) as initial conditions. Crucially, we do not want to specify $a_i$ and $b_i$. Rather we want our solution in terms of them.

For simplicity we break $\Omega$ into a uniform “grid” of elements, all of the same width $h$ and order $p \leq N$. For initial data that is truly arbitrary, initial conditions (4.153) and (4.156)
are accurate up to order $h^N$ and $h^{N-1}$ respectively. To simulate a realistic situation, where the initial conditions introduce error equivalent to the order of the discretization scheme, we set $N = p$.

In principle, we have an infinite number of elements since our domain is the real line. In practice, however, we can examine a finite number $K$ of elements, spanning some interval $I = [-A, A] \subset \Omega$ as long as that number is sufficiently large so that no information from elements near the boundary of $I$ has time to propagate to elements near the centre of $I$ in time $T$.

We then write the positions of the nodes within elements, $x^k_i$ as a multiple of the element width $h$, which can be calculated by finding how many elements away from the origin the element $\Omega^k$ is and the “local” coordinates of $x^k_i$ within $\Omega^k$. We also define $\phi^k_i$ and $\psi^k_i$ as the restrictions of the fields $\phi$ and $\psi$ onto the nodes within elements, $x^k_i$.

**4.D.4  Comparing to the Continuum Solution**

Once we define our fields, we integrate them using a $p^{th}$-order explicit Runge-Kutta scheme and compare to the analytic solution. We perform this calculation for both second- and fourth-order stencils and the results seem to be generic. We use an explicit integrator of the same order as the OLDG stencil we wish to test. Our implementation of a second-order Runge-Kutta, for example, integrator is given by the following code.

\[
\text{dt} = \text{cfl} \times h; \quad (\text{\emph{factor is arbitrary}})
\]
RK2::usage = "Integrate \( y \) via RK2. The state vector is \( y \).

\[ \text{RK2}[y, t, f] := \text{Module}[\{k1, k2, yNew, tNew\}, \]
\[ \begin{align*}
  k1 &= f[t, y]; \\
  k2 &= f[t + (2/3)*dt, y + (2/3)*dt*k1]; \\
  yNew &= y + dt*((1/4)*k1 + (3/4)*k2); \\
  tNew &= t + dt;
\end{align*} \]
\[ \{tNew, yNew\}]; \]

Since the initial conditions are arbitrary, we only need to integrate by one time step. After integration, we subtract the true solution, given by equations (4.154) and (4.155), from the integrated solution and calculate the error. Because the wave equation is homogeneous, it is sufficient to study an element in the center of \( I \). For second-order elements (for example), this error is of the form

\[ \begin{align*}
  \phi^k_0 &= -h^2 \alpha (a_2 - b_2) \quad (4.157) \\
  \phi^k_1 &= 0 \quad (4.158) \\
  \phi^k_2 &= -\phi^k_0 \quad (4.159)
\end{align*} \]
for $\phi$ and

$$
\psi_0^k = -h\alpha(a_2 - b_2) \quad (4.160)
$$

$$
\psi_1^k = 0 \quad (4.161)
$$

$$
\psi_2^k = -\psi_0^k \quad (4.162)
$$

for $\psi$, where $\alpha$ is the Courant-Friedrichs-Lewy factor in this context. For a second-order element, initial conditions (4.153) and (4.156) also have errors of leading order $h^2$ and $h^1$ respectively, so convergence is retained. More generally, we find that the error in $\phi$ is of order $O(h^p)$ and that the error in $\psi$ is of order $O(h^{p-1})$.

### 4.E  Making Contact with Standard Discontinuous Galerkin Methods

In this section we provide an example of how our wide derivative operator relates to more traditional DGFE formulations. This calculation also provides a simple example of how OLDG methods can be used to discretize more complicated systems.

Consider the linear wave equation in first-order, flux-conservative, form

$$
\partial_t \phi + \partial_x f(\phi) = 0, \quad (4.163)
$$
where we have introduced the complex variable

\[ \phi = \psi + i\pi \]  

(4.164)

and flux

\[ f(\phi) = \left( \pi + ic^2\psi \right) \]  

(4.165)

for some continuum functions

\[ \psi(t, x) : [0, T] \rightarrow \mathbb{R} \]

and

\[ \pi(t, x) : [0, T] \rightarrow \mathbb{R} \]

and constants \( T > 0 \in \mathbb{R} \) and \( c \in \mathbb{R} \) subject to appropriate initial and boundary conditions.

For simplicity assume the domain is an interval \( \Omega = [X_l, X_r] \).

With the usual choice of Legendre basis functions, this translates to the semi-discrete system

\[ 0 = \partial_t \phi_i^k + D^k [f(\phi)]_i^k \]

for all elements \( \Omega^k \) and all \( i = 0, \ldots, P^k \), where \( P^k \) is the order of the element. If we expand the wide derivative operator \( D^k \), we obtain

\[ 0 = \partial_t \phi_i^k + d^k f_i^k - \frac{1}{2} b^k \left[ 1 - (b^{-1} F)^k \right] f_i^k, \]  

(4.166)
where we have now suppressed the dependence of $f$ on $\phi$.

To obtain the usual representation of a DGFE method, we must take the inner product with respect to a test function $\Phi_j^k(x)$, which is its own interpolant. Recall from equation (4.39) that

$$V_{ij}^k = \Phi_j^k(x_i^k)$$

and from equation (4.119) that

$$\phi^k(t, x) = \sum_{i=0}^N \phi_i^k l_i^k(x),$$

where the $l_i^k(x)$ are Lagrange interpolants. Finally recall that derivatives $d^k$ of polynomials of order $P^k$ and lower are exact, since this is how the narrow derivative operator is defined.

We thus have

$$0 = \left[ \sum_{i=0}^{P^k} \left( \Phi_j^k(x), l_i^k(x) \right) \frac{\partial}{\partial x} \phi_i^k \right]_{\Omega^k} + \left( \Phi_j^k, d^k f^k \right)_{\Omega^k} - \frac{1}{2} \left( \Phi_j^k, d^k \left[ 1 - (b^{-1} F)^k \right] f^k \right)_{\Omega^k}$$

$$= \left[ \max_{i=0}^{P^k} \left( \Phi_j^k(x), l_i^k(x) \right) \frac{\partial}{\partial x} \phi_i^k \right]_{\Omega^k} + \left( \sum_{i=0}^{P^k} \left( \Phi_j^k, \frac{\partial l_i^k(x)}{\partial x} f_i^k \right) \right)_{\Omega^k} - \frac{1}{2} \left( \Phi_j^k, b^k \left( 1 - (b^{-1} F)^k \right) f^k \right)_{\Omega^k}$$

$$= \mathcal{M}^k \frac{\partial \phi^k}{\partial x} + \mathcal{S}^k f^k - \frac{1}{2} \left( \Phi_j^k, b^k \left( 1 - (b^{-1} F)^k \right) f^k \right)_{\Omega^k},$$

(4.167)

where

$$\mathcal{M}^k_{ij} = \left( \Phi_j^k, l_i^k \right) \text{ and } \mathcal{S}^k_{ij} = \left( \Phi_j^k, \frac{\partial l_i^k}{\partial x} \right)$$

(4.168)
are the mass and stiffness matrices from standard discontinuous Galerkin methods and where we have suppressed the index notation to recover a matrix form within each element.

The last term in equation (4.167) still requires some massaging, however. Recall from equation (4.46) that \( w^k b^k \) is nonzero only on the boundary and it is always \( \pm 1 \). Then we can do away with the integral over the boundary and recover

\[
\mathcal{M}^k \partial_t \phi^k + S(\alpha \phi^k) = \frac{1}{2} \Phi^k \left( f_-^k - f_+^k \right) \bigg|_{x_i^k}
\]

or

\[
\mathcal{M}^k \partial_t \phi^k + S(\alpha \phi^k) = \Phi^k \left[ f_-^k - \frac{1}{2} \left( f_-^k + f_+^k \right) \right] \bigg|_{x_i^k}
\]

(4.169)

where \( f_-^k \) and \( f_+^k \) are the interior and exterior values of the flux on an element respectively. In other words \( f_-^k \) returns values on the boundary within an element and \( f_+^k \) returns values on the boundary from neighbouring elements.

We recognize equation (4.169) as a standard DGFE method in strong form with a simple central numerical flux

\[
f^*(\phi_-^k, \phi_+^k) = \frac{1}{2} C(\phi_+^k + \phi_-^k),
\]

(4.170)

for the interior and exterior values of \( \phi \) at the boundary of \( \Omega^k \). Therefore, in the simplest cases at least, our scheme matches traditional DGFE methods.
4.F Calculating Self-Convergence

Here we derive the test for self convergence (4.116) given in section 4.6.5. We follow a procedure first proposed by Richardson [188]. Suppose we are evolving the BSSN equations. Based on equation (4.74), suppose that the error in the $xx$-component of the metric is of the form:

$$(\gamma_{xx})_i = \gamma_{xx} + \mathcal{E} h^P_i,$$  \hspace{1cm} (4.171)

where $i$ indexes three resolutions, such that $i = 1$ is the coarsest and $i = 3$ is the finest. $\gamma_{xx}$ is the true solution, $\mathcal{E}$ is an “error” function, and $h$ and $P$ are the element width and order as usual. For notational simplicity, we assume all elements have the same width and order and we therefore suppress the element index $k$. We also suppress dependence on $x$ and $t$.

We now combine formula (4.171) for different values of $i$:

$$(\gamma_{xx})_1 - (\gamma_{xx})_3 = \gamma_{xx} + \mathcal{E} h^P_1 - \gamma_{xx} - \mathcal{E} h^P_3$$

$$= \mathcal{E} \left( h^P_1 - h^P_3 \right)$$  \hspace{1cm} (4.172)

and similarly

$$(\gamma_{xx})_2 - (\gamma_{xx})_3 = \mathcal{E} \left( h^P_2 - h^P_3 \right).$$  \hspace{1cm} (4.173)

If we combine equations (4.172) and (4.173), we find equation (4.116):

$$\frac{(\gamma_{xx})_1 - (\gamma_{xx})_3}{h^P_1 - h^P_3} = \frac{(\gamma_{xx})_2 - (\gamma_{xx})_3}{h^P_2 - h^P_3},$$
so we can check for self convergence by constructing the left- and right-hand-sides of equation (4.116) and comparing them. Self-convergence is a weaker statement than convergence, since it does not guarantee that a numerical solution converges to the \textit{true} solution.

It could, in principle, converge to something else.
Chapter 5

Nonlinear Stability of Kerr Black Holes

with Proca Hair

5.1 Context

This project began December 2016 and is a work in progress. The question was first posed to me by my collaborator, William East. The analytic work presented here is developed by myself. The time-dependent numerical studies are performed by me using code developed by W. East [86]. I am in the process of performing simulations for and writing a publication based on this work.
5.2 Abstract

Herdeiro, Radu, and Rúnarsson recently presented axisymmetric stationary black hole spacetimes with massive vector hair—evasions of the no hair theorems. A key condition for the existence of these solutions is that the system must be at the threshold of superradiance, the wave analogue of the Penrose process. This suggests that these stationary solutions might be the end-result of the superradiant instability. We seek to test this hypothesis by numerically studying the nonlinear stability of these solutions against perturbation. In the process, we will examine whether these solutions evade the intended physical consequences of the no hair theorems or only their technical assumptions.

5.3 Introduction

Light, massive bosons emerge naturally from quantum chromodynamics (QCD) [169, 233, 234] and Kaluza-Klein compactifications of antisymmetric tensors in higher dimensional models [175, 105, 106]. In the scalar case, these fields are called “axions” and are one proposed model of dark matter. A minimal coupling of massive vector fields—called Proca fields [182]—to gravity is a convenient toy model not only for string theory-motivated bosons, but for modified theories of gravity with a preferred reference frame [122] or with vector modes [131].

A wave with frequency $\omega$ can superradiantly scatter off of a spinning black hole$^1$ with horizon frequency $\Omega_H$ and mine it for energy in an analogue of the Penrose process if

$$\omega \leq m\Omega_H.$$  

---

$^1$Superradiant scattering is also possible for a charged black hole if the wave is a force carrier for the charge.
where $m$ is the azimuthal winding number of the wave [242, 205]. If the wave is confined, via reflecting boundaries or simply via mass, then positive feedback can drive a superradiant instability, which Press and Teukolsky originally called a “black hole bomb” [176].

Based on the early work of Israel [129, 130] and Carter [62], Bekenstein showed in [44, 43] that all asymptotically flat stationary black hole solutions that (1) possess no naked singularities and (2) share symmetries between the metric and matter fields must be described by just a few numbers: the mass, spin, and charge of the black hole. However, a complex boson field with harmonic time dependence can be made compatible with a stationary metric. This scenario evades the second assumption of Bekenstein’s no hair theorems. In [170] Peña and Sudarsky were able to extend Bekenstein’s result to the complex boson case in spherical symmetry.

In [117], Herdeiro and Radu developed stationary axisymmetric black hole solutions with massive scalar boson hair (KBHsSH), evading both Bekenstein’s and Peña and Sudarsky’s no hair theorems. In [115] Herdeiro, Radu, and Rúnarsson (HRR) built on this work and presented solutions with Proca hair (KBHsPH). A key ingredient in these constructions is that the boson field is at the very threshold of superradiance. Superradiant dynamics could therefore act either as a stabilizing influence, driving systems towards these solutions, or as a destabilizing one, making them unstable against perturbations that activate the instability. Which of these scenarios is borne out is an open question.

Recent nonlinear simulations of the superradiant instability provide some evidence. In

---

2For a review, see [52].
3Some presentations also assume the null energy condition [115].
4Alternatively, one may use two real minimally coupled boson fields.
[86], East and Pretorius study the superradiant scattering of a complex Proca field with harmonic time and angular dependence such that the metric is axisymmetric (although the Proca field is not) and the Proca field has a fixed azimuthal winding number $m = 1$. They find that the system dynamically approaches the superradiant threshold, at which point the field reaches a quasistationary state. This may be evidence that HRR’s hairy black holes are not only stable, but the end state of the superradiant instability.

Demanding that the metric is axisymmetric is a severe restriction, however, since it suppresses (most but not all) gravitational wave emission. It is possible that HRR’s solutions (and/or Will and Pretorius’ solutions) are unstable to nonaxisymmetric perturbations or to perturbations with $m \geq 2$. We seek to resolve these ambiguities by performing simulations to determine the nonlinear stability of HRR’s KBHsPH against perturbations. If they are unstable, they cannot possibly be the end state of the superradiant instability. If they are stable, then perhaps the no hair theorems are not as broadly applicable as previously believed.

HRR’s original construction is in a coordinate system that is singular at the horizon. Therefore, an important first step towards studying the dynamical stability of these solutions is to move them into a horizon-penetrating coordinate system, which will be more amenable to numerical calculations. We present the required coordinate transformation and subsequent new coordinate system in this chapter. In section 5.4, we review HRR’s construction. In section 5.5, we present our coordinate transformation. In section 5.6, we apply our coordinate transformation in some simple test cases and validate its efficacy. In
section 5.7, we provide some concluding thoughts and present the next steps in our project.

5.4 HRR’s Construction of Kerr Black Holes with Proca Hair

We now review HRR’s construction of KBHsPH [115]. We begin by describing the Einstein-Proca system. We then discuss HRR’s ansatz and how this ansatz can be used to evade the no hair theorems.

5.4.1 The Einstein-Proca System

HRR study a complex vector field $\mathcal{A}$ with mass $\mu$ minimally coupled to general relativity via the action$^5$

$$S = \int d^4x \sqrt{-g} \left( \frac{1}{16\pi} R - \frac{1}{4} F_{\alpha \beta} \tilde{F}^{\alpha \beta} - \frac{1}{2} \mu^2 \mathcal{A}_\alpha \tilde{\mathcal{A}}^\alpha \right),$$  \hspace{1cm} (5.2)

where

$$\mathcal{F}_{\mu \nu} = 2 \partial_{[\mu} \mathcal{A}_{\nu]}$$  \hspace{1cm} (5.3)

is the field tensor associated with $\mathcal{A}$. The associated equations of motion are

$$\nabla_\alpha \mathcal{F}^{\alpha \beta} = \mu^2 \mathcal{A}^\beta$$  \hspace{1cm} (5.4)

$^5$Where, as usual, $G = c = 1$
for the Proca field and

$$R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} = 8\pi T_{\alpha\beta} \quad (5.5)$$

for gravity, where

$$T_{\alpha\beta} = \frac{1}{2} \left( F_{\alpha}^{\sigma} F_{\beta}^{\sigma} + \tilde{F}_{\alpha}^{\sigma} \tilde{F}_{\beta}^{\sigma} \right) - \frac{1}{4}g_{\alpha\beta} F_{\sigma\gamma} \tilde{F}^{\sigma\gamma} + \frac{1}{2} \mu^2 \left( A_{\alpha} \tilde{A}_{\beta} + \tilde{A}_{\alpha} A_{\beta} - g_{\alpha\beta} A_{\sigma} \tilde{A}^{\sigma} \right)$$

(5.6)

is the energy momentum tensor for a complex Proca field. Note that, in contrast to massless vector fields, equations (5.3) and (5.4) imply that $A$ is completely determined by $F$.

Moreover, one can show that the structure of these equations implies the Lorentz condition:

$$\nabla_{\alpha} A^{\alpha} = 0. \quad (5.7)$$

It is not a gauge choice.

### 5.4.2 The Metric Ansatz

Given the action (5.2), HRR assume the following axisymmetric, stationary metric ansatz

$$ds^2_0 = -e^{2F_0} N dt^2 + e^{2F_1} \left( \frac{dr^2}{N} + r^2 d\theta^2 \right) + e^{2 F_2} r^2 \sin^2(\theta) (d\phi - W dt)^2, \quad (5.8)$$
where \( W \) and \( F_i, i = 0, 1, 2 \) are functions of \( r \) and \( \theta \) that vanish as \( r \to \infty \) (for asymptotically flat spacetimes) and where

\[
N = 1 - \frac{r_h}{r} \tag{5.9}
\]

for the event horizon radius \( r_h \).\(^6\) Metric (5.8) is singular at \( r = r_h \). This implies that each \( F_i, i = 0, 1, 2 \) must satisfy appropriate regularity conditions at the horizon. One way of imposing this is to define

\[
x = \sqrt{r^2 - r_h^2} \tag{5.10}
\]

and demand that

\[
\partial_x F_i = 0 \ \forall \ i = 0, 1, 2. \tag{5.11}
\]

HRR also impose that

\[
\lim_{r \to r_h} W = \Omega_H, \tag{5.12}
\]

where \( \Omega_H \) is the horizon frequency of the black hole.

For a vacuum black hole, the ansatz functions are known analytically to be [117]:

\[
e^{2F_1} = \left(1 + \frac{b}{r}\right)^2 + b(b + r_h) \frac{\cos^2 \theta}{r^2} \tag{5.13}
\]

\[
e^{2F_2} = e^{-2F_1} \left\{ \left[ \left(1 + \frac{b}{r}\right)^2 + \frac{b(b + r_h)}{r^2} \right]^2 - b(b + r_h) \left(1 - \frac{r_h}{r}\right) \frac{\sin^2 \theta}{r^2} \right\} \tag{5.14}
\]

\[
F_0 = -F_2 \tag{5.15}
\]

\[
W = e^{-2(F_1 + F_2)} \sqrt{b(b + r_h)(r_h + 2b)} \left(\frac{1 + \frac{b}{r}}{r^3}\right) \tag{5.16}
\]

\(^6\)This ansatz can also be used to construct scalar and vector boson stars if \( r_h \to 0 \).
where $b$ is the “spheroidal prolateness” parameter. It is related to the mass and angular momentum of the black hole by

$$b = \frac{a^2}{R_+}$$  (5.17)

where

$$a = \frac{J}{M}$$  (5.18)

is the black hole angular momentum per unit mass and

$$R_+ = M + \sqrt{M^2 - a^2}$$  (5.19)

is the radius of the outer event horizon in Boyer-Lindquist coordinates.

HRR’s radial coordinate $r$ is related to the Boyer-Lindquist radius $R$ by a radial shift

$$r = R - b.$$  (5.20)

The ADM mass and angular momentum can be calculated in terms of $r_h$ and $b$ as:

$$M = \frac{1}{2}(r_h + 2b)$$  (5.21)

$$J = \frac{1}{2}\sqrt{b(b + r_h)(r_h + 2b)}$$  (5.22)
and the horizon frequency as

\[ \Omega_H = \frac{1}{r_h + 2b} \sqrt{\frac{b}{r_h + b}}. \]  

(5.23)

More generally, the ADM quantities in an asymptotically flat spacetime can be calculated by examining the leading-order falloff of the metric as \( r \to \infty \). That is,

\[ \lim_{r \to \infty} (g_{00})_{tt} = 1 - \frac{2GM}{r} + \ldots \]  

(5.24)

and

\[ \lim_{r \to \infty} (g_{00})_{\phi t} = -\frac{2GJ}{r} \sin^2(\theta) + \ldots. \]  

(5.25)

5.4.3 The Proca Ansatz

The complex Proca one-form \( A_0 \) is assumed to be of the form

\[ A_0 = \Phi \tilde{A}_0 \]  

(5.26)

with

\[ \Phi = e^{i(m\phi-\omega t)} \]  

(5.27)

and

\[ \tilde{A}_0 = uV dt + H_1 dr + H_2 d\theta + iH_3 \sin \theta d\phi \]  

(5.28)
for winding number $m \in \mathbb{N}$ and frequency $\omega$. $H_i, i = 1, 2, 3$ and $V$ are functions of $r$ and $\theta$ subject to similar regularity conditions at the horizon:

$$\partial_\theta H_i = 0 \forall i = 1, 2, 3. \quad (5.29)$$

Since the coordinate system is axisymmetric, the following symmetry conditions apply at the symmetry axis:

$$H_1 \bigg|_{\theta = 0, \pi} = \partial_\theta H_2 \bigg|_{\theta = 0, \pi} = \partial_\theta H_3 \bigg|_{\theta = 0, \pi} = V \bigg|_{\theta = 0, \pi} = \partial_\theta F_i \bigg|_{\theta = 0, \pi} = \partial_\theta W \bigg|_{\theta = 0, \pi} = 0. \quad (5.30)$$

Asymptotic flatness implies that the Proca field must be square integrable and thus $V$ and $H_i$ must vanish as $r^{-1}$ or faster as $r \rightarrow \infty$ for all $i = 1, 2, 3$.

Note that metric (5.8) is manifestly time-independent. The spacetime is stationary—i.e., it possesses an asymptotically timelike Killing vector. However, the matter field (5.28) is decidedly time-dependent. The two ansatzes are compatible because the energy momentum tensor (5.6) is time-independent—the phase factor (5.27) cancels out. This is the first ingredient in evading the no hair theorems. Bekenstein’s no hair theorems assume that both the matter and the metric possess the appropriate Killing vector, which is not true here [44, 43].

If the metric is to be stationary, the flux of energy momentum across the event horizon must vanish. This implies that

$$\lim_{r \rightarrow r_h} \xi^\alpha A_\alpha = 0 \quad (5.31)$$
for horizon null generator
\[ \xi^\alpha \partial_\alpha = \partial_t + \Omega H \partial_\phi. \] (5.32)

This implies that
\[ \omega = m \Omega H \] (5.33)

and that
\[ \lim_{r \to r_h} \left[ V + \frac{\omega}{m} H_3 \sin \theta \right] = 0. \] (5.34)

The frequency given by equation (5.33) is at exactly the threshold for superradiance. And this in turn hints at how HRR have evaded Peña and Sudarsky’s no hair theorem [170]. Here the black hole must be rotating and the Proca field must have harmonic azimuthal dependence, which is not possible under Peña and Sudarsky’s assumptions of spherical symmetry and statisticity.

Perhaps unintuitively, this implies that, for a given fixed Proca mass \( \mu \) and winding number \( m \), there is no \( \omega \to 0 \) limit of HRR’s KBHsPH.\(^7\) This limit can be understood by examining the non-relativistic test-field limit of the system. The Proca field forms a bound state around the black hole that has a hydrogen-like spectrum and that is determined by the horizon mass and spin of the black hole and four quantum numbers \( (n, l, m, j) \) corresponding to the spin-weighted spherical harmonics. The ADM mass and angular momentum of the system are then also determined by these quantum numbers. This extra compati-

\(^7\)For a given horizon mass and frequency, it is always possible to choose a winding number \( m \) and Proca mass \( \mu \) such that a hairy solution exists. However, if the winding number and mass are fixed, this freedom is lost.
bility condition implies KBHsPH exist only within a (continuous) subspace of the space parametrized by ADM mass and angular momentum and are unique only within a smaller subspace. See, for example, figure 6 of [115]. For a detailed discussion of this restriction for the scalar case, see [117].

5.4.4 Finding Solutions that Satisfy the Ansatz

When the ansatz described in sections 5.4.2 and 5.4.3 is fed into the Einstein-Proca equations (5.3), (5.4), (5.5), (5.6), and (5.7), and combined with appropriate boundary conditions, one attains a two-dimensional overdetermined elliptic system of differential equations. HRR impose the boundary conditions by introducing the compactification

\[ X = \frac{x}{x + 1}, \tag{5.35} \]

where \( x \) is given by equation (5.10). This maps the domain \( x \in [0, \infty) \) to \( X \in [0, 1] \). Asymptotically Minkowski boundary conditions may then be imposed at \( X = 1 \). HRR have provided sample data of solutions to this system [116]. We plot the Proca one-form for one of HRR’s solutions with horizon radius \( r_h = 0.2475 \) and Proca frequency \( \omega = 0.9775 \) in figure 5.1. We are also in the process of developing our own infrastructure to solve the Einstein-Proca system based on [157].

\[ \text{8We also compactify. However, we use a different compactification scheme. For elliptic systems, we use a scheme due to Boyd [50]. For time-evolutions, we use the scheme described by Pretorius in [180] and [179].} \]
Figure 5.1: The components of the Proca one-form for $r_h = 0.2475$ and $\omega = 0.9775$ generated by HRR in [116]: (a) $\text{Im}(A_t)$, (b) $\text{Re}(A_r)$, (c) $\text{Re}(A_\theta)$, and (d) $\text{Im}(A_\phi)$. The inner boundary is the event horizon $r_h$. The outer boundary is spacelike infinity.

5.5 Transformation to Horizon Penetrating Coordinates

As shown in figure 5.2, coordinate system (5.8) is singular at the horizon. With the application of regularity (and perhaps infall-type) boundary conditions, this doesn’t pose a problem for elliptic or perturbative calculations. Indeed, the interior of the black hole is not physically observable, so the horizon is a natural boundary. However, when one includes backreaction, this singularity can become problematic for two reasons. First, as the metric
Figure 5.2: Lapse $\alpha$ and volume form $\sqrt{\gamma}$ for HRR’s coordinates for a vacuum black hole spacetime with $M = 0.1945$ and $a = 0.95M$ in code units. Note that the lapse collapses and the volume form diverges at the inner boundary, which is the horizon. This indicates the coordinates are singular there. The outer boundary of the plot is $r = \infty$. The $\alpha \to 1$ as $r \to \infty$ and $\sqrt{\gamma} \to r^2 \sin^2 \theta$, which is the volume form for spherical coordinates.

The equations of motion for the metric can become exponentially unstable in the presence of a coordinate singularity without some regularization technique.

One standard way to avoid these difficulties is to move to a horizon-penetrating coordinate system, where this singularity is not present. The classic Kerr-Schild coordinates are one such coordinate system [74]. Because of the presence of the numerically-determined ansatz functions $F_i, W, V, H_i$, we cannot easily access Kerr-Schild coordinates.

We can, however, access a coordinate system that shares some of the desirable properties of Kerr-Schild coordinates—namely, it is non-singular at the event horizon; foliated by spacelike hypersurfaces, each of which is asymptotically Euclidean; and manifestly independent of the asymptotically timelike and azimuthal coordinates. We call our new coordinate system “Kerr-Schild-Like” or “Generalized Kerr-Schild.” (We will justify this
name in section 5.6.1.) We now present this new coordinate system for the first time and its derivation by coordinate transformation away from HRR’s coordinate ansatz (5.8).

### 5.5.1 A Toy Example

Before we consider the full transformation, we will use a simple toy example to motivate our construction. Consider the two-dimensional “toy” line element

$$\tilde{ds}^2 = -dt^2 e^{2F_0} N + e^{2F_1} \frac{dr^2}{N}.$$  \hspace{1cm} (5.36)

The $1/N$ term in equation (5.36) diverges at the horizon. We wish to regularize this. To do so, we define a new coordinate $v = t + r^*$ such that $\tilde{g}_{rr}$ vanishes. For this to work out, we need to satisfy the following condition:

$$dr^* = \frac{e^{F_1 - F_0}}{N} dr$$ \hspace{1cm} (5.37)

or

$$r^* = \int \frac{e^{F_1 - F_0}}{N} dr.$$ \hspace{1cm} (5.38)

The resulting “tortoise” line element is

$$\tilde{ds}_T^2 = -dt^2 e^{2F_0} N + 2e^{F_0 + F_1} dv dt.$$

\hspace{1cm} (5.39)
Vectors normal to constant $v$ slices are not guaranteed to be timelike in this coordinate system. We can enforce a foliation by spacelike hypersurfaces by introducing

$$\tau = v - r$$

(5.40)

so that our line element becomes

$$\tilde{ds}_{KSL}^2 = -e^{2F_0}N d\tau^2 + 2 \left( e^{F_0+F_1} - e^{2F_0}N \right) d\tau dr + \left( 2e^{F_0+F_1} - e^{2F_0}N \right) dr^2.$$  

(5.41)

This “Kerr-Schild-like” line element can be foliated by spacelike hypersurfaces at constant time $\tau$.

### 5.5.2 Transforming the Full Line Element to Tortoise-Like Coordinates

Consider the more general transformation into “tortoise-like” coordinates

$$x^\mu = (t, r, \theta, \phi) \rightarrow x'^\mu = (v, r, \theta, \psi)$$

(5.42)

where

$$v = t + r^*$$

(5.43)

$$\psi = \phi + r^\#.$$  

(5.44)
Our goal is a coordinate system that is regular at the horizon \( r = r_h \) when \( N = 0 \). Motivated by our result in section 5.5.1, we set

\[
\begin{align*}
&&
\end{align*}
\]

and

\[
\begin{align*}
&&
\end{align*}
\]

where \( C(\theta) \) is a constant of integration. (We will choose \( C \) on physical grounds shortly.)

Now, however, \( F_1 \) and \( F_0 \) depend on \( \theta \). Therefore,

\[
\begin{align*}
&&
\end{align*}
\]

and

\[
\begin{align*}
&&
\end{align*}
\]

We choose \( C(\theta) \) such that \( \Xi \to 0 \) as \( r \to \infty \). (We plot \( \Lambda \) and \( \Xi \) for an example metric in figure 5.3.) Then our Jacobian becomes

\[
\begin{align*}
\end{align*}
\]
Figure 5.3: Example values of (a) $\Lambda$ and (b) $\Xi$ in the $xz$-plane for a vacuum black hole spacetime with $M = 0.1945$ and $a = 0.95M$ in code units. The inner boundary is the horizon and the outer boundary is spacelike infinity.

and our inverse Jacobian becomes

$$
\left( \frac{\partial x^\mu}{\partial x'^\nu} \right)_T = \begin{pmatrix}
1 & -\frac{e^{F_1-F_0}}{N} & -\Lambda & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & -We^{F_1-F_0} & -\Xi & 1
\end{pmatrix}.
$$

Then the metric

$$(g_T)_{\mu'\nu'} = (g_0)_{\mu\nu} \frac{\partial x^\mu}{\partial x'^\mu} \frac{\partial x^\nu}{\partial x'^\nu}. $$
becomes the following:

\[
(g_T)_{vv} = -e^{2F_0}N + e^{2F_2}W^2 r^2 \sin^2 \theta \tag{5.51}
\]

\[
(g_T)_{vr} = (g_T)_{rv} = e^{F_0 + F_1} \tag{5.52}
\]

\[
(g_T)_{v\theta} = (g_T)_{\theta v} = e^{2F_0}N\Lambda - e^{2F_2}r^2 W^2 \mathcal{B} \sin^2 \theta \tag{5.53}
\]

\[
(g_T)_{v\psi} = (g_T)_{\psi v} = -e^{2F_2}W r^2 \sin^2 \theta \tag{5.54}
\]

\[
(g_T)_{rr} = 0 \tag{5.55}
\]

\[
(g_T)_{r\theta} = (g_T)_{\theta r} = -e^{F_0 + F_1}\Lambda \tag{5.56}
\]

\[
(g_T)_{r\psi} = (g_T)_{\psi r} = 0 \tag{5.57}
\]

\[
(g_T)_{\theta\theta} = r^2 e^{2F_1} - e^{2F_0}N\Lambda^2 + e^{2F_2}r^2 \mathcal{B}^2 \sin^2 \theta \tag{5.58}
\]

\[
(g_T)_{\theta\psi} = (g_T)_{\psi\theta} = e^{2F_2}r^2 \sin^2 \theta \tag{5.59}
\]

\[
(g_T)_{\psi\psi} = e^{2F_2}r^2 \sin^2 \theta \tag{5.60}
\]

where

\[
\mathcal{B} = W\Lambda - \Xi. \tag{5.61}
\]

(We plot \(\mathcal{B}\) for an example black hole spacetime in figure 5.5(a).) This metric will be regular if both \(\Lambda\) and \(\mathcal{B}\) are regular. We show this regularity in appendices 5.A and 5.B.
5.5.3 Transforming the Full Line Element to Spherical Kerr-Schild-Like Coordinates

As in example 5.5.1, we now introduce a second mapping

\[ x^\mu = v, r, \theta, \psi \rightarrow x^{\mu''} = \tau, r, \theta, \psi \]

given by

\[ \tau = v - r, \]  \hspace{1cm} (5.62)

which takes us into a coordinate system we call “spherical Kerr-Schild-like” (SKSL).

This transformation has Jacobian

\[
\left( \frac{\partial x^{\mu''}}{\partial x^{\mu'}} \right)_{\text{sksl}} = \begin{pmatrix}
1 & -1 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]  \hspace{1cm} (5.63)

and inverse Jacobian

\[
\left( \frac{\partial x^{\mu'}}{\partial x^{\mu''}} \right)_{\text{sksl}} = \begin{pmatrix}
1 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}.
\]  \hspace{1cm} (5.64)
When we apply it to our line element, the metric

$$(g_{SKSL})_{\mu'\nu'} = (g_T)_{\mu'\nu'} \left(\frac{\partial x^\mu'}{\partial x^{\mu'}}\right)_{skl} \left(\frac{\partial x^\nu'}{\partial x^{\nu'}}\right)_{skl}$$

becomes the following:

$$(g_{skl})_{\tau\tau} = -e^{2F_0} N + e^{2F_2} W^2 r^2 \sin^2 \theta$$ (5.65)

$$(g_{skl})_{\tau\theta} = (g_{skl})_{\theta\tau} = e^{F_0 + F_1} - e^{2F_0} N + e^{2F_2} W^2 r^2 \sin^2 \theta$$ (5.66)

$$(g_{skl})_{\tau\psi} = (g_{skl})_{\psi\tau} = -e^{2F_2} W^2 r^2 \sin^2 \theta$$ (5.68)

$$(g_{skl})_{\tau\psi} = (g_{skl})_{\psi\tau} = -e^{2F_0} N \Lambda + e^{2F_2} W^2 r^2 \sin^2 \theta$$ (5.67)

$$(g_{skl})_{\tau\psi} = (g_{skl})_{\psi\tau} = -e^{2F_0} N \Lambda + e^{2F_2} W^2 r^2 \sin^2 \theta$$ (5.69)

$$(g_{skl})_{rr} = 2e^{F_0 + F_1} - e^{2F_0} N + e^{2F_2} W^2 r^2 \sin^2 \theta$$ (5.70)

$$(g_{skl})_{r\theta} = (g_{skl})_{\theta r} = -e^{F_0 + F_1} + e^{2F_0} N \Lambda - e^{2F_2} W^2 r^2 \sin^2 \theta$$ (5.71)

$$(g_{skl})_{r\psi} = (g_{skl})_{\psi r} = -e^{2F_2} W^2 r^2 \sin^2 \theta$$ (5.72)

$$(g_{skl})_{\theta\theta} = -(e^{2F_0} N \Lambda^2 + r^2 e^{2F_1} + e^{2F_2} W^2 r^2 \sin^2 \theta)$$ (5.73)

$$(g_{skl})_{\theta\psi} = (g_{skl})_{\psi\theta} = e^{2F_2} W^2 r^2 \sin^2 \theta$$ (5.74)

Note that our metric in SKSL coordinates contains a logarithmic divergence as $r \to \infty$. (See appendix 5.C for more details.) However, this divergence vanishes when we transform to Cartesian coordinates, as we show in appendix 5.G.
5.5.4 Transforming the One-Form into SKSL Coordinates

After we apply inverse Jacobians (5.50) and (5.64) to the one-form $\tilde{A}_0$ defined in equation (5.28), it becomes

$$(\tilde{A}^0_{\text{sksl}})_{\mu''} = \begin{pmatrix} iV \\ H_1 + i e^{F_1-F_0} \Sigma \\ H_2 - i [V \Lambda + \Xi H_3 \sin \theta] \\ i \sin \theta H_3 \end{pmatrix},$$

where $\Sigma$ is defined to be

$$\Sigma = \frac{V + W H_3 \sin \theta}{N}. \quad (5.75)$$

Although it appears as though it may be problematic at the horizon, $\Sigma$ is actually regular, as shown in figure 5.4 (a) and as we show in section 5.D.

We also wish to understand how $\Phi$, defined in equation (5.27), is changed by transformations (5.43) and (5.44). Note that

$$t = \tau - r^* + r$$

$$= \tau + r - \int_{r_h}^{r} \frac{e^{F_1-F_0}}{N} dr - C \quad (5.76)$$

and $\phi = \psi - r^#$

$$= \psi - \int_{r_h}^{r} W \frac{e^{F_1-F_0}}{N} dr - \Omega_H C, \quad (5.77)$$
and consider

\[-i \ln \Phi = m\phi - \omega t\]

\[= m \left( \psi - \int_{r_h}^{r} W e^{F_1-F_0} \frac{e^{F_1-F_0}}{N} dr - \Omega_H C \right) - \omega \left( \tau + r - \int_{r_h}^{r} \frac{e^{F_1-F_0}}{N} dr - C \right)\]

\[= m\psi - \omega \tau - \omega r + m \int_{r_h}^{r} e^{F_1-F_0} \left( \Omega_H - W \right) \frac{\Omega_H - W}{N} dr + (\omega - m\Omega_H) C\]

\[= m\psi - \omega \tau - \omega r + m \int_{r_h}^{r} e^{F_1-F_0} \rho dr\]

\[= m\psi - \omega \tau - \omega r + m\Delta\]  \hspace{1cm} (5.78)

where

\[\Delta = \int_{r_h}^{r} e^{F_1-F_0} \frac{\Omega_H - W}{N} dr.\]  \hspace{1cm} (5.79)

We plot $\Delta$ in figure 5.5(b). Although it is not obvious, $\Delta$ is regular at the horizon. We

Figure 5.4: Example values of (a) $\Sigma$ and (b) $-i \ln \Psi$ in the $xz$-plane for Proca data with $r_h = 0.2475$ and $\omega = 0.9775$ generated by HRR in [116]. The inner boundary is the horizon and the outer boundary is spacelike infinity. Note that $\Sigma$ is regular at the horizon.
Figure 5.5: Example values of (a) $\hat{b}$ and (b) $\Delta$ in the $xz$-plane for a vacuum black hole spacetime with $M = 0.1945$ and $a = 0.95M$ in code units. The inner boundary is the horizon and the outer boundary is spacelike infinity.

will show this in section 5.E. The phase is indeterminant as $r \to \infty$. However, this is not a difficulty since $A$ vanishes in this limit.

It will be convenient to absorb the $e^{-i\omega r}$ term into $A_{\text{sksl}}$ so that $\Phi$ depends only on $\tau$ and $\psi$. Therefore we redefine

$$A_{\text{sksl}} = \Phi_{\text{sksl}} \tilde{A}_{\text{sksl}}$$

(5.80)

where

$$\Phi_{\text{sksl}} = \exp \left[ i (m\psi - \omega\tau) \right]$$

(5.81)
and

\[
(\mathcal{A}_{\text{sksl}})_{\mu''} = \Psi_{\text{sksl}} \begin{pmatrix}
iV \\
H_1 + i e^{F_1 - F_0} \Sigma \\
H_2 - i [V \Lambda + \Xi H_3 \sin \theta] \\
i \sin \theta H_3
\end{pmatrix},
\] (5.82)

such that

\[
\Psi_{\text{sksl}} = \exp \left[ i (m \Delta - \omega r) \right]
\] (5.83)

is an additional phase introduced by the coordinate transformation. We plot an example phase angle for \(\Psi\) in figure 5.4 (b).

### 5.5.5 The 3+1 Split for SKSL Coordinates

In the ADM (or 3+1) decomposition of spacetime, we write the line element as\(^9\)

\[
ds_{\text{sksl}}^2 = - \left( \gamma_{\text{sksl}} \right)_{ij} dx^i dx^j
\] (5.84)

where \(\left(\gamma_{\text{sksl}}\right)_{ij}\) is the 3-metric,

\[
\left(\beta_{\text{sksl}}\right)_i = \begin{pmatrix}
e^{F_0 + F_1} - e^{2F_0} N + e^{2F_2} \rho^2 W^2 \sin^2 \theta \\
e^{2F_0} N \Lambda - e^{2F_2} \rho^2 W \sin^2 \theta \\
-e^{2F_2} \rho^2 \sin^2 \theta
\end{pmatrix},
\] (5.85)

\(^9\)As usual, Greek indices range from 0 to 3 and Latin indices range from 1 to 3.
such that
\[
\kappa = 2 e^{F_1} r^2 - e^{F_0} N r^2 - e^{F_0} \Lambda^2
\] (5.87)
is the shift and
\[
\alpha_{\text{skl}} = e^{F_1} \left( -1 + 2 e^{F_1} - e^{F_0} + \frac{r h}{r} - \frac{\Lambda^2}{r^2} \right)^{-1/2}
\] (5.88)
is the lapse. The unit vector orthogonal to the hypersurfaces with intrinsic metric \(\gamma_{\text{skl}}\) is given by
\[
(n_{\text{skl}})^\mu = \frac{1}{\alpha} \left( 1, (\beta_{\text{skl}})^i \right).
\] (5.89)

A straightforward calculation shows that
\[
(n_{\text{skl}})_\mu (n_{\text{skl}})^\mu = (g_{\text{skl}})_{\mu\nu} (n_{\text{skl}})^\mu (n_{\text{skl}})^\nu = -1,
\] (5.90)
indicating that our hypersurfaces are spacelike.

For the 3+1 decomposition of the Proca field, we follow [244, 86] and define\(^{10}\)
\[
\mathcal{A}_\mu = \chi_\mu + n_\mu \chi,
\] (5.91)

\(^{10}\)We suppress the subscript specifying a coordinate system because these formulae are valid in any coordinate system, so long as \(n\) is timelike.
which can be computed explicitly by

\[ \chi = n^\mu A_\mu \]  

(5.92)

and

\[ \chi_\mu = A_\mu - n_\mu \chi. \]  

(5.93)

We also define an “electric” field

\[ E_i = \gamma^\mu_i F_{\mu\nu} n^\nu \]  

(5.94)

and “magnetic” field

\[ B^i = \epsilon^{ijk} D_j \chi_k, \]  

(5.95)

where

\[ \gamma^\mu_\nu = \delta^\mu_\nu + n^\mu n_\nu \]  

(5.96)

is the projector and \( D_j \) is the covariant derivative associated with the 3-metric \( \gamma \) and

\[ F_{\mu\nu} = 2\partial_{[\mu} A_{\nu]} \]  

(5.97)

is the field two-form. The magnetic field is not an independent variable. Rather, it is a shorthand for the antisymmetrized gradient of \( \chi \). Note that it is often beneficial to raise the
index on the electric field, which can be effected by applying the inverse metric:

$$E^i = \gamma^{ij} E_j.$$ (5.98)

For convenience, we also define “tilde” versions of $\chi$, $E$, and $F$, which depend only on $r$ and $\theta$:

$$\chi_{\text{sksl}} = \Phi_{\text{sksl}} \tilde{\chi}_{\text{sksl}}$$ (5.99)

$$(\chi_{\text{sksl}})_i = \Phi_{\text{sksl}} (\tilde{\chi}_{\text{sksl}})_i$$ (5.100)

$$(E_{\text{sksl}})_i = \Phi_{\text{sksl}} (\tilde{E}_{\text{sksl}})_i$$ (5.101)

$$(F_{\text{sksl}})_{\mu\nu} = \Phi_{\text{sksl}} (\tilde{F}_{\text{sksl}})_{\mu\nu}$$ (5.102)

where $\Phi_{\text{sksl}}$ is defined in equation (5.81). Here we have re-introduced the coordinate-subscript since this decomposition can only be performed in appropriate coordinate systems (and metrics). In this decomposition, we can utilize the known structure of the field to calculate the field tensor. For more details see section 5.F.

### 5.5.6 Cartesian Kerr-Schild-Like Coordinates

To avoid the coordinate singularity at the symmetry axis, we would like to map to Cartesian Kerr-Schild-like coordinates (CKSL) via the transformation

$$x^{\mu''} = (\tau, r, \theta, \psi) \rightarrow x^{\mu'''} = (\tau, x, y, z)$$
where

\[
\tau = \tau \quad (5.103)
\]

\[
x = r \sin \theta \cos \psi \quad (5.104)
\]

\[
y = r \sin \theta \sin \psi \quad (5.105)
\]

\[
z = r \cos \theta \quad (5.106)
\]

with Jacobian

\[
\left( \frac{\partial x^\mu'''}{\partial x^{\nu''}} \right)_{\text{cksl}} = 
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & \cos \psi \sin \theta & r \cos \theta \cos \psi & -r \sin \theta \sin \psi \\
0 & \sin \theta \sin \psi & r \cos \theta \sin \psi & r \cos \psi \sin \theta \\
0 & \cos \theta & -r \sin \theta & 0
\end{pmatrix} \quad (5.107)
\]

and inverse Jacobian

\[
\left( \frac{\partial x^{\mu''}}{\partial x^\nu'''} \right)_{\text{cksl}} = 
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & \cos \psi \sin \theta & \sin \theta \sin \psi & \cos \theta \\
0 & \frac{1}{r} \cos \theta \cos \psi & \frac{1}{r} \cos \theta \sin \psi & -\frac{1}{r} \sin \theta \\
0 & -\frac{1}{r} \csc \theta \sin \psi & \frac{1}{r} \csc \theta \cos \psi & 0
\end{pmatrix}, \quad (5.108)
\]
which we can use to calculate the metric in CKSL coordinates via

\[
(g_{\text{cksl}})_{\mu''\nu''} = (g_{\text{sksl}})_{\mu''\nu''} \left( \frac{\partial x^\mu''}{\partial x^\mu'\mu''} \right)_{\text{cksl}} \left( \frac{\partial x^\nu''}{\partial x^\nu'\nu''} \right)_{\text{cksl}}
\] (5.109)

Note that the usual choice for Kerr-Schild coordinates is slightly different [74]. We make our choice such that in our coordinates the outer event horizon is at a constant coordinate radius \( r_h \).

Because of the algebraic nastiness involved, we present the metric only for \( \psi = 0 \) (i.e., in the \( xz \)-plane):

\[
(g_{\text{cksl}})_{\tau\tau} = -e^{2F_0} N + e^{2F_2} r^2 W^2 \sin^2 \theta \tag{5.110}
\]

\[
(g_{\text{cksl}})_{\tau x} = (g_{\text{cksl}})_{x\tau} = \psi \cos \theta + \zeta \sin \theta + e^{2F_2} \left( rW \sin \theta - \psi \cos \theta \right) rW \sin^2 \theta \tag{5.111}
\]

\[
(g_{\text{cksl}})_{\tau y} = (g_{\text{cksl}})_{y\tau} = -e^{2F_2} rW \sin \theta \tag{5.112}
\]

\[
(g_{\text{cksl}})_{\tau z} = (g_{\text{cksl}})_{z\tau} = \zeta \cos \theta - \psi \sin \theta + e^{2F_2} \left( rW \cos \theta + \psi \sin \theta \right) rW \sin^2 \theta \tag{5.113}
\]

\[
(g_{\text{cksl}})_{xx} = e^{2F_1} \cos^2 \theta - \frac{\Lambda}{r} \cos^2 \theta + \left( e^{F_0 + F_1} + \zeta \right) \sin^2 \theta + e^{2F_2} \psi^2 \cos^2 \theta \sin^2 \theta + e^{2F_2} r^2 W^2 \sin^4 \theta \tag{5.114}
\]

\[
-2 \frac{\Lambda}{r} \zeta \cos \theta \sin \theta - 2 e^{2F_2} \psi rW \cos \theta \sin^3 \theta
\]
\[(g_{kst})_{xy} = (g_{kst})_{yx} = e^{2F_2} \cot \theta \sin \theta - e^{2F_2} r W \sin^2 \theta \]  
\[(g_{kst})_{xz} = (g_{kst})_{zx} = \frac{1}{2} \sin(\theta) \left[-2\eta \cos(\theta) + e^{2F_2} \sin(\theta) \left(2rW\Xi \cos(2\theta) + (rW - \Xi)(rW + \Xi) \sin(2\theta)\right)\right] + \left(\frac{\Lambda}{r}\right) e^{F_0} \xi \cos(2\theta) - e^{2F_2} r W \sin^2(\theta) \left(rW \cos(2\theta) - \Xi \sin(2\theta)\right) + \left(\frac{\Lambda}{r}\right)^2 \cos \theta \sin \theta \left(e^{2F_0} N - e^{2F_2} r^2 W^2 \sin^2(\theta)\right)\]
\[(g_{kst})_{yy} = e^{2F_2} \]  
\[(g_{kst})_{yz} = (g_{kst})_{zy} = -e^{2F_2} r W \cos \theta \sin \theta - e^{2F_2} \cot \theta \sin^2 \theta\]  
\[(g_{kst})_{zz} = \left(e^{F_0 + F_1} + \zeta\right) \cos^2 \theta + e^{2F_1} \sin^2 \theta - \frac{\Lambda}{r} \psi \sin^2 \theta + e^{2F_2} r^2 W^2 \cos^2 \theta \sin^2 \theta + 2e^{2F_2} r W \cot \theta \sin^3 \theta + e^{2F_2} \cot^2 \sin^4 \theta + \frac{\Lambda}{r} \zeta \sin(2\theta)\]

where
\[
\zeta = e^{F_0 + F_1} - e^{2F_0} N, \]  
\[
\eta = e^{2F_1} - 2e^{F_0 + F_1} + e^{2F_0} N, \]  
\[
\xi = e^{F_0} N - e^{F_1},\]

and
\[
\psi = \frac{e^{2F_0} N \Lambda}{r}.\]
In CKSL coordinates, the lapse $\alpha$ is given by equation (5.88), the shift by

$$(\beta_{\text{CKSL}})^i = \frac{r}{\kappa} \left( e^{F_0} \Lambda \cos \theta + e^{F_1} r \sin \theta - e^{F_0} N r \sin \theta \right)$$

$$\left( -e^{F_1} r^2 W + e^{F_0} \Lambda \Xi \right) \sin \theta$$

$$e^{F_1} r \cos \theta - e^{F_0} N r \cos \theta - e^{F_0} \Lambda \sin \theta)$$

where $\kappa$ is given by equation (5.87). The volume form is given by

$$\sqrt{\gamma_{\text{CKSL}}} = e^{F_0 + 2(F_1 + F_2)} \left[ 2 e^{F_1} - e^{F_0} N - e^{F_0} \left( \frac{\Lambda}{r} \right)^2 \right],$$

where we have again assumed $\phi = 0$.

The CKSL coordinate system has a number of distinct advantages over SKSL coordinates: It lacks the coordinate singularity at $\theta = 0, \pi$ typical of spherical coordinates and it is asymptotically Minkowski, meaning we can use Minkowski space as a boundary condition in dynamical simulations. We show this in section 5.G.

### 5.5.7 Implementation Strategy

We now describe details of our implementation of the coordinate transform described in this section.
Change to SKSL Coordinates

If we numerically solve for a stationary, hairy spacetime using HRR’s construction, we will have access to the eight ansatz functions defined in equations (5.8), (5.26) (5.27), and (5.28). We can then calculate the horizon-penetrating metric and associated Proca field in SKSL coordinates via equations (5.65), (5.66), (5.67), (5.68), (5.69), (5.70), (5.71), (5.72), (5.73), (5.74), (5.82), (5.81), and (5.80). Note that during this process the variables $\Lambda$, $\Xi$, and $\Delta$ must be calculated and regularized as described in sections 5.A, 5.B, and 5.E.

Once in SKSL coordinates, we can calculate the field tensor via equations (5.155), (5.156), (5.157), (5.158), (5.159), and (5.160) and then we can perform the 3+1 split as described in section 5.5.5. Finally, we can apply Jacobian (5.107) and inverse Jacobian (5.108) to map into CKSL coordinates. Due to the complexity of the coordinate system, this last step is performed only numerically rather than analytically.

Moreover, we take advantage of the simple functional dependence on $\psi$ to calculate the metric and Proca quantities only in the $\psi = 0$ plane. If we want to calculate these quantities (which are in CKSL coordinates) at $\psi \neq 0$, we utilize the transformation

$$
\left( g_{\text{cksl}}^\psi \right)_{\mu\nu} = \left( g_{\text{cksl}}^{\psi=0} \right)_{\rho\sigma} \left( R_{-\psi} \right)_\rho^\mu \left( R_{-\psi} \right)_\sigma^\nu
$$
for the metric and

\[
\chi_{\text{skl}}^\psi = e^{im\psi} \chi_{\text{skl}}^{\psi=0} \tag{5.127}
\]

\[
(\chi_{\text{skl}}^\psi)_{\mu} = e^{im\psi} (\chi_{\text{skl}}^{\psi=0})_{\nu} (R_{-\psi})_{\mu}^{\nu} \tag{5.128}
\]

\[
(E_{\text{skl}}^\psi)_{\mu} = e^{im\psi} (E_{\text{skl}}^{\psi=0})_{\nu} (R_{\psi})_{\mu}^{\nu}, \tag{5.129}
\]

where

\[
(R_{\psi})_{\nu}^{\mu} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & \cos(\psi) & -\sin(\psi) & 0 \\
0 & \sin(\psi) & \cos(\psi) & 0 \\
0 & 0 & 0 & 1
\end{pmatrix} \tag{5.130}
\]

is the rotation matrix for a rotation angle of $\psi$.

**Excision and Extrapolation**

The physical black hole singularity is present at the origin of the spacetime both in HRR’s coordinate system and in KSL coordinates. We avoid the singularity via standard excision techniques. More formally, we choose a radius\(^{11}\) $0 < r_e < r_h$, where we terminate simulation. ($r_h$ is the horizon radius.) For hyperbolic formulations of the Einstein equations, all characteristics in the black hole interior move inward and thus no bound-

\(^{11}\)Since our evolution domain is actually Cartesian, we actually use a discrete approximation of an ellipse that best-fits the horizon, as described in [200] and [180, 179].
ary conditions at the excision surface are necessary; error at the excision surface cannot contaminate the rest of the domain [196, 33].

HRR’s construction described in section 5.4 describes the black hole exterior at all times $t$. However, we need data on a single spacelike hypersurface at $t = t_0$ on the domain $[r_e, \infty)$, where $r_e$ is the excision radius. The black hole exterior contains a Cauchy surface. So, formally, we could extract the black hole interior by evolving forward in time. However, thanks to the causal structure of the spacetime, we can get away with a much simpler approach. We may approximate the black hole interior by simple extrapolation. This is obviously unphysical. However, so long as the first two derivatives of the metric are continuous, this unphysical solution will not contaminate the black hole exterior. We summarize the situation in figure 5.6.

We have found that the Proca fields are best approximated not by polynomial extrapolation, but by “odd parity” extrapolation. I.e., if $A_{\mu}(\tau, r, \theta, \psi)$ are the components of the Proca one-form, we demand that

$$A_{\mu}(\tau, -(r - r_h), \theta, \psi) - A_{\mu}(\tau, r_h, \theta, \psi) = - [A_{\mu}(\tau, r - r_h, \theta, \psi) - A_{\mu}(\tau, r_h, \theta, \psi)],$$

(5.131)

which guarantees that the zeroth and first derivatives of the Proca fields are continuous at the horizon $r_h$.

---

12In practice, the presence or absence of boundary conditions at the excision surface depends on the formulation. Some authors have needed to impose “in-fall” type boundary conditions to ensure stability. See [5] for a review of the literature.
We have data for the exterior $t = \text{constant}$ in HRR’s coordinates and $\tau = \text{constant}$ in SKSL coordinates. The constant $t$ slice for HRR’s coordinates does not extend beyond the horizon $\mathcal{H}$. The constant $\tau$ slice for SKSL coordinates contains extrapolated information within the horizon but does not extend beyond the excision radius $r_e$. Both slices are flat at spacelike infinity $i^0$. For simplicity, we present the $\Omega_{HH} = 0$ case.

**Notation**

When they numerically search for their hairy black hole solutions, HRR rescale the Proca field such that

$$G_{\mu\nu} = 2T_{\mu\nu}$$

(in units of $G = 1$). To rectify this, we must multiply the Proca field by a factor of

$$\frac{1}{\sqrt{4\pi}}$$

(5.132)
to rescale it with respect to the other quantities. Finally, note that we will often abuse notation and map

\[ \tau \rightarrow t \quad \text{(5.133)} \]

\[ \psi \rightarrow \phi \quad \text{(5.134)} \]

since this matches the notation within standard numerical relativity codes.

## 5.6 Validation of our Coordinate Transformation

Here we check the validity of our coordinate transform. We examine our coordinates both for vacuum and hairy black hole solutions.

### 5.6.1 Vacuum Solutions

Equations (5.13), (5.14), (5.15), and (5.16) give us the analytic form for the ansatz functions \( F_i \) and \( W \) for vacuum spacetimes. Therefore, we can study our Kerr-Schild-like coordinates for exact vacuum solutions. In the \( \Omega_H = 0 \) limit,

\[ F_i = W = \Lambda = \Xi = 0 \quad \forall \ i = 0, 1, 2 \]
and the SKSL metric reduces to the Schwarzschild metric in Kerr-Schild form:

\[
\lim_{\Omega_H \to 0} (g_{\text{SKSL}})_{ij} = \begin{pmatrix}
-N & 1 - N & 0 & 0 \\
1 - N & 2 - N & 0 & 0 \\
0 & 0 & r^2 & 0 \\
0 & 0 & 0 & r^2 \sin^2 \theta
\end{pmatrix},
\]

where \( N = 1 - r_h/r \) and \( r_h = 2M \). This justifies our name “Kerr-Schild-like.” Then the CKSL metric indeed become the Schwarzschild metric in Cartesian Kerr-Schild form.

On the other hand, when \( \Omega_H > 0 \), our coordinates are only Kerr-Schild like. Figure 5.7 shows the lapse \( \alpha \) and the square root of the determinant of the 3-metric \( \sqrt{\gamma} \) for CKSL coordinates for a vacuum black hole spacetime with \( M = 0.1945 \) and \( a = 0.95M \) in code units. Compare to figure 5.2 and note that the lapse and volume form are both now finite and strictly positive near the horizon. In true Kerr-Schild coordinates, the volume form is unity near the horizon, and we don’t attain this. We plot the shift for the same example in figure 5.8. Note that the shift contains a small, but non-zero \( \theta \) component. This is distinct from traditional Kerr-Schild coordinates, where the shift contains only radial and azimuthal components.
Figure 5.7: Example (a) lapse $\alpha$ and (b) volume form $\sqrt{\gamma}$ for CKSL coordinates for a vacuum black hole spacetime with $M = 0.1945$ and $a = 0.95M$ in code units. Note that the lapse and volume form are both finite and strictly positive near the horizon (inner boundary). The outer boundary of the plot is $r = \infty$. $\alpha \to 1$ and $\sqrt{\gamma} \to 1$ as $r \to \infty$, which is consistent with the fact that the coordinates are asymptotically Minkowski.

5.6.2 Hairy Black Hole Solutions

As a first simple demonstration of our coordinate system, we present the components $\Xi$, $\Xi_i$, and $E_i$ of the Proca field after the 3+1 split (discussed in section 5.5.5) in CKSL coordinates. Figure 5.9 shows these components for example data with $r_h = 0.2475$ and $\omega = 0.9775$ provided by HRR in [116]. We do not know what to expect from a Proca evolution, so the validations we can perform are minimal. However, we can check that the constraints converge to zero. Unfortunately, this requires access to several resolutions of solution to the Einstein-Proca equations, which we have not yet independently attained. This calculation is in-progress.
Figure 5.8: The shift $\beta$ in CKSL coordinates for a vacuum black hole spacetime with $M = 0.1945$ and $a = 0.95M$ in code units. The quiver plot is the component of the shift in-plane, while the color map shows the out-of-plane component. (Note the non-compactified $x$ and $z$ coordinates.)

5.7 Next Steps

To summarize, we have presented HRR’s construction of stationary Kerr black holes with Proca hair and we have developed horizon-penetrating coordinate system compatible with HRR’s ansatz. These efforts are the first preliminary steps required to study the nonlinear stability of HRR’s construction under perturbations. The next step, which we are currently implementing, is to take HRR’s hairy black holes, take a $\tau =$-constant slice, perturb it, and use it as initial data for a numerical relativity simulation.
5.8 Acknowledgements

The authors thank Stephen Green, Nester Ortiz, Jorge Alejandro Lopez, Hugo Roussille, Hugo Marrochio, Olivier Sarbach, and especially Erik Schnetter and Luis Lehner for many helpful discussions.

We are grateful to the countless developers contributing to software projects (both open and closed source) on which we relied in this work: gh3d2m [180, 179], Data Vault [178], HDF5 [214], Python [189], and numpy and scipy [221, 132], and Matplotlib [126].

The authors acknowledge support from the Natural Sciences and Engineering Research Council of Canada and from the National Science Foundation of the USA (OCI 0905046, PHY 1212401). Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Research and Innovation.

This research used computing facilities at the Perimeter Institute, the Center for Computation & Technology at LSU, the Shared Hierarchical Academic Research Computing Network (SHARCNET) and Compute Canada, the Extreme Science and Engineering Discovery Environment (XSEDE) (TG-PHY100033).
Figure 5.9: Components of the Proca field in CKSL coordinates after the 3+1 split. We use example data with $r_h = 0.2475$ and $\omega = 0.9775$ provided by HRR in [116]. We set $\psi = \tau = 0$. The inner boundary is the event horizon, the outer boundary is asymptotic infinity. In the vector plots, the colormap represents the $y$-direction, while the arrows are in-plane. **Left column:** real part. **Right column:** imaginary part. **First row:** $\chi$, the scalar potential. **Second row:** $\chi$, the 3-dimensional potential one-form. **Third row:** $E_x$, the electric field one-form.
Appendix

5.A On the Regularity of \( \Lambda \)

Recall that \( \Lambda \) is defined by equation (5.47) to be

\[
\Lambda = \int \frac{\partial \theta e^{F_1 - F_0}}{N} \, dr = \int \frac{r \partial \theta e^{F_1 - F_0}}{r - r_h} \, dr,
\]

which will be regular if

\[
\lim_{r \to r_h} \frac{\partial}{\partial \theta} e^{F_1 - F_0} = \sum_{i=1}^{\infty} a_i(\theta)(r - r_h)^i
\]

(5.136)

for some coefficients \( a_i, i \in \mathbb{N} \), which may depend on \( \theta \), so that

\[
\lim_{r \to r_h} \Lambda = \int r \sum_{i=1}^{\infty} a_i(r - r_h)^{i-1} \, dr.
\]

(5.137)

Although we do not know what \( F_1 \) and \( F_0 \) are outside the horizon, we know what it needs to be at the horizon. Since the spacetime is stationary and since there is no Proca flux through
the horizon, the interior of the black hole can be assumed to be vacuum. Therefore, we can use equations (5.13), (5.14), and (5.15) to inform the $\theta$ dependence of $e^{F_1-F_0}$.

We find that, in vacuum,

$$
\left(e^{F_1-F_0}\right)_{\text{vacuum}} = \frac{1}{r^2} \sqrt{(b + r)^2 + b(b + r_h)^2 - br(r - r_h)(b + r_h)\sin^2 \theta} \quad (5.138)
$$

and

$$
\left(\partial_\theta e^{F_1-F_0}\right)_{\text{vacuum}} = -(r - r_h) \frac{b(b + r_h)\cos \theta \sin \theta}{r \left[(b + r)^2 + b(b + r_h)^2 - br(r - r_h)(b + r_h)\sin^2 \theta\right]^{1/2}}. \quad (5.139)
$$

Note that although there are factors of $(r - r_h)$ in the denominator, near the horizon equation (5.139) reduces to

$$
\lim_{r \to r_h} \left(\partial_\theta e^{F_1-F_0}\right)_{\text{vacuum}} = -(r - r_h) \frac{b(b + r_h)}{r_h \sqrt{r_h^2 + 3br_h + 2b^2}} \cos \theta \sin \theta, \quad (5.140)
$$

which is linear in $(r - r_h)$ as desired.

### 5.B On the Regularity of $\mathcal{F}$ and $\Xi$

Recall from equation (5.61) that $\mathcal{F}$ is

$$
\mathcal{F} = W_{\Lambda} - \Xi.
$$
One way to show that it is regular is to show that

$$\Xi = \int \frac{\partial_b \left( W e^{F_1 - F_0} \right)}{N} dr$$

is regular. Then, since we know $\Lambda$ is regular, regularity will be guaranteed. We can play the same game we played in section 5.A.

Equations (5.13), (5.14), and (5.15) tell us that, in vacuum,

$$(W e^{F_1 - F_0})_{\text{vacuum}} = \frac{(b + r)(2b + r_h)(b(b + r_h)((b + r)^2 + b(b + r_h))^2 - br(r - r_h)(b + r_h) \sin^2 \theta)^{1/2}}{r^2 \left(2b^2 + r^2 + b(2r + r_h))^2 - br(r - r_h)(b + r_h) \sin^2 \theta \right)}.$$

which has derivative

$$\left( \partial_\theta W e^{F_1 - F_0} \right)_{\text{vacuum}} = \frac{[b(b + r_h)]^{3/2} (b + r)(2b + r_h)(r - r_h) \cos \theta \sin \theta}{r \left([(b + r)^2 + b(b + r_h))^2 - br(r - r_h)(b + r_h) \sin^2 \theta \right)^{3/2}}.$$

Near the horizon, this becomes

$$\lim_{r \to r_h} \left( \partial_\theta W e^{F_1 - F_0} \right)_{\text{vacuum}} = (r - r_h) \frac{b}{2br_h + r_h^2} \sqrt{\frac{b}{(b + r_h)(2b + r_h)^2}} \cos \theta \sin \theta,$$

which means that we can write

$$\lim_{r \to r_h} \frac{\partial}{\partial \theta} W e^{F_1 - F_0} = \sum_{i=1}^{\infty} b_i(\theta)(r - r_h)^i.$$
for some coefficients $b_i$, $i \in \mathbb{N}$, which may depend on $\theta$ so that

$$\lim_{r \to r_h} \Xi = \int r \sum_{i=1}^{\infty} b_i (r - r_h)^{i-1} \, dr,$$

which is regular. Thus, since $\Lambda$ and $\Xi$ are regular, $\mathcal{B}$ is regular too.

### 5.C The Asymptotic Limit of SKSL Coordinates

Here investigate the limit of $g_{sksl}$ as $r \to \infty$. Recall from section 5.4.2 that

$$\lim_{r \to \infty} F_i = \lim_{r \to \infty} W = \lim_{r \to \infty} V = \lim_{r \to \infty} H_j = 0 \forall i \in \{0, 1, 2\}, \ j \in \{1, 2, 3\}. \quad (5.145)$$

This gives us some hint that $\Lambda$ and $\Xi$ are well-behaved at infinity. However, to actually check this, we rely on the vacuum ansatz given by equations (5.13), (5.14), (5.15), and (5.16). In section 5.A, we derived that the integrand for $\Lambda$ in vacuum is given by equation (5.139):

$$\left( \partial_\theta e^{F_1 - F_0} \right)_{\text{vacuum}} = -(r - r_h) \frac{b(b + r_h) \cos \theta \sin \theta}{r \left[ (b + r)^2 + b(b + r_h)^2 - br(r - r_h)(b + r_h) \sin^2 \theta \right]^{1/2}},$$

so that, in the vacuum, asymptotic limit, $\Lambda$ becomes

$$\lim_{r \to \infty} \Lambda_{\text{vacuum}} = - \ln(r) \frac{\sqrt{2}b^2}{\sqrt{2 - b^2 + b^2 \cos 2\theta}} \cos \theta \sin \theta, \quad (5.146)$$
which is logarithmically divergent.

On the other hand, equation (5.16) tells us that in vacuum

\[
\lim_{r \to \infty} W_{\text{vacuum}} \sim \frac{1}{r^3}
\]

so that (as given by equation (5.141))

\[
\lim_{r \to \infty} \left( \partial_\theta W e^{F_1 - F_0} \right)_{\text{vacuum}} \sim \frac{1}{r^5}
\]  

(5.147)

\[
\lim_{r \to \infty} W \Lambda = \lim_{r \to \infty} \Xi = \lim_{r \to \infty} \mathcal{B} = 0.
\]  

(5.148)

Then in this limit,

\[
\lim_{r \to \infty} g_{\text{sksl}} = \begin{pmatrix}
-1 & 0 & \Lambda & 0 \\
0 & 1 & 0 & 0 \\
\Lambda & 0 & r^2 & 0 \\
0 & 0 & 0 & r^2 \sin^2 \theta
\end{pmatrix},
\]

(5.149)

which is Minkowski space in spherical coordinates with a shift in \( \theta \) that grows logarithmically with \( r \). In other words, it is a coordinate system co-rotating in the \( \theta \) direction. This is a bit strange and a little troubling. However, as we shall see in section 5.G, this problematic behaviour vanishes in Cartesian coordinates.
5.D On the Regularity of $\Sigma$

Here we show that $\Sigma$, defined as

$$\Sigma = \frac{V + WH_3 \sin \theta}{N},$$

is regular. Boundary condition (5.34) guarantees that, near the horizon

$$\lim_{r \to r_h} [V + WH_3 \sin \theta] = \sum_{i=1}^{\infty} c_i(\theta)(r - r_h)^i$$  \hspace{1cm} (5.150)

for some coefficients $c_i$, $i \in \mathbb{N}$, which may depend on $\theta$. So,

$$\lim_{r \to r_h} \Sigma = r \sum_{i=1}^{\infty} c_i(r - r_h)^{i-1},$$  \hspace{1cm} (5.151)

thus guaranteeing that the entirety of $\tilde{A}_0$ is regular.

5.E On the Regularity of $\Delta$

We now show that $\Delta$, defined as

$$\Delta = \int e^{F_1 - F_0} \frac{\Omega_H - W}{N} dr$$

is regular. To do so, we show that the integrand is regular, thus ensuring that the integral is regular as well. We can use the same tricks as in sections 5.A and 5.B. Equation (5.12) tells us that, near the horizon,

\[
\lim_{r \to r_h} (\Omega_H - W) = \sum_{i=1}^{\infty} d_i (r - r_h)^i
\]

so that

\[
\lim_{r \to r_h} \frac{\Omega_H - W}{N} = r \sum_{i=1}^{\infty} d_i (r - r_h)^{i-1}.
\]

Then, near the horizon, \(\Delta\) becomes

\[
\lim_{r \to r_h} \Delta = \int re^{F_1 - F_0} \sum_{i=1}^{\infty} d_i (r - r_h)^{i-1} dr,
\]

which is manifestly regular.
5.F Calculating the Field Tensor

We can utilize the symmetries of $\mathcal{A}_{\text{sksl}}$ defined in equation (5.80) to help calculate $\tilde{F}_{\text{sksl}}$:

\begin{align}
\left(\tilde{F}_{\text{sksl}}\right)_{r\tau} &= - \left(\tilde{F}_{\text{sksl}}\right)_{\tau r} = - \left[ i \omega \left(\tilde{\mathcal{A}}_{\text{sksl}}\right)_r + \partial_r \left(\tilde{\mathcal{A}}_{\text{sksl}}\right)_\tau \right] \quad (5.155) \\
\left(\tilde{F}_{\text{sksl}}\right)_{r\theta} &= - \left(\tilde{F}_{\text{sksl}}\right)_{\theta r} = - \left[ i \omega \left(\tilde{\mathcal{A}}_{\text{sksl}}\right)_\theta + \partial_\theta \left(\tilde{\mathcal{A}}_{\text{sksl}}\right)_r \right] \quad (5.156) \\
\left(\tilde{F}_{\text{sksl}}\right)_{r\psi} &= - \left(\tilde{F}_{\text{sksl}}\right)_{\psi r} = - \left[ i \omega \left(\tilde{\mathcal{A}}_{\text{sksl}}\right)_\psi + m \left(\tilde{\mathcal{A}}_{\text{sksl}}\right)_r \right] \quad (5.157) \\
\left(\tilde{F}_{\text{sksl}}\right)_{\theta r} &= - \left(\tilde{F}_{\text{sksl}}\right)_{r\theta} = \left[ \partial_r \left(\tilde{\mathcal{A}}_{\text{sksl}}\right)_\theta - \partial_\theta \left(\tilde{\mathcal{A}}_{\text{sksl}}\right)_r \right] \quad (5.158) \\
\left(\tilde{F}_{\text{sksl}}\right)_{\theta\psi} &= - \left(\tilde{F}_{\text{sksl}}\right)_{\psi\theta} = \left[ \partial_\theta \left(\tilde{\mathcal{A}}_{\text{sksl}}\right)_\psi - m \left(\tilde{\mathcal{A}}_{\text{sksl}}\right)_r \right] \quad (5.159) \\
\left(\tilde{F}_{\text{sksl}}\right)_{\psi r} &= - \left(\tilde{F}_{\text{sksl}}\right)_{r\psi} = \left[ \partial_r \left(\tilde{\mathcal{A}}_{\text{sksl}}\right)_\psi - m \left(\tilde{\mathcal{A}}_{\text{sksl}}\right)_\theta \right] \quad (5.160)
\end{align}

We can then utilize equation (5.102) to reconstruct the full tensor.

5.G The Asymptotic Limit of CKSL Coordinates

Recall from section 5.C that SKSL coordinates have a logarithmically divergent shift as $r \to \infty$. In CKSL coordinates, inverse Jacobian (5.108) maps

$$
\Lambda \to \frac{\Lambda}{r}
$$
in the angular directions so that these problematic terms vanish as \( r \to \infty \). Thus, in this limit,

\[
\lim_{r \to \infty} g_{\text{cksl}} = \begin{pmatrix}
-1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix},
\]

which is just Minkowski space. Thus, CKSL coordinates are asymptotically Minkowski.
Chapter 6

Computational Gravity

More is different.

P. W. Anderson [32]

Over the last several hundred pages, we have explored three topics in computational gravity and drawn connections between lattice quantum gravity and numerical relativistic astrophysics. I now take a few paragraphs to explore why these topics are connected. Obviously as the name implies, both topics require computers, but there are deeper principles at work—after all, the first numerical algorithms were performed with pen (or perhaps quill) and paper, not a computer.\(^1\) Lattice quantum gravity and numerical relativity have two important traits in common: (1) the need to wrestle with the gauge freedom of general relativity and (2) the care one must take when matching a discrete system to a continuum one.

\(^1\)For example, see [81] for a brief history on Newton’s method for root finding. It was discovered earlier than Newton.
6.1 Gauge Freedom

Practitioners of both classical and quantum gravity must wrestle with the gauge freedom of general relativity. Path integral formulations in quantum gravity must wrestle with the ambiguity that emerges from this gauge freedom. Classically, one must choose a gauge to work in and the best gauges for initial value problems are different than the best gauges for boundary value problems. Therefore, when one wishes to connect these two classes of problem, one may need to perform a coordinate transformation.

As alluded to in chapters 2 and 3, one major motivation for CDT is the fact that the measure $d\mu(g)$ in the gravitational path integral

$$A[\gamma] = \int_{g|\partial M = \gamma} d\mu(g) e^{iS_{cl}[g]}$$

(6.1)

is not known. The problem is the ambiguity of the metric $g$—many metrics correspond to the same physical geometry. By replacing a continuum geometry with a piecewise flat one, curvature can be encoded in the connectivity of a dual graph. In this coordinate-free representation, the metric ambiguity is no longer present. This ambiguity in the metric is, of course, nothing more than the gauge freedom of general relativity. This discussion is not unique to CDT. Other discrete theories of quantum gravity are motivated by the same measure problem and resolve it by similar techniques. Indeed, understanding and wrestling with this gauge freedom is one of the major themes of quantum gravity research [203].

Studies of relativistic astrophysics must also contend with this gauge freedom, albeit
not in the same way. To formulate the Einstein equations as a well-posed initial value problem, it is necessary (but not sufficient) to choose gauge conditions that ensure that, for hyperbolic formulations, characteristic perturbations of the system propagate at finite and bounded coordinate velocity \[192\]. Chapter 4 is devoted to developing a discretization scheme compatible with one particularly successful formulation of the Einstein equations, the BSSN formulation \[199, 42, 53, 3, 7\].

Not all problems are best formulated as initial-value problems. The no hair theorems assert that (under certain assumptions) all stationary black hole spacetimes are completely described by three numbers: the mass, angular momentum, and charge of the black hole. If one seeks to evade these theorems, one searches for time-independent solutions to the Einstein equations. This is an elliptic boundary-value problem.

In \[115\], Herdeiro, Radu, and Rúnarsson solve this elliptic problem for black holes with massive vector hair. However, to do so, they choose a coordinate system ill-suited for initial-value problems. I am in the process of simulating time-dependent perturbations to Herdeiro, Radu, and Rúnarsson’s hairy black holes, which requires I first translate the solutions into a form compatible with GR as an initial-value problem. Chapter 5 is devoted to this necessary first step.

\[2\] Of course, the Einstein equations do not need to be formulated as a hyperbolic system. One may choose a gauge that makes them either elliptic or parabolic. In the former case, they become a boundary-value problem, rather than an initial-value problem.
6.2 Discretuum and Continuum

When one discretely approximates a partial differential equation, it is tempting to believe that only one choice has been made—one discretization scheme. In fact, one has made at least two choices: how to discretely approximate the differential equations, and how to approximate their solution. These choices are not necessarily compatible and this potential ambiguity is characteristic not only of partial differential equations, but of any continuum system approximated by a finite number of degrees of freedom.

Consider system

\[ R^i[u^i, \partial_t u^i, \partial_j u^i, \partial_j \partial_k u^i] = \partial_t u^i - L^i[u^i, \partial_j u^i, \partial_j \partial_k u^i] = 0 \]  \hspace{1cm} (6.2)

where \( u^i \) is a collection of variables and \( L^i \) is a nonlinear operator that acts on \( u \) and its spacelike gradients. Further suppose that \( L^i \) is constructed such that, subject to appropriate initial and boundary conditions, equation (6.2) forms a well-posed initial value problem. We wish to approximate equation (6.2) and its solutions \( u^i \) with a discrete, finite number of degrees of freedom \( N \).

Three properties are necessary for a successful discrete approximation of a well-posed initial-value problem in the continuum: consistency, convergence, and stability. A discrete approximation \( \tilde{R}^i \) is consistent with \( R^i \) if it converges to \( R^i \) as \( N \to \infty \). A discrete approximation \( \tilde{R}^i \) is convergent if the discrete solutions \( \tilde{u}^i \) of \( \tilde{R}^i \) converge to the solutions
$u^i$ of $\mathcal{R}^i$ as $N \to \infty$. A discrete approximation $\bar{\mathcal{R}}^i$ is stable if the infinity norm\textsuperscript{3} $|\bar{u}|_\infty$ of discrete solutions $\bar{u}^i$ grows no more quickly than exponentially in time, with a rate that is bounded independent of the number of degrees of freedom $N$. More simply: consistency is the agreement of the discrete approximation of the differential operator with the continuum system, convergence is the agreement of the discrete solutions to the initial-value problem with the continuum solutions, and stability is the well-posedness of the discrete initial value problem. In their landmark theorem, Lax and Richtmyer demonstrated that a finite-differences approximation of equation (6.2) system is convergent if and only if it is both consistent and stable. In fact, if a system possesses any two of these properties, it is guaranteed to possess the third.

In fact, a finite differences discretization possesses one of these three properties if it also possesses the other two.

The theorem I just described is only applicable to hyperbolic differential equations. However, it hints at a deeper truth: it is possible to choose a discrete approximation of a problem and a discrete approximation to solutions of that problem which are incompatible. There is, in other words, a degeneracy of discretizations. In chapter 4, the aforementioned theory is directly applicable. My collaborator and I developed a new discretization scheme, and we were careful to ensure that our discretizations of general relativity were consistent, convergent, and stable. However, I argue that Lax-Richtmyer theory is also indirectly applicable to the discussions of CDT in chapters 2 and 3. In particular, the distinction between

\textsuperscript{3}Or any appropriate norm.
consistency and convergence provides a useful lens through which to view lattice quantum gravity.

The CDT path sum

\[ A[\Gamma] = \sum_{T_{c} \mid \partial T_{c} = \Gamma} \mu(T_{c}) e^{iS_{R}[T_{c}]} \]  

is a consistent approximation of the path integral (6.1)

\[ A[\gamma] = \int_{g \mid \partial M = \gamma} d\mu(g) e^{iS_{cl}[g]}. \]

However, it is not clear that the triangulations \( T_{c} \) that dominate the path sum (6.3) converge to the spacetimes that dominate the path integral (6.1). To ensure that Wick rotation is well defined in CDT, we demand that our triangulations possess a foliation of spacelike hypersurfaces and that these spacelike hypersurfaces must all possess the same topology. It is not clear whether these restrictions are justified—i.e., that all physically meaningful quantum spacetimes will possess these properties.\(^4\) In vacuum and with positive cosmological constant, the dominant triangulations in the path sum converge to de Sitter space. This was one of the major early triumphs of CDT, indicating it may have the appropriate classical limit. However, this success is far from a formal proof. It is not known if the low-energy limit of CDT matches general relativity in more general settings.

The search for effective field theories that are well approximated by CDT can also be cast into this language. We know that the triangulations that dominate the CDT path sum

\(^4\)This is different from the gauge a ambiguity discussed in section 6.1. Rather it is a technical condition of CDT. Indeed, there is some evidence that the foliation condition can be relaxed [133].
converge to the solutions of some effective system in the continuum, but we don’t know what it is. In other words, we are searching for a continuum system with which our discrete system is consistent.

6.3 Conclusion

The deep connection between field theory (both classical and quantum) and statistical physics has long been understood. If one has a continuum of bodies, then many-body physics becomes field theory. The much-utilized converse relationship is rarely remarked upon, however: the countable approximation of a field theory is a many-body problem. It is this converse relationship that one utilizes when one regularizes a field theory with an ultraviolet cutoff or approximates a PDE system numerically. In this work, I have explored this converse relationship through three topics in computational gravity: causal dynamical triangulations, operator-based discontinuous Galerkin methods, and black holes with massive vector hair.
References


[188] L. F. Richardson. The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, pages 307–357, 1911.


