Causal Set Research Notes

Will Cunningham

Northeastern University

(Krioukov Research Group)

(Dated: September 10, 2014)

CONTENTS

I. INTRODUCTION

This project is a continuation of the work described in [2]. We hypothesize that a space-time with only dark energy has a structure which may be described by discrete events uniformly distributed across a de Sitter manifold. Two events are connected if they share a causal relation - i.e. they lie within each other's light cones. A positive cosmological constant Λ in Einstein's field equations - which corresponds to a positive dark energy density - results in a solution described by a de Sitter manifold. [**wc:** `Why asymptotically? De Sitter is the exact solution for the dark energy only case.` *-kz-*] [**kz:** `Re-worded` *-wc-*] When we add matter in the form of uniformly distributed dust, a singularity (or Big Bang) emerges naturally while the structure at late times remains the same. This connects the hubs of the networks, which are recognized as the earliest events. Subsequently this allows us to construct a complex network with physical properties which match those of our universe, as well as mathematical properties which match those of other well-studied complex networks.

The rescaled degree distributions for the pure-energy causal set have been previously studied in [2]. In this project we compare them to the same distributions for a universe with matter over a range of physical parameters, including the ratio of matter to dark energy. We go further to study the clustering, connectedness, and navigability of our universe. Studying navigability alone requires addressing how to measure geodesic distances: this remains an unsolved analytical problem. We attempt to solve this by embedding our universe in a higher spatial dimension, and we implement several validation tests to determine if this is a legitimate approach. We also would like to address the *why now* problem: why are the matter and dark energy densities currently the same order of magnitude? To study this we will look for special behaviors of the network when the rescaled time is set to unity.

II. MOTIVATION

In recent years, experimental observations have made us increasingly certain that we live in a flat universe [**wc:** `How can it be both flat and closed? The spatial curvature` K `is either zero or not.` *-kz-*] [**kz:** `Fixed. I was thinking flat refers to local geometry and closed refers to global topology, but the topology is still unknown I think.` *-wc-*] which is expanding at an accelerating rate. The ΛCDM Model

has provided a robust theoretical framework for describing the cosmological evolution of our universe, though it relies on six free parameters whose origins have not yet been elucidated. Rather than simply relinquish this issue to the anthropic principle, we aim to identify how a causal set framework might shed light on the nature of these constraints and how they may be related. Complex networks have given insight into problems in many unrelated fields in the past decade, so it is not unreasonable to assume they could be help out here as well.

In our work, we are interested in studying two degrees of freedom: the sum of the baryon and dark matter density, and the dark energy density [**wc: What is the 3rd parameter? -*kz*-**] [**kz: Edited. -*wc*-**]. The matter and energy densities both enter Einstein's equations (1), so they naturally both affect the manifold on which the causal set events lie.

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \tag{1}$$

We can instantly recognize that if these variables are free parameters, they are somehow chosen to minimize the Einstein-Hilbert action

$$S = \int \left[\frac{c^4}{16\pi G} R + \mathcal{L}_M \right] \sqrt{-g}\, \mathrm{d}^4 x \tag{2}$$

The dark energy density is manifested in the cosmological constant

$$\rho_\Lambda = \frac{8\pi G}{3} \Lambda \tag{3}$$

Starting from a quantum field theory, with the mass fields described by the Lagrangian density $\mathcal{L}_M$ we can obtain the stress-energy tensor by applying the Euler-Lagrange equations:

$$T_{\mu\nu} = -2\frac{\delta \mathcal{L}_M}{\delta g^{\mu\nu}} + g_{\mu\nu}\mathcal{L}_M \tag{4}$$

The matter density is then simply proportional to the first element of this tensor, $c^2 \rho_M = T_{00}$.

This implies that the properties of the complex network we are studying will be modified when these physical parameters are modified. In turn, if certain physical properties are extremized, so must properties of the causal set. It is our goal to search for these properties and show through variation of these parameters that there is a local (perhaps global) extremum. Conversely, if we take the physical scenario to arise from an underlying complex network, we

5

may provide an explanation for why our universe has the measured ratio of matter to dark energy, and we may even be able to offer a response to the *why now* question.

A primary consideration has to do with the form of the scale factor we will choose for the causal set of the universe. The general FLRW metric is given by

$$ds^2 = -dt^2 + R^2(t) \frac{dx^2 + dy^2 + dz^2}{\left(1 + \frac{K}{4}(x^2 + y^2 + z^2)\right)^2} \tag{5}$$

where K can have the values $\{0, \pm 1\}$ and $R(t)$ is the scale factor we want to find. We can rewrite this equation using a different parameterization as

$$ds^2 = \begin{cases} -dt^2 + R^2(t)\left(d\theta_1^2 + \theta_1^2\left(d\theta_2^2 + \sin^2\theta_2 d\theta_3^2\right)\right) &, & K = 0 \\ -dt^2 + R^2(t)\left(d\theta_1^2 + \sin^2\theta_1\left(d\theta_2^2 + \sin^2\theta_2 d\theta_3^2\right)\right) &, & K = 1 \\ -dt^2 + R^2(t)\left(d\theta_1^2 + \sinh^2\theta_1\left(d\theta_2^2 + \sin^2\theta_2 d\theta_3^2\right)\right) &, & K = -1 \end{cases} \tag{6}$$

where $\theta_2 \in [0, \pi]$, $\theta_3 \in [0, 2\pi)$, and $\theta_1 \in [0, \infty)$ for K = 0, -1 and $\theta_1 \in [0, \pi]$ for K = 1.

The evolution of the FLRW metric is described by the Friedmann equation:

$$\frac{\dot{R}^2}{R^2} = \frac{\Lambda}{3} - \frac{K}{R^2} + \frac{c}{R^{3g}} \tag{7}$$

where Λ is the cosmological constant and g = 1 for matter in the form of dust. The constant c can be shown to equal

$$c = \frac{8\pi G}{3} R_0^3 \rho_M = H_0^2 R_0^3 \Omega_M = \frac{\alpha^3}{a^2} \tag{8}$$

where R_0 is the scale factor at the present time, ρ_M is the density of matter, H_0 is Hubble's constant at the present time, Ω_M is the fraction of matter in the universe, a is the de Sitter pseudoradius [**wc:** the de Sitter pseudoradius -*kz*-] [**kz:** Fixed. -*wc*-], and α is an unphysical constant used later. We can identify several explicit solutions to the Friedmann equation which each represent different physical situations:

1. Vacuum FLRW Spacetime (No Matter)

    (a) $\Lambda = 0$ (Minkowski)

$$R(t) = t \tag{9}$$

(b) $\Lambda < 0$ (Anti de Sitter)

$$R(t) = a \cos\left(\frac{t}{a}\right)$$

$$K = -1 \tag{10}$$

$$a = \sqrt{\frac{3}{|\Lambda|}}$$

(c) $\Lambda > 0$ (de Sitter) [**wc:** Let us don't confuse future generations and be careful with parentheses $\exp(t/a)$ -kz-] [**kz:** Fixed. -wc-]

$$R(t) = \begin{cases} ae^{t/a} & , & K = 0 \\ a \cosh\left(\frac{t}{a}\right) & , & K = 1 \\ a \sinh\left(\frac{t}{a}\right) & , & K = -1 \end{cases} \tag{11}$$

$$a = \sqrt{\frac{3}{\Lambda}}$$

2. FLRW Spacetime with Matter (No Dark Energy)

$$R(\psi) = \begin{cases} \left(\frac{3}{2}\sqrt{c}t\right)^{2/3} & , & K = 0 \quad t = \psi \\ c \sin^2 \psi & , & K = 1 \quad t = c(\psi - \sin\psi\cos\psi) \\ c \sinh^2 \psi & , & K = -1 \quad t = c(\sinh\psi\cosh\psi - \psi) \end{cases} \tag{12}$$

3. General FLRW Spacetime (Matter and Dark Energy) For this case we will only study the flat foliation:

$$R(t) = \begin{cases} \left(ca^2 \sin^2 \frac{3t}{2a}\right)^{1/3} & , & \Lambda < 0 \\ \left(ca^2 \sinh^2 \frac{3t}{2a}\right)^{1/3} & , & \Lambda > 0 \end{cases} \tag{13}$$

For the other foliations the solution is much more complicated.

We tried to find a function $R(\lambda, t)$ which would reduce to the case of no matter for one set of parameters and the case of no dark energy for another set of parameters, but this is an impossible task due to the conformal nature of the theory. To study the effect of the ratio of matter to dark energy we ultimately chose our variable parameter to be the pseudoradius of the manifold.

III.   NUMERICAL METHODS

A.   High-Level Tasks

The following subroutines may be accessed from the main method:

- Create and Initialize a Causal Set Network

- Load a Network into Memory via Position and/or Edge Lists

- Measure Certain Properties of any Network

- Print All Data Variables to File

- Print Benchmarking Information to File

- Free All Memory / Identify Memory Leaks

B.   Low-Level Tasks

Of all the high-level tasks mentioned above, the single most computationally intensive involves generating graphs. Certain measurement subroutines may take longer but the process of generating a graph efficiently requires a more elaborate process. The low-level tasks may be divided into five steps, with the last being performed either serially on the CPU or in parallel on the GPU:

1.   Constraining the System

The universe causal set model must be constrained before any other step. There are ten parameters constrained by six equations (14) - (19) so we have four free parameters which define a network. A network is taken to mean an ensemble of graphs, so identifying a random seed as well defines a particular graph.

$$N = \frac{\pi^2}{3} \delta a \alpha^3 \left(\sinh\left(3\tau_0\right) - 3\tau_0\right) \tag{14}$$

$$\bar{k} = \delta a^4 \bar{\kappa}\left(\tau_0\right) \tag{15}$$

$$\tau_0 = \frac{2}{3} \sinh^{-1} \sqrt{\frac{\Omega_\Lambda}{\Omega_M}} \tag{16}$$

$$1 = \Omega_\Lambda + \Omega_M \tag{17}$$

$$\Lambda = \frac{3}{a^2} \tag{18}$$

$$R_0 = \alpha \sinh^{\frac{2}{3}} \left(\frac{3}{2}\tau_0\right) \tag{19}$$

The software is prepared to take any four of these parameters which do not explicitly conflict with each other, and the other six are then taken to be dependent variables. Typically though, we use N, $\bar{k}$, δ, and a as our independent variables.

There are only two cases when solving for dependent variables does not boil down to trivial algebra. The first is when the maximum rescaled time τ_0 is a dependent variable. In this case, (14) must be solve numerically using either the Newton-Raphson or bisection method due to the sinh term. The second is when the theoretical average degrees $\bar{k}$ is begin calculated, as described by (15). This requires the numerical evaluation of a 2-dimensional integral given by (28), which is performed using the NINTLIB numerical library. This uses Monte Carlo sampling over roughly half a million points to evaluate the integral. [**wc:** `Why not to use numerical 2D integration? And why not to use a lookup table that you mention on p.12 -kz-`] [**kz:** `Since this step is performed only once for a graph, it is better to trade speed for accuracy. The lookup table is used when this step would be needed to be performed many times. -wc-`]

The derivation of (28) is straightforward (thanks to Kostia). Since the rescaled scale factor $r\left(\tau\right) = \sinh^{\frac{2}{3}}\left(\frac{3}{2}\tau\right)$ is a monotonic function of τ, we can use r as a measure of time instead of τ.

$$\bar{\kappa}\left(r_0\right) = \int_0^{r_0} \bar{\kappa}\left(r|r_0\right) \rho_r \left(r|r_0\right) \, \mathrm{d}r \tag{20}$$

[**wc:** `ρ_r(r|r_0) is missing -kz-`] [**kz:** `Fixed. -wc-`] where we use the following definitions:

9

$$r_0 = \sinh^{\frac{2}{3}}\left(\frac{3}{2}\tau_0\right) \tag{21}$$

$$\bar{\kappa}(r_0) = \frac{\bar{k}(r_0)}{\delta a^4} \tag{22}$$

$$\bar{\kappa}(r|r_0) = \frac{\bar{k}(r|r_0)}{\delta a^4} \tag{23}$$

Since $\tau \sim \rho_\tau\left(\tau|\tau_0\right) = \frac{\sinh^2\left(\frac{3}{2}\tau\right)}{\mathcal{N}(\tau_0)}$ where the normalization factor $\mathcal{N}(\tau_0) = \frac{1}{6}\left(\sinh\left(3\tau_0\right) - 3\tau_0\right)$ [wc: This is not a definition, that is corollary. -kz-] [kz: Fixed. -wc-] is determined as a corollary, we find

$$\begin{aligned}
\rho_r\left(r|r_0\right) &= \rho_\tau\left(\tau\left(r|\,\tau_0\right)\right)\frac{\mathrm{d}\tau}{\mathrm{d}r}\Big|_{\tau=\tau(r)} \\
&= \frac{r^3}{\mathcal{N}\left(\tau_0\right)}\frac{\sqrt{r}}{\sqrt{1+r^3}}
\end{aligned} \tag{24}$$

The rescaled average degree at "time" r is given by the sum of the rescaled in- and out-degrees at this same r: $\bar{\kappa}\left(r|r_0\right) = \bar{\kappa}_i\left(r|r_0\right) + \bar{\kappa}_o\left(r|r_0\right)$ which leads us to (108) and (109) from [3].

$$\zeta(r) = 2\sqrt{r}\,_2F_1\left(\frac{1}{6}, \frac{1}{2}; \frac{7}{6}; -r^3\right) \tag{25}$$

$$\bar{\kappa}_i\left(r|r_0\right) = \frac{4\pi}{3}\int_r^{r_0}\left[\zeta(x) - \zeta(r)\right]^3\frac{x^2}{\sqrt{1+x^{-3}}}\,\mathrm{d}x \tag{26}$$

By following a similar approach and using (62), (88), and (108) from [3] we obtain a similar expression for the average out-degrees:

$$\bar{\kappa}_o\left(r|r_0\right) = \frac{4\pi}{3}\int_0^r\left[\zeta(r) - \zeta(x)\right]^3\frac{x^2}{\sqrt{1+x^{-3}}}\,\mathrm{d}x \tag{27}$$

Adding the last two equations yields the desired result:

$$\bar{\kappa}(r_0) = \frac{4\pi}{3\mathcal{N}(\tau_0)}\int_0^{r_0}\int_0^{r_0}\frac{\left|\zeta(x) - \zeta(r)\right|^3 x^2 r^2}{\sqrt{1+x^{-3}}\sqrt{1+r^{-3}}}\,\mathrm{d}x\,\mathrm{d}r \tag{28}$$

[wc: Let us write it in a more symmetric form: $\bar{\kappa}(\tau_0) = \frac{4\pi}{3\mathcal{N}(\tau_0)}\int_0^{r_0}\int_0^{r_0}\frac{\left|\zeta(x)-\zeta(y)\right|^3 x^2 y^2}{\sqrt{1+x^{-3}}\sqrt{1+y^{-3}}}\,\mathrm{d}x\,\mathrm{d}y$ -kz-] [kz: Fixed. -wc-]
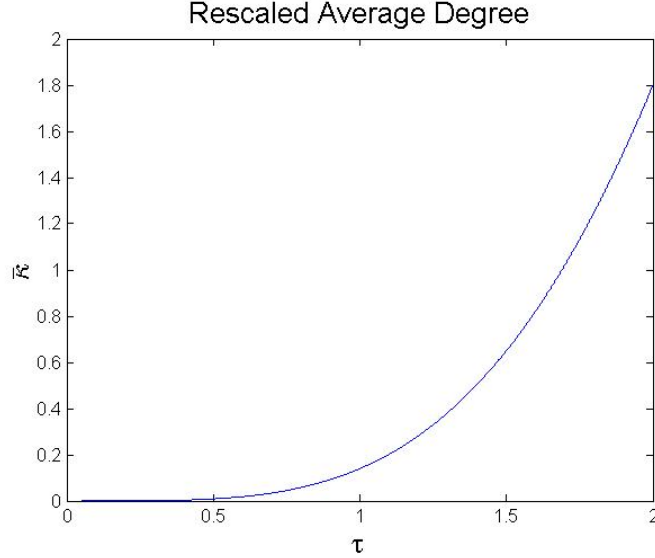
FIG. 1. Rescaled Average Degree of the Universe Causet

The relationship between κ and τ_0 is shown in Fig. 1. Since the function is monotonic in the interval it is easy to solve for τ_0 when κ is provided as an independent variable by taking advantage of a lookup table. Currently it has a resolution of 0.001 and uses linear interpolation between these points.

2. Memory

The next step requires allocating the data structures used to define a graph. A graph is defined by the coordinates of its events as well as its adjacency list. The list of coordinates takes 4 x 4 x N bytes at floating point precision. The number of in- and out-degrees associated with each node is also saved - these use 4 x N bytes each. Since edges are sparse it is more economical to use an adjacency list rather than an adjacency matrix. Both the past and future edges (associated with in- and out-degrees respectively) are saved so that any given node can be analyzed independent from the rest of the graph. Each takes 4 x (N x $\bar{k}$ / 2) bytes. Since the actual average number of edges for a graph may be above the average of the ensemble, a buffer of 20% is added. [**wc:** How do you know that 20% is enough? What will happen if the number of links exceed the buffer? *-kz-*] [**kz:** 20% is not always enough, but only in rare cases does it overflow. In

11

this case, it just prints a `warning message and the program exits.` *-wc-*] Additionally, two indices pointing to locations in the adjacency lists are associated with each node to indicate where to start for a given node (the sparse pointers). These take 4 x N bytes each. Finally, a fraction of nodes indicated by the *core_edge_ratio* parameter is allocated for an adjacency matrix so that subroutines looking for edges among the hubs (earliest events) do not suffer from long lookup times in the adjacency lists. This by default is set to 1%. Memory is allocated and deallocated in other locations in the code as well, such as in low-level GPU functions, but the memory segments described here represent the minimal requirement to store a single graph. Moreover, all of these variables remain active and accessible throughout the duration of the simulation and are not freed until the very end.

3.   The Maximum Time

One of the most basic constrained parameters is the maximum conformal and rescaled time of the graph. It serves to define the temporal boundary on the volume of the spacetime patch. Hence, if nodes are uniformly sprinkled at constant density then this maximum time may be calculated from the total number of nodes. In a 1+1 patch of spacetime without matter it is easy to derive the relation from the definition of the differential volume element

$$
\begin{aligned}
\mathrm{d}V &= (a \sec \eta)^{\mathrm{d}+1} \, \mathrm{d}\eta \, \mathrm{d}\Phi_d \\
&= a^2 \sec^2 \eta \, \mathrm{d}\eta \, \mathrm{d}\theta_3
\end{aligned}
\tag{29}
$$

[**wc:**   To be consisnent with (6), θ must be repalce by ϕ. This applies also to the next equation. *-kz-*] [**kz:** Fixed. *-wc-*]
Integration yields the volume

$$
\begin{aligned}
V &= a^2 \int_0^{\eta_0} \mathrm{d}\eta \int_0^{2\pi} \mathrm{d}\theta_3 \, \sec^2 \eta \\
&= 2\pi a^2 \int_0^{\eta_0} \sec^2 \eta \, \mathrm{d}\eta \\
&= 2\pi a^2 \tan \eta_0
\end{aligned}
\tag{30}
$$

and with some algebra this simplifies to the following relation

$$N_{1+1} = 2\pi\delta a^2 \tan\eta_0 \tag{31}$$

The same procedure may be applied to derive a similar relation in a 3+1 patch of spacetime without matter

$$dV = a^4 \sec^4\eta \, d\eta \, \sin^2\theta_1 \, \sin\theta_2 \, d\theta_1 \, d\theta_2 \, d\theta_3 \tag{32}$$

[**wc:** To be consistent with (6), ϕ is χ, χ is θ, and θ is ϕ -kz-] [**kz:** Fixed. -wc-]

Integrate the differential volume element to get the total volume

$$\begin{aligned}
V &= a^4 \int_0^{\eta_0} d\eta \int_0^{2\pi} d\theta_3 \int_0^{\pi} d\theta_2 \int_0^{\pi} d\theta_1 \, \sec^4\eta \, \sin^2\theta_1 \, \sin\theta_2 \\
&= 2\pi a^4 \int_0^{\eta_0} d\eta \int_0^{\pi} d\theta_2 \int_0^{\pi} d\theta_1 \, \sec^4\eta \, \sin^2\theta_1 \, \sin\theta_2 \\
&= 4\pi a^4 \int_0^{\eta_0} d\eta \int_0^{\pi} d\theta_1 \, \sec^4\eta \, \sin^2\theta_1 \\
&= 2\pi^2 a^4 \int_0^{\eta_0} \sec^4\eta \, d\eta \\
&= \frac{2}{3}\pi^2 a^4 \left(2 + \sec^2\eta_0\right) \tan\eta_0
\end{aligned} \tag{33}$$

and multiply by the scaled density to get the number of nodes

$$N_{3+1} = \frac{2}{3}\pi^2\delta a^4 \left(2 + \sec^2\eta_0\right) \tan\eta_0 \tag{34}$$

Recall in a 3+1 patch of spacetime with matter, the maximum rescaled time is an independent parameter. [**wc:** Why? N, $\bar{k}$, δ, and a are parameter, and τ_0 is computed from (14) -kz-] [**kz:** When $\bar{k}$ is specified, τ_0 is constrained, and vice versa -wc-] In this case, the function solves for the maximum conformal time from the maximum rescaled time. The transformation is given by equations (52), (85), and (86) in [3]:

$$\eta = \int_0^t \frac{dt'}{R(t')} \tag{35}$$

$$R(t) = \alpha \sinh^{\frac{2}{3}}\left(\frac{3t}{2a}\right) \tag{36}$$

These combine to produce

$$\eta_0 = \frac{1}{\alpha} \int_0^{a\tau_0} \sinh^{-\frac{2}{3}} \left(\frac{3t'}{2a} \right) dt' \tag{37}$$

[**wc:** dt' `is missing.` -kz-] [**kz:** `Fixed.` -wc-]
This is calculated numerically using 1-D QNG integration in the GNU Scientific Library. It may also be approximated with a power series. This technique is described in detail in the appendix.

It is often desirable to hold the average degree [**wc:** `Do you mean ''average degree''`
`?` -kz-] [**kz:** `Fixed.` -wc-] approximately constant rather than the density of nodes or the pseudoradius. In the cases with no matter (i.e. pure energy) the maximum conformal time may be solved via a transcendental equation if the average degrees is specified as an independent parameter. In a 1+1 spacetime patch the distribution of nodes in the temporal dimension is given by

$$\rho = \frac{\sec^2 \eta}{\tan \eta_0} \tag{38}$$

with the volume of the past light cone of the event at η [**wc:** `of the event at temporal`
`coordinate` η -kz-] [**kz:** `Fixed.` -wc-] written as

$$V_p = 2a^2 \ln \sec \eta \tag{39}$$

[**wc:** `A factor` $2a^2$ `is missing here and below.` -kz-] [**kz:** `Fixed.` -wc-] With some algebra the expected volume becomes [**wc:** `The expected volume` $\bar{V}_p$ `is then` -kz-] [**kz:**
`Fixed.` -wc-]

$$\begin{aligned}
V_p &= 2a^2 \int_0^{\eta_0} \frac{\sec^2 \eta}{\tan \eta_0} \ln \sec \eta \, d\eta \\
&= \frac{2a^2}{\tan \eta_0} \left[\eta + \tan \eta \left(\ln \sec \eta - 1 \right) \right]_0^{\eta_0} \\
&= \frac{2a^2}{\tan \eta_0} \left[\eta_0 + \tan \eta_0 \left(\ln \sec \eta_0 - 1 \right) \right] \\
&= 2a^2 \left(\frac{\eta_0}{\tan \eta_0} + \ln \sec \eta_0 - 1 \right)
\end{aligned} \tag{40}$$

Similarly, the volume of the future light cone of the event at η is

$$V_f = 2a^2 \left((\eta_0 - \eta) \tan \eta_0 + \ln \left(\frac{\sec \eta}{\sec \eta_0} \right) \right) \tag{41}$$

and the expected volume of the future light cone becomes [**wc:** `The expected volume:` *-kz-*] [**kz:** `Fixed.` *-wc-*]

$$
\begin{aligned}
V_f &= 2a^2 \int_0^{\eta_0} \frac{\sec^2 \eta}{\tan \eta_0} \left[(\eta_0 - \eta) \tan \eta_0 + \ln \left(\frac{\sec \eta}{\sec \eta_0} \right) \right] \mathrm{d}\eta \\
&= \frac{2a^2}{\tan \eta_0} \left[\eta_0 \tan \eta_0 \int_0^{\eta_0} \sec^2 \eta \, \mathrm{d}\eta - \tan \eta_0 \int_0^{\eta_0} \eta \sec^2 \eta \, \mathrm{d}\eta \right. \\
&\quad + \int_0^{\eta_0} \sec^2 \eta \ln \sec \eta \, \mathrm{d}\eta - \ln \sec \eta_0 \int_0^{\eta_0} \sec^2 \eta \, \mathrm{d}\eta \Big] \\
&= \frac{2a^2}{\tan \eta_0} \left[\eta_0 \tan^2 \eta_0 - \eta_0 \tan^2 \eta_0 + \tan \eta_0 \ln \sec \eta_0 \right. \\
&\quad + \eta_0 + \tan \eta_0 (\ln \sec \eta_0 - 1) - \tan \eta_0 \ln \sec \eta_0 \Big] \\
&= 2a^2 \left(\frac{\eta_0}{\tan \eta_0} + \ln \sec \eta_0 - 1 \right)
\end{aligned}
\tag{42}
$$

We now see the two expected [**wc:** `expected` *-kz-*] [**kz:** `Fixed.` *-wc-*] volumes are indeed equal

$$V_p = V_f \tag{43}$$

$$\bar{k} = \delta \left(V_p + V_f \right) \tag{44}$$

[**wc:** $2a^2$ `is not needed here.` *-kz-*] [**kz:** `Fixed.` *-wc-*] and we arrive at the final result for a 1+1 patch:

$$\bar{k} = 4\delta a^2 \left(\frac{\eta_0}{\tan \eta_0} + \ln \sec \eta_0 - 1 \right) \tag{45}$$

By taking together the two equations (31) and (45) we find the maximum conformal time may be found by solving the transcendental equation

$$\frac{\bar{k}}{N} = \frac{2}{\pi} \frac{\eta_0 / \tan \eta_0 + \ln \sec \eta_0 - 1}{\tan \eta_0} \tag{46}$$

A similar strategy may applied to a 3+1 patch of spacetime with no matter. The volume of the past light cone is given by the expression

15

$$V_p(\eta) = \int_0^\eta \mathrm{d}\eta' \int_0^{\eta-\eta'} \mathrm{d}\theta_1 \int_0^{2\pi} \mathrm{d}\theta_3 \int_0^\pi \mathrm{d}\theta_2 \, a^4 \sec^4 \eta' \sin^2 \theta_1 \sin \theta_2 \tag{47}$$

$$= 2a^4 \int_0^\eta \mathrm{d}\eta' \int_0^{\eta-\eta'} \mathrm{d}\theta_1 \int_0^{2\pi} \mathrm{d}\theta_3 \, \sec^4 \eta' \sin^2 \theta_1 \tag{48}$$

$$= 4\pi a^4 \int_0^\eta \mathrm{d}\eta' \int_0^{\eta-\eta'} \mathrm{d}\theta_1 \, \sec^4 \eta' \sin^2 \theta_1 \tag{49}$$

$$\int_0^{\eta-\eta'} \sin^2 \theta_1 \, \mathrm{d}\theta_1 = \frac{1}{2} \left[\eta - \eta' + \sin \eta \cos \eta \left(1 - 2\cos^2 \eta' \right) \right.$$
$$\left. - \sin \eta' \cos \eta' \left(1 - 2\cos^2 \eta \right) \right] \tag{50}$$

$$= 2\pi a^4 \int_0^\eta \mathrm{d}\eta' \, \sec^4 \eta' \left[\eta - \eta' + \left(1 - 2\cos^2 \eta' \right) \sin \eta \cos \eta \right.$$
$$\left. - \sin \eta' \cos \eta' \left(1 - 2\cos^2 \eta \right) \right] \tag{51}$$

$$\int_0^\eta \eta \sec^4 \eta' \, \mathrm{d}\eta' = \frac{\eta}{3} \tan \eta \left(2 + \sec^2 \eta \right) \tag{52}$$

$$\int_0^\eta \eta' \sec^4 \eta' \, \mathrm{d}\eta' = \frac{1}{3} \left[-\ln \sec^2 \eta - \frac{1}{2} \sec^2 \eta + \frac{1}{2} \right.$$
$$\left. + 2\eta \tan \eta + \eta \sec^2 \eta \tan \eta \right] \tag{53}$$

$$\int_0^\eta \left(1 - 2\cos^2 \eta' \right) \sec^4 \eta' \sin \eta \cos \eta \, \mathrm{d}\eta' = -\frac{4}{3} \sin^2 \eta + \frac{1}{3} \tan^2 \eta \tag{54}$$

$$\int_0^\eta \sin \eta' \cos \eta' \sec^4 \eta' \left(1 - 2\cos^2 \eta \right) \, \mathrm{d}\eta' = \frac{1}{2} \left(2\cos^2 \eta + \sec^2 \eta - 3 \right) \tag{55}$$

$$= \frac{2\pi a^4}{3} \left[\ln \sec^2 \eta - \sec^2 \eta + 4 - 4\sin^2 \eta + tan^2 \eta - 3\cos^2 \eta \right] \tag{56}$$

$$= \frac{2\pi a^4}{3} \left[\ln \sec^2 \eta - \sec^2 \eta + \tan^2 \eta + \cos^2 \eta \right] \tag{57}$$

The temporal distribution of nodes is described by

$$\rho = \frac{3 \sec^4 \eta}{\left(2 + \sec^2 \eta_0 \right) \tan \eta_0} \tag{58}$$

and so we obtain the average out-degrees

16

$$\bar{k}_o = \delta \int_0^{\eta_0} \rho(\eta) V_p(\eta) \, d\eta \tag{59}$$

$$= \frac{2\pi\delta a^4}{(2 + \sec^2 \eta_0) \tan \eta_0} \int_0^{\eta_0} \left[\sec^4 \eta \ln \sec^2 \eta - \sec^6 \eta + \tan^2 \eta \sec^4 \eta + \sec^2 \eta \right] d\eta \tag{60}$$

$$\int_0^{\eta_0} \sec^4 \eta \ln \sec^2 \eta \, d\eta = \frac{4}{3}\eta_0 - \frac{10}{9} \tan \eta_0 + \frac{4}{3} \tan \eta_0 \ln \sec \eta_0$$
$$- \frac{2}{9} \sec^2 \eta_0 \tan \eta_0 + \frac{2}{3} \sec^2 \eta_0 \tan \eta_0 \ln \sec \eta_0 \tag{61}$$

$$\int_0^{\eta_0} \sec^6 \eta \, d\eta = \frac{8}{15} \tan \eta_0 + \frac{4}{15} \tan \eta_0 \sec^2 \eta_0 + \frac{1}{5} \tan \eta_0 \sec^4 \eta_0 \tag{62}$$

$$\int_0^{\eta_0} \tan^2 \eta \sec^4 \eta \, d\eta = -\frac{2}{15} \tan \eta_0 - \frac{1}{15} \tan \eta_0 \sec^2 \eta_0 + \frac{1}{5} \tan \eta_0 \sec^4 \eta_0 \tag{63}$$

$$\int_0^{\eta_0} \sec^2 \eta \, d\eta = \tan \eta_0 \tag{64}$$

$$= \frac{2\pi\delta a^4}{2 + \sec^2 \eta_0} \left[\frac{4}{3} \frac{\eta_0}{\tan \eta_0} - \frac{7}{9} + \frac{4}{3} \ln \sec \eta_0 - \frac{5}{9} \sec^2 \eta_0 + \frac{2}{3} \sec^2 \eta_0 \ln \sec \eta_0 \right] \tag{65}$$

Once again recognizing the two expected [**wc:** `expected -kz-`] [**kz:** `Fixed. -wc-`] volumes are equal saves some work

$$\bar{k} = 2\bar{k}_o \tag{66}$$

and we arrive at the final result:

$$\bar{k} = \frac{4}{9} \frac{\pi\delta a^4}{2 + \sec^2 \eta_0} \left[12 \left(\frac{\eta_0}{\tan \eta_0} + \ln \sec \eta_0 \right) + (6 \ln \sec \eta_0 - 5) \sec^2 \eta_0 - 7 \right] \tag{67}$$

Again, the resulting transcendental equation is this time obtained by combining (34) and (67) to get

$$\frac{\bar{k}}{N} = \frac{2}{3\pi} \frac{12 \left(\eta_0 / \tan \eta_0 + \ln \sec \eta_0 \right) + (6 \ln \sec \eta_0 - 5) \sec^2 \eta_0 - 7}{(2 + \sec^2 \eta_0)^2 \tan \eta_0} \tag{68}$$

In a pure-energy de Sitter universe we have the relation between rescaled and conformal times given by

$$\sec \eta = \cosh \frac{t}{a} \tag{69}$$

so at this point it is easy to obtain the rescaled time for pure-energy graphs.

17

4. Generating Nodes

Generating the spacetime coordinates of the nodes is the next task. In $d+1$ dimensions we draw d spatial coordinates from the surface of a unit d-sphere. In a 1+1 de Sitter graph this consists of uniformly drawing an angle θ_3 between 0 and 2π. In a 3+1 spacetime two more angles θ_1 and θ_2 are drawn from the distributions $(2/\pi)\sin^2\theta_1$ and $(1/2)\sin\theta_2$ respectively between 0 and π. [**wc:** We need to use a consistent notation for the angles throughout the notes and future paper, be it ϕ, θ, χ, or $\theta_1, \theta_2, \theta_3$. -kz-]

On the other hand, generating the temporal coordinates is different for each topology. When there is no matter, conformal times are drawn between 0 and η_0:

$$\rho_{1+1}\left(\eta|\eta_0\right) = \frac{\sec^2\eta}{\tan\eta_0} \tag{70}$$

$$\rho_{3+1}\left(\eta|\eta_0\right) = \frac{3\sec^4\eta}{\left(2 + \sec^2\eta_0\right)\tan\eta_0} \tag{71}$$

When dust matter is present rescaled times are drawn between 0 and τ_0 from the distribution

$$\rho_{3+1}\left(\tau|\tau_0\right) = \frac{6\sinh^2\left(\frac{3}{2}\tau\right)}{\sinh\left(3\tau_0\right) - 3\tau_0} \tag{72}$$

This distribution is shown in Figure 2 for the indicated parameter [**wc:** This distribution has only one parameter, τ_0 -kz-] [**kz:** Fixed. -wc-]. For each node, both its conformal and rescaled time are stored in memory to avoid redundant calculations in the measurement stages. This requires only 4 x N extra bytes of memory so the tradeoff is worthwhile.

5. Linking Nodes with the CPU

Two events are said to be connected if there is a causal relation between them. Mathematically, this means if the Lorentzian inner product $\langle \vec{x} - \vec{y}, \vec{x} - \vec{y}\rangle$ is negative then the events are time-like separated and the two nodes located at coordinates x and y are connected. If the inner product is positive the events are space-like separated and the two nodes remain disconnected. Define a four vector $\vec{x} = (\eta, \theta_1, \theta_2, \theta_3)$ with the metric signature $(-1, +1, +1, +1)$. The temporal difference is easy:
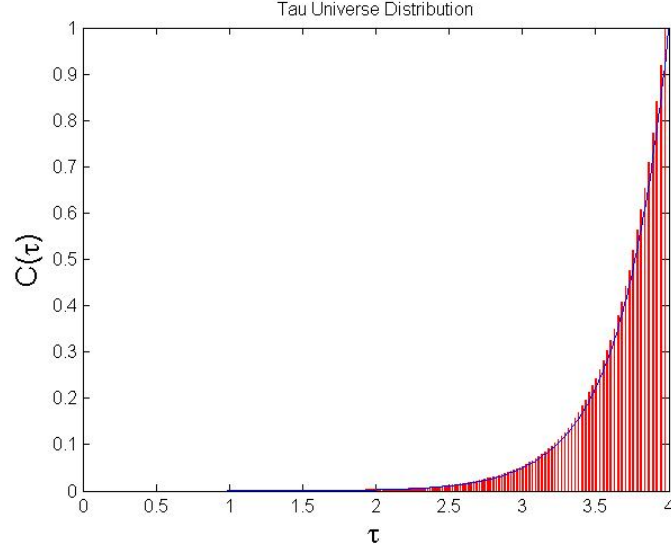
FIG. 2. Cumulative distribution function of rescaled times for $\tau = 4$. The blue curve represents the theoretical distribution.

$$\Delta \eta = |x_0 - y_0| \tag{73}$$

where η is the conformal time of an event. For a single spatial dimension, the spatial distance is defined as the shortest arc length between two points on a d-sphere. It always lies between 0 and π. In one spatial dimension this is given by

$$\Delta \theta = \pi - \left| \pi - |x_3 - y_3| \right| \tag{74}$$

and in three spatial dimensions it is found using the spherical law of cosines to be

$$\begin{aligned} \cos \left(\Delta \theta \right) = {} & \cos x_1 \cos y_1 + \\ & \sin x_1 \cos x_2 \sin y_1 \sin y_2 + \\ & \sin x_1 \sin x_2 \cos x_3 \sin y_1 \sin y_2 \cos y_3 + \\ & \sin x_1 \sin x_2 \sin x_3 \sin y_1 \sin y_2 \sin y_3 \end{aligned} \tag{75}$$

Now if the spacetime interval defined by the metric signature as $ds^2 = -d\eta^2 + d\theta^2$ is negative, two events are connected, so for every pair of nodes generated they will be linked if $\Delta \theta < \Delta \eta$.
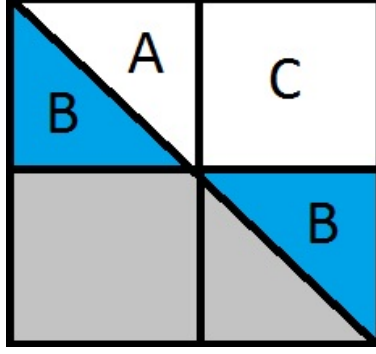
19

FIG. 3. The lower right triangle B is mapped to the upper left quadrant, creating two square matrices which can be handled much more efficiently by the GPU.

When a pair is recognized as begin causally connected, several things then happen in the algorithm. If both of the nodes are represented in the adjacency matrix, its binary value at this spot (and the one representing the reciprocal link) is set to true. After, the earlier node's out-degree counter is incremented and the later node's in-degree counter is incremented. Then, the earlier node's index is added to the list of past edges and the later node's index is added to the list of future edges. Each time a node gets it's first neighbor added to an edge list, the index of the new edge's location in the edge list is added to the respective sparse pointer list. If a node never has any past/future/any neighbors the sparse pointer values are set to -1. Nodes must be temporally ordered for this to work correctly.

After this operation, some non-zero percent of the nodes generated will be completely isolated and the rest will form connected components, often creating a giant connected component. This makes it easy to identify the resulting averge degree and number of nodes (i.e. size of giant component). It is important to have these two values be very close to their target values so that there are not large statistical fluctuations between different graphs in a network. Small fluctuations mean there is a large degree of control over the graphs produced by any given parameters, so the measured values have higher precision. The number of nodes of degree two or higher is also calculated since it's a useful statistical quantity.

6. Linking Nodes with the GPU

By performing the node-linking step in parallel, the network generation can occur tens or hundreds of times faster. The pairwise comparison may be viewed as non-zero elements of an upper triangular matrix (Fig. 3). In greater detail, splitting this matrix into four quadrants yields a triangular portion in both the upper-left and lower-right quadrants and a full square matrix in the upper-right quadrant. By mapping the lower-right quadrant's values to the lower triangle in the upper-left quadrant, we are left with two densely packed square matrices in the upper half of the original matrix, each of width *width*.

If $N(N-1)/4$ threads are launched, each thread can compare two pairs of nodes - one in [**wc:** in? -*kz*-] [**kz:** Fixed. -*wc*-] the upper left quadrant and one in the upper-right quadrant of the original matrix. For the purposes of this algorithm, threads indexed in the upper triangle of the upper-left quadrant is designated with the a label, the lower triangle of the same quadrant with the b label, and the upper right-quadrant with the c label. Hence, each thread addresses an a or b pair as well as a c pair. The method for determining the four nodes assigned to a given thread is shown below:

```
unsigned int tid = threadIdx.x;
unsigned int i = blockIdx.y;
unsigned int j = blockDim.x * blockIdx.x + threadIdx.x;
int do_map = i >= j;


unsigned int i_ab = i + do_map * (((width - i) << 1) - 1);
unsigned int j_ab = j + do_map * ((width - j) << 1);


unsigned int i_c = i;
unsigned int j_c = j + width;
```

Several of these parameters are related to the block and grid size send to the kernel. Here, there are 128 threads per block, so *tid* can range from 0 to 127. It is a 1-D thread block, so $blockDim.y = 1$ and $threadIdx.y = 0$ for all threads.

Since the thread block is linear, one of the nodes read by each thread will be identical among all threads in the block. To minimize reads to global GPU memory, this value is read

to *shared* memory by a single thread, after which the other 127 threads in the block will read it from the shared memory cache. This is demonstrated by the following code snippet:

```
__shared__ float4 shr_node0_c;
if (!tid)
  shr_node0_c = nodes[i_c];
__syncthreads();


float4 node0_c = shr_node0_c;
float4 node0_ab = do_map ? nodes[i_ab] : node0_c;
```

In this example, the threads must be synchronized so that no reads from shared memory occur before the value is written from global memory. Further, it is possible elements in the *a* triangular matrix will use the same value from shared memory, as demonstrated by the last line.

While connections will in general be sparse, there will typically be multiple edges found in a given thread block. To avoid serializing writes to global memory with the *atomicAdd* function, it is best to perform a reduction operation in the shared memory cache. The basic math involved in calculating dx and dt has been omitted.

```
__shared__ int n_a[BLOCK_SIZE];
__shared__ int n_b[BLOCK_SIZE];
__shared__ int n_c[BLOCK_SIZE];
int edge_ab = dx_ab < dt_ab;
int edge_c = dx_c < dt_c;


n_a[tid] = edge_ab * !do_map;
n_b[tid] = edge_ab * do_map;
n_c[tid] = edge_c;
__syncthreads();


int stride;
for (stride = 1; stride < BLOCK_SIZE; stride <<= 1) {
```

```
    if (!(tid % (stride << 1)) {

      n_a[tid] += n_a[tid+stride];

      n_b[tid] += n_b[tid+stride];

      n_c[tid] += n_c[tid+stride];

    }

    __syncthreads();

  }
```

This results in the total number of out-degrees for each region being stored in the inital spot of the respective arrays. This allows a single atomic operation to be performed per thread block for the out-degrees. The in-degrees must still be written by each thread, but this is guaranteed to avoid serialization (to the greatest extent possible) by the nature of how the thread block is designed.

```
  if (edge_ab)

    atomicAdd(&k_in[j_ab], 1);

  if (edge_c)

    atomicAdd(&k_in[j_c], 1);


  if (!tid) {

    if (n_a[0])

      atomicAdd(&k_out[i], n_a[0]);

    if (n_b[0])

      atomicAdd(&k_out[i_ab], n_b[0]);

    if (n_c[0])

      atomicAdd(&k_out[i_c], n_c[0]);

  }
```

Finally, the edges are written to a single edge list of 64-bit integers by storing each pair (i, j) in global memory, in a segment reserved via a final atomic operation on a global index. It is written in this manner because it is both easy to sort and it allows for contiguous writes to a single global memory segment.

```
  int idx = 0;
```

23

```
if (edge_ab | edge_c)
  idx = atomicAdd(g_idx, edge_ab + edge_c);
if (edge_ab)
  edges[idx++] = ((uint64_t)i_ab) << 32 | ((uint64_t)j_ab);
if (edge_c)
  edges[idx] = ((uint64_t)i_c) << 32 | ((uint64_t)j_c);
```

IV.   MEASUREMENTS

A.   Average Degrees

We expect the expected average degrees of a network to be a Poissonian distribution about the mean value $\langle \bar{k} \rangle$. Because points are scattered via a Poisson process, the fluctuations from the average degrees will be Poissonian in their form. For each of the sets of parameters listed below, a sample of 1000 graphs was produced and compared to the theoretical curves:

B.   Connectivity

It is of interest to understand how the connectivity of the network changes as a function of the independent variables of the model. By studying 100 graphs we obtained statistics about the number of isolated nodes, the number of connected components, and the size of the giant component.

1.   Isolated Nodes

The number of isolated nodes may be solve analytically for a 1+1 spacetime with no matter. For a poisson point process the probability function has the form

$$P(x) = \frac{(\delta V)^x}{x!} e^{-\delta V} \tag{76}$$

The probability a node is isolated at conformal time η is given by this expression, where $x = 0$ is the number of nodes [**wc:** just the number of nodes, not expected. -kz-]
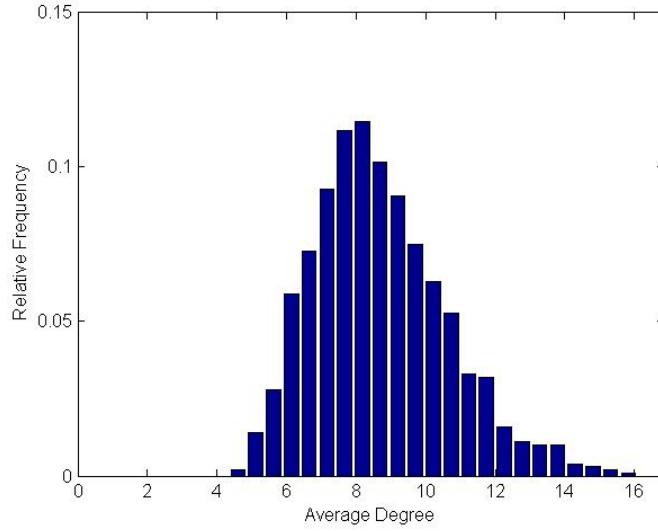
24

FIG. 4. Distribution of the resulting average degree for $N = 10240$, $\tau_0 = 4$, $a = 1$, $\delta = 1$. The target average degree was constrained by these parameters to be 12.8788.[**wc:** The expected average degree seems to be much smaller, about 9. So this plot shows that for this set of parameters we can't really generate graphs with the desired $\bar{k}$. In this plot, is it the average degree in the whole graph, or only in its giant component? (BTW, why do you use τ_0 is a parameter, and not $\bar{k}$?). It is interestig to see the plot $\bar{k}$ (target) versus expected average degree in generated graphs (both full graphs and giant components) $\langle \bar{k} \rangle$ (obtained), for fixed values of N, δ, and a. *-kz-*] [**kz:** Yes, this is correct. What I found is that the average degree does not differ much for the whole graph vs giant component – they are nearly equivalent. Also, as I mentioned earlier, τ_0 and $\bar{k}$ are interchangeable as constraints. Here I used the former to avoid confusion. I am going to add more of these graphs for different parameters to see how things scale. *-wc-*]

[**kz:** Fixed. *-wc-*] in the sum of the light cone volumes $V = V_p(\eta) + V_f(\eta)$. Manipulating this expression yields

$$P(0) = e^{-2\delta a^2 \left[(\eta_0 - \eta)\tan \eta_0 + \ln \sec^2 \eta - \ln \sec \eta_0\right]}$$

$$= e^{-2\delta a^2 [\eta_0 \tan \eta_0 - \ln \sec \eta_0]} e^{-2\delta a^2 \left[\ln \sec^2 \eta - \eta \tan \eta_0\right]} \tag{77}$$

$$= \xi e^{-2\delta a^2 \left[\ln \sec^2 \eta - \eta \tan \eta_0\right]}$$

Then, the expected number of isolated nodes is given by

$$\langle N(0) \rangle = N \int_0^{\eta_0} \rho(\eta) P(0) \, d\eta$$

$$= \frac{N\xi}{\tan \eta_0} \int_0^{\eta_0} \sec^2 \eta \, e^{-2\delta a^2 \left[\ln \sec^2 \eta - \eta \tan \eta_0\right]} \, d\eta \tag{78}$$

$$e^{-2\delta a^2 \ln \sec^2 \eta} = (\cos \eta)^{4\delta a^2}$$

$$= \frac{N\xi}{\tan \eta_0} \int_0^{\eta_0} (\cos \eta)^{4\delta a^2 - 2} e^{\left(2\delta a^2 \tan \eta_0\right)\eta} \, d\eta$$

It is possible to simplify this further by recognizing the following identity:

$$\int_0^a \cos^b x e^{cx} \, dx = \frac{\left(1 + e^{2ia}\right) e^{ac} \cos^b a}{c - ib} {}_2F_1\left(1, \frac{b}{2} + 1 - i\frac{c}{2}; 1 - \frac{b}{2} - i\frac{c}{2}; -e^{2ia}\right)$$

$$- \frac{2}{c - ib} {}_2F_1\left(1, \frac{b}{2} + 1 - i\frac{c}{2}; 1 - \frac{b}{2} - i\frac{c}{2}; -1\right) \tag{79}$$

[**wc:** exp(ac) -kz-] [**kz:** Fixed. -wc-]

so in the present case we find

$$\int_0^{\eta_0} (\cos \eta)^{4\delta a^2 - 2} e^{2\delta a^2 \eta \tan \eta_0} \, d\eta = \left[\frac{\left(1 + e^{2i\eta_0}\right) e^{2\delta a^2 \eta_0 \tan \eta_0} (\cos \eta_0)^{4\delta a^2 - 2}}{2\delta a^2 \tan \eta_0 - i\left(4\delta a^2 - 2\right)}\right.$$

$$\left. \times {}_2F_1\left(1, 2\delta a^2 - i\delta a^2 \tan \eta_0; 2 - 2\delta a^2 - i\delta a^2 \tan \eta_0; -e^{2i\eta_0}\right)\right] \tag{80}$$

$$- \left[\frac{1}{\delta a^2 \tan \eta_0 - i\left(2\delta a^2 - 1\right)}\right.$$

$$\left. \times {}_2F_1\left(1, 2\delta a^2 - i\delta a^2 \tan \eta_0; 2 - 2\delta a^2 - i\delta a^2 \tan \eta_0; -1\right)\right]$$

therefore we find the number of isolated nodes can be written as

26

$$\bar{N}_0 \sim N \frac{e^{-2\delta a^2 \eta_0 \tan \eta_0} (\cos \eta_0)^{-2\delta a^2}}{\tan \eta_0} \frac{\left(1 + e^{2i\eta_0}\right) e^{2\delta a^2 \eta_0} (\cos \eta_0)^{4\delta a^2 - 2}}{2\delta a^2 \tan \eta_0 - i\left(4\delta a^2 - 2\right)}$$

$$\times \,_2F_1 \left(1, 2\delta a^2 - i\delta a^2 \tan \eta_0; 2 - 2\delta a^2 - i\delta a^2 \tan \eta_0; -e^{2i\eta_0}\right)$$

$$\sim N \frac{\left(1 + e^{2i\eta_0}\right) (\cos \eta_0)^{2\delta a^2 - 2}}{2\delta a^2 (\tan \eta_0)^2} \,_2F_1 \left(1, 2\delta a^2 - i\delta a^2 \tan \eta_0; 2 - 2\delta - i\delta a^2 \tan \eta_0; -e^{2i\eta_0}\right)$$

$$\sim N \frac{\left(1 + e^{2i\eta_0}\right) (\cos \eta_0)^{2\delta a^2}}{2\delta a^2} \,_2F_1 \left(1, 2\delta a^2 - i\delta a^2 \tan \eta_0; 2 - 2\delta - i\delta a^2 \tan \eta_0; -e^{2i\eta_0}\right)$$

$$(81)$$

This can be reduced by implementing the Pfaff transformation: [**wc:** this is not a definition, it is a theorem, a property of the hypergeometric function. -kz-] [**kz:** Fixed. -wc-]

$$\,_2F_1 (a, b; c; z) = (1 - z)^{-a} \,_2F_1 \left(a, c - b; c; \frac{z}{z - 1}\right) \tag{82}$$

we find the relation

$$\bar{N}_0 \sim N \frac{\left(1 + e^{2i\eta_0}\right) (\cos \eta_0)^{2\delta a^2}}{2\delta a^2} \left(1 + e^{2i\eta_0}\right)^{-1}$$

$$\times \,_2F_1 \left(1, 2 - 4\delta a^2; 2 - 2\delta a^2 - i\delta a^2 \tan \eta_0; \frac{e^{2i\eta_0}}{1 + e^{2i\eta_0}}\right)$$

$$(83)$$

and by using the approximations (as $\eta_0 \to \frac{\pi}{2}$)

$$e^{2i\eta_0} \to -1; \quad 1 + e^{2i\eta_0} \to 2i \left(\frac{\pi}{2} - \eta_0\right); \quad \tan \eta_0 \to \frac{1}{\frac{\pi}{2} - \eta_0} \tag{84}$$

we arrive at the equation

$$\bar{N}_0 \sim N \frac{(\cos \eta_0)^{2\delta a^2}}{2\delta a^2} \,_2F_1 \left(1, 2 - 4\delta a^2; \frac{-i\delta a^2}{\frac{\pi}{2} - \eta_0}; \frac{-1}{2i \left(\frac{\pi}{2} - \eta_0\right)}\right)$$

$$\sim N \frac{(\cos \eta_0)^{2\delta a^2}}{2\delta a^2} \,_2F_1 \left(1, 2 - 4\delta a^2; -i\delta a^2 \tan \eta_0; \frac{i}{2} \tan \eta_0\right)$$

$$(85)$$

Since the average number of isolated nodes is known to be a real number, it makes sense to continue to reduce this expression to eliminate the complex quantities. To do this, we need to apply four new relations:

1. For $a - b \notin \mathbb{Z}$ and $x \notin (0, 1)$,

$$
\begin{aligned}
{}_2F_1\left(a, b; c; x\right) = & \frac{\Gamma\left(b - a\right)\Gamma\left(c\right)}{\Gamma\left(b\right)\Gamma\left(c - b\right)}\left(-x\right)^{-a} {}_2F_1\left(a, a - c + 1; a - b + 1; \frac{1}{x}\right) + \\
& \frac{\Gamma\left(a - b\right)\Gamma\left(c\right)}{\Gamma\left(a\right)\Gamma\left(c - b\right)}\left(-x\right)^{-b} {}_2F_1\left(b, b - c + 1; -a + b + 1; \frac{1}{x}\right)
\end{aligned}
\tag{86}
$$

[**wc:** This is true if $a - b \notin \mathbb{Z}$, and $x \notin (0, 1)$. -*kz*-] [**kz:** Added above equation. -*wc*-]

2. The Kummer hypergeometric function is [**wc:** Again, it is not a definition, it is (an elementary) property of the Kummer hypergeometric function proved in my notes. -*kz*-] [**kz:** Fixed. -*wc*-]

$$
{}_1F_1\left(a; c; bd\right) = \lim_{x \to \infty} {}_2F_1\left(a, bx; c; \frac{d}{x}\right)
\tag{87}
$$

3.

$$
\lim_{x \to \infty} \frac{\Gamma\left(x\right)}{\Gamma\left(x - a\right)} = x^a
\tag{88}
$$

4.

$$
{}_1F_1\left(1; a; b\right) = \left(a - 1\right) e^b b^{1-a}\left(\Gamma\left(a - 1\right) - \Gamma\left(a - 1, b\right)\right)
\tag{89}
$$

[**wc:** $\exp(b)$ -*kz*-] [**kz:** Fixed. -*wc*-]

By using the definitions $\alpha \equiv 2 - 4\delta a^2$, $\beta \equiv -i\delta a^2$, $\gamma \equiv \frac{i}{2}$, and $x \equiv \tan\eta_0$ and the relations

$$
\begin{aligned}
{}_2F_1\left(1, \alpha; \beta x; \gamma x\right) = & \frac{\Gamma\left(\alpha - 1\right)\Gamma\left(\beta x\right)}{\Gamma\left(\alpha\right)\Gamma\left(\beta x - 1\right)}\left(-\gamma x\right)^{-1} {}_2F_1\left(1, 2 - \beta x; 2 - \alpha; \frac{1}{\gamma x}\right) \\
& + \frac{\Gamma\left(1 - \alpha\right)\Gamma\left(\beta x\right)}{\Gamma\left(1\right)\Gamma\left(\beta x - \alpha\right)}\left(-\gamma x\right)^{-\alpha} {}_2F_1\left(\alpha, \alpha + 1 - \beta x; \alpha; \frac{1}{\gamma x}\right) \\
& \to \frac{\beta x - 1}{\left(\alpha - 1\right)\left(-\gamma x\right)} {}_1F_1\left(1; 2 - \alpha; -\frac{\beta}{\gamma}\right) \\
& + \frac{\Gamma\left(1 - \alpha\right)\left(\beta x\right)^\alpha}{\left(-\gamma x\right)^\alpha} {}_1F_1\left(\alpha; \alpha; -\frac{\beta}{\gamma}\right)
\end{aligned}
\tag{90}
$$

[**wc:** The last $=$ should be replaced by $\to$ -*kz*-] [**kz:** Fixed. -*wc*-] as well as the limit with (89), we obtain [**wc:** Using (89), we obtain -*kz*-] [**kz:** Fixed. -*wc*-]

$$
\lim_{x \to \infty} {}_2F_1\left(1, \alpha; \beta x; \gamma x\right) = e^{-\beta/\gamma}\left(-\frac{\beta}{\gamma}\right)^\alpha \Gamma\left(1 - \alpha, -\frac{\beta}{\gamma}\right)
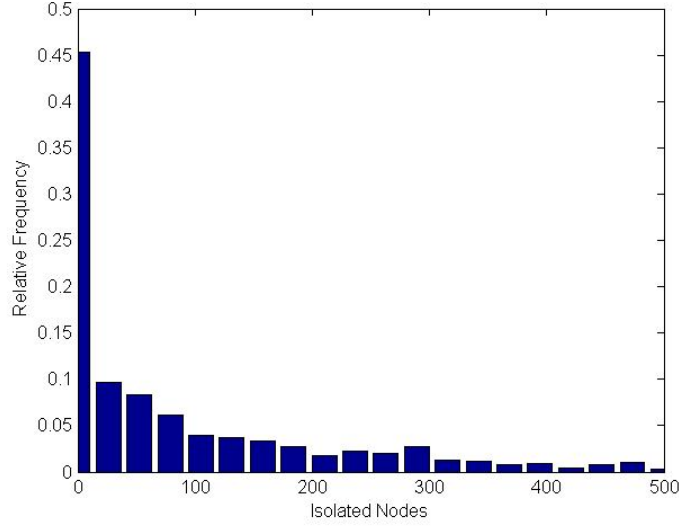\tag{91}
$$

FIG. 5. The distribution of isolated nodes taken over a sample of 1000 graphs. These graphs were constrained by the parameters $N = 10240$, $\tau_0 = 4$, $a = 1$, and $\delta = 1$. [**wc:** It seems that the expected number of isolated nodes is less than 1% of the full network. So they alone should not strongly affect the average degree. But the problem is that there are non-isolated nodes that do not belong to the giant component. Their fraction is bigger. -*kz*-] [**kz:** Yes, this is correct, but from the size of the giant component you can see most of the time the giant component strongly dominates. -*wc*-]

and finally arrive at the result

$$\bar{N}_0 \sim N \left(e \cos \eta_0 \right)^{2\delta a^2} \left(2\delta a^2 \right)^{1-4\delta a^2} \Gamma \left(4\delta a^2 - 1, 2\delta a^2 \right) \tag{92}$$

[**wc:** It is not exp, it is a number e :) -*kz*-] [**kz:** Fixed. -*wc*-]

For a 3+1 spacetime it does not appear to be possible to obtain a closed-form analogue of this solution.

For a 3+1 universe with matter, we have found the following results for the number of isolated nodes:
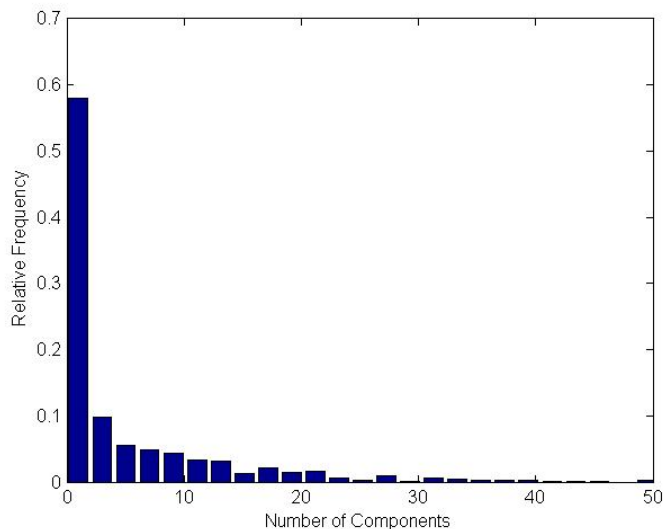
FIG. 6. The distribution of connected components taken over a sample of 1000 graphs. These were constrained by the parameters $N = 10240$, $\tau_0 = 4$, $a = 1$, and $\delta = 1$.

2. Number of Connected Components

In this project we are interested in networks above the percolation threshold, so in all cases the average degree has been manipulated to meet this condition. The number of connected components was calculated by implementing a breadth first search algorithm, assigning each node an identification number associated with a component. Beginning at the earliest node, the algorithm works by assigning each past and future neighbor the same component ID as the current node, so the algorithm acts recursively. Once each pass is completed, the second earliest node is checked to see if it has been assigned a component ID. If it has not, the recursive process is repeated. This is done for every node until every single one has a component ID.

3. Size of Giant Component

The size of the giant component is similar to the number of isolated nodes except when the system is below the percolation threshold. Since all components except for the giant component are ignored it was useful to identify any possible deviations in statistical behavior through this extra quantity. When the breadth first search algorithm is used, if it is
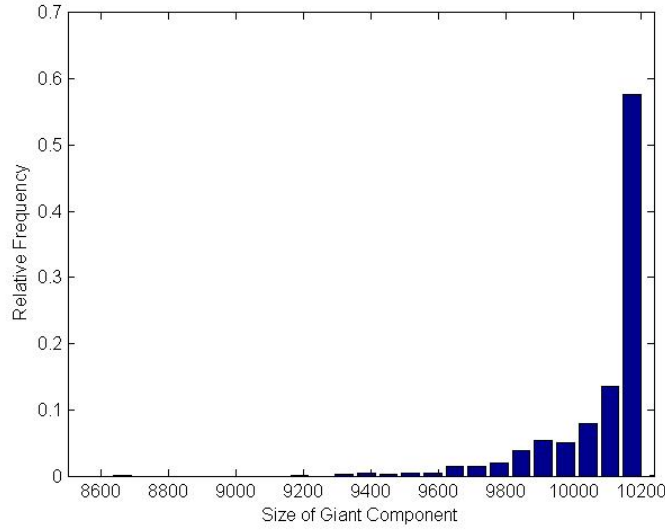
30

FIG. 7. The size of the giant component in a sample of 1000 graphs. These were constrained by the parameters $N = 10240$, $\tau_0 = 4$, $a = 1$, and $\delta = 1$. [**wc:** It would be interesting to see the plot of the expected relative size of the giant component $\langle \mathcal{G} \rangle / N$ as a function of $\bar{k}$ and a, *-kz-*] [**kz:** I'll add this. *-wc-*]

performed first on the earliest node (the largest hub) then any nodes with the same component ID as the hub are recognized as part of the giant component. Therefore by counting the number of nodes which have matching IDs it is easy to find the size of the giant component.

C. Degree Distribution

Numerical evaluation of the degree distributions for a universe with matter can be done using these equations from [3]: [**wc:** More details are needed here. Where exactly do you use these equations? In Eq. for $\bar{k}_i$, should be $\eta'' - \eta'$ insted of $\eta' - \eta''$. *-kz-*] [**kz:** Fixed the equations. They are used in a MATLAB script Dima provided. *-wc-*]
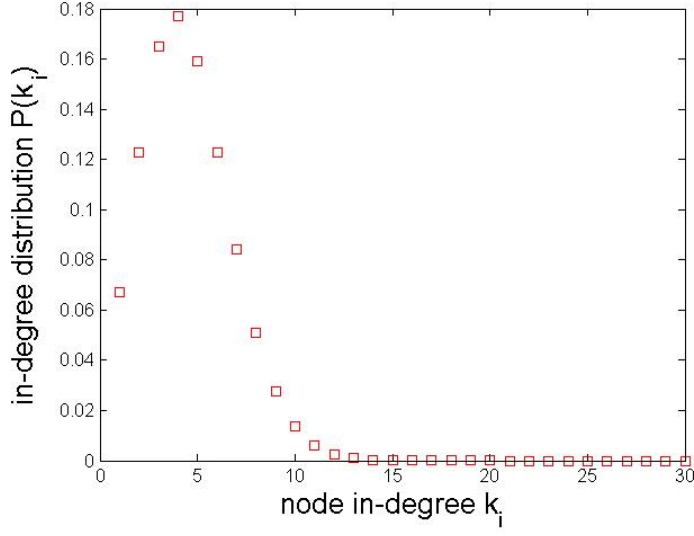
31

FIG. 8. Distribution of in-degrees observed over 1000 graphs.

$$\bar{k}_i \left(\eta' | \eta \right) = \delta \frac{\pi}{K} \int_{\eta'}^{\eta} \left\{ 2 \left(\eta'' - \eta' \right) - \frac{1}{\sqrt{K}} \sin \left[2\sqrt{K} \left(\eta'' - \eta' \right) \right] \right\} R^4 \left(\eta'' \right) \, \mathrm{d}\eta'' \qquad (93)$$

$$\bar{k}_o \left(\eta' | \eta \right) = \delta \frac{\pi}{K} \int_{0}^{\eta'} \left\{ 2 \left(\eta'' - \eta' \right) - \frac{1}{\sqrt{K}} \sin \left[2\sqrt{K} \left(\eta'' - \eta' \right) \right] \right\} R^4 \left(\eta'' \right) \, \mathrm{d}\eta'' \qquad (94)$$

$$p \left(k | \eta', \eta \right) = \frac{1}{k} \left[\bar{k} \left(\eta' | \eta \right) \right]^k e^{-\bar{k} \left(\eta' | \eta \right)} \qquad (95)$$

$$\rho \left(\eta' | \eta \right) = \frac{R^4 \left(\eta' \right)}{\mathcal{N} \left(\eta \right)} \qquad (96)$$

$$P \left(k, \eta \right) = \int_{0}^{\eta} p \left(k | \eta', \eta \right) \rho \left(\eta' | \eta \right) \, \mathrm{d}\eta' \qquad (97)$$

1. In-Degree Distribution

The in-degrees describe connections coming into a given node from an earlier node. Here we study the degree distribution for N = 10240, $\tau_0 = 4$, a = 1, and $\delta = 1$.

2. Out-Degree Distribution

Likewise, the out-degrees describe connections exiting a given node toward a node at a later time. The same parameters have been tested to produce these plots:
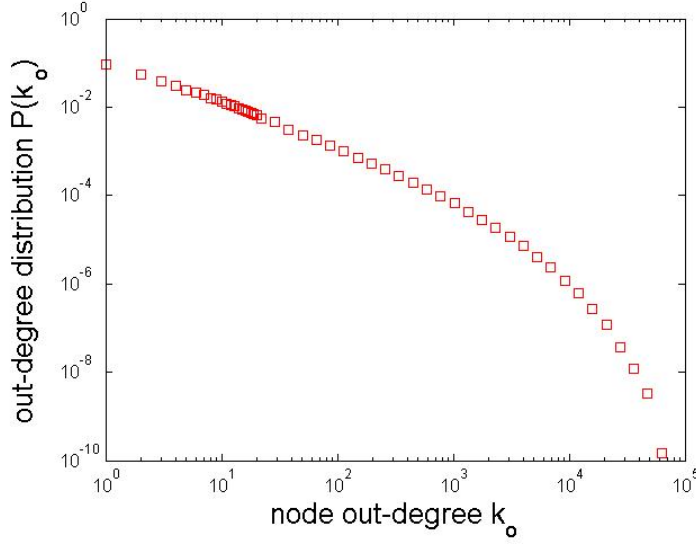
32

FIG. 9. Distribution of out-degrees observed over 1000 graphs. [**wc:** Network size is $N = 10240$. What are those squares going towards 10^5? *-kz-*] [**kz:** I believe this is summed over all 1000 graphs. *-wc-*]

D. Clustering

1. Average Clustering

2. Clustering Distribution

E. Navigability in the Universe

1. Geodesic Distances in the Universe

Measuring geodesic distances in an FLRW universe with a positive dark energy density and dust matter is currently an unsolved problem. We hypothesize that we may embed the 3+1 manifold isometrically in a higher spatial dimension so distances may be calculated using a 4+1 Lorentzian metric. This is not a problem for a simple de Sitter manifold, but the alterations to the topology resulting from the inclusion of matter may render this an incorrect approach. Here we propose a mathematical framework to justify the extension of the embedding technique to the new manifold, as well as several validation tests which may be performed in the causal set simulation.

33

We begin by considering a flat five-dimensional Minkowski space $\mathbb{M}^{4+1}$

$$ds^2 = -dz_0^2 + dz_1^2 + dz_2^2 + dz_3^2 + dz_4^2, \tag{98}$$

and the de Sitter manifold that is embedded in $\mathbb{M}^{4+1}$, $d\mathbb{S}^{3+1} \subset \mathbb{M}^{4+1}$

$$-z_0^2 + z_1^2 + z_2^2 + z_3^2 + z_4^2 = a^2, \quad a = \sqrt{3/\Lambda} \tag{99}$$

$d\mathbb{S}^{3+1}$ can be covered by the coordinates $(t, \theta_1, \theta_2, \theta_3)$. Let us rotate the system such that all the angular coordinates of a node [**wc:** of a node -kz-] [**kz:** Fixed. -wc-] are zero. Then we have

$$\begin{cases} z_0 & = a \sinh \frac{t}{a} \\ z_1 & = a \cosh \frac{t}{a} \\ z_2 & = z_3 = z_4 = 0 \end{cases} \tag{100}$$

Now we want to find a function $z_1 = f(z_0)$ such that

$$\begin{cases} z_1 = R(t) = a \cosh \frac{t}{a} \\ L\left[(x, f(x)), x \in [0, z_0] \right] = t \end{cases} \tag{101}$$

where L is the length of the curve along the de Sitter manifold in the temporal direction. In the Minkowski metric,

$$L = \int_0^{z_0} \sqrt{\left| f'(x)^2 - 1 \right|} \, dx \tag{102}$$

Therefore, using (101) we have

$$f(z_0) = R\left(\int_0^{z_0} \sqrt{\left| f'(x)^2 - 1 \right|} \, dx \right) \tag{103}$$

For $d\mathbb{S}^{3+1}$, we know that $R(t) = a \cosh \frac{t}{a}$ and $f(z_0) = \sqrt{z_0^2 + a^2}$. As a test, let us check that (103) is satisfied for this special case.

34

$$\int_0^{z_0} \sqrt{\left| f'(x)^2 - 1 \right|}\, dx = \int_0^{z_0} \sqrt{\left| \frac{x^2}{x^2 + a^2} - 1 \right|}\, dx$$

$$= \int_0^{z_0} \frac{a}{\sqrt{x^2 + a^2}}\, dx$$

$$= a \int_0^{z_0} \frac{d\frac{x}{a}}{\sqrt{\left(\frac{x}{a}\right)^2 + 1}} \tag{104}$$

$$= a \int_0^{z_0/a} \frac{dy}{\sqrt{y^2 + 1}}$$

$$= a \sinh^{-1} \frac{z_0}{a}$$

$$f(z_0) = a \cosh\left(\sinh^{-1} \frac{z_0}{a} \right)$$

$$= a\sqrt{\frac{z_0^2}{a^2} + 1} \tag{105}$$

$$= \sqrt{z_0^2 + a^2}$$

Since the scale factor $R(t)$ is monotonic, we can invert $R(t)$ [**wc:** We can invert $R(t)$, not the equation. -kz-] [**kz:** Fixed. -wc-] and obtain the transcendental differential equation

$$\frac{f'(z_0)}{R'\left(R^{-1}\left(f(z_0) \right) \right)} = \sqrt{\left| 1 - f'(z_0)^2 \right|} \tag{106}$$

For a ΛCDM spacetime we obtain from this

$$R(t) = \alpha \sinh^{\frac{2}{3}} \left(\frac{3t}{2a} \right) \tag{107}$$

$$R'(t) = \frac{\alpha}{a} \frac{\cosh\left(\frac{3t}{2a} \right)}{\sinh^{1/3}\left(\frac{3t}{2a} \right)} \tag{108}$$

$$R^{-1}(t) = \frac{2a}{3} \sinh^{-1} \left(\frac{t}{\alpha} \right)^{3/2} \tag{109}$$

$$R'\left(R^{-1}(t) \right) = \frac{\alpha}{a} \frac{\sqrt{1 + \left(\frac{t}{\alpha} \right)^3}}{\sqrt{\frac{t}{\alpha}}} \tag{110}$$

35

so we get the result

$$\frac{f'(z_0)}{\sqrt{\left|1 - f'(z_0)^2\right|}} = \frac{\alpha^{3/2}}{a} \frac{\sqrt{1 + \left(\frac{f(z_0)}{\alpha}\right)^2}}{\sqrt{f(z_0)}} \tag{111}$$

Now, we would like to obtain an analogue of this equation for the function $z_0 = f(z_1)$. We begin the derivation as before:

$$z_1 = R\left(\int_0^{z_1} \sqrt{\left|1 - f'(x)^2\right|}\, dx\right) \tag{112}$$

$$R^{-1}(z_1) = \int_0^{z_1} \sqrt{\left|1 - f'(x)^2\right|}\, dx \tag{113}$$

$$\frac{d}{dz_1} R^{-1}(z_1) = \sqrt{\left|1 - f'(z_1)^2\right|} \tag{114}$$

$$\left|1 - f'(z_1)^2\right| = \left(\frac{d}{dz_1} R^{-1}(z_1)\right)^2 \tag{115}$$

$$f'(z_1) = \sqrt{1 \pm \left(\frac{d}{dz_1} R^{-1}(z_1)\right)^2} \tag{116}$$

and we find the resulting general expression to be

$$f(z_1) = \int_0^{z_1} \sqrt{1 \pm \left(\frac{d}{dx} R^{-1}(x)\right)^2}\, dx \tag{117}$$

When the scale factor for the ΛCDM spacetime is inserted into this expression we arrive at

$$f(z_1) = z_0 = \int_0^{z_1} \sqrt{1 + \frac{a^2 x \alpha^2}{\alpha^3 + x^3}}\, dx \tag{118}$$

where we have taken the positive sign in the $\pm$ to represent the physical situation. A plot of this equation is shown in Fig. 10 for several choices of parameters.

The inner product between two vectors is always invariant under Lorentz transformations [**wc:** Counterexample: inner product is not invariant under scaling, $x \mapsto \lambda x$:-) invariant under what transformation of vectors? -kz-] [**kz:** Fixed. -wc-], so the two vectors may be rotated such that each of their spatial coordinates is zero [**wc:** becomes zero -kz-] [**kz:** Fixed. -wc-]. The scalar product of the two 5-D vectors in terms
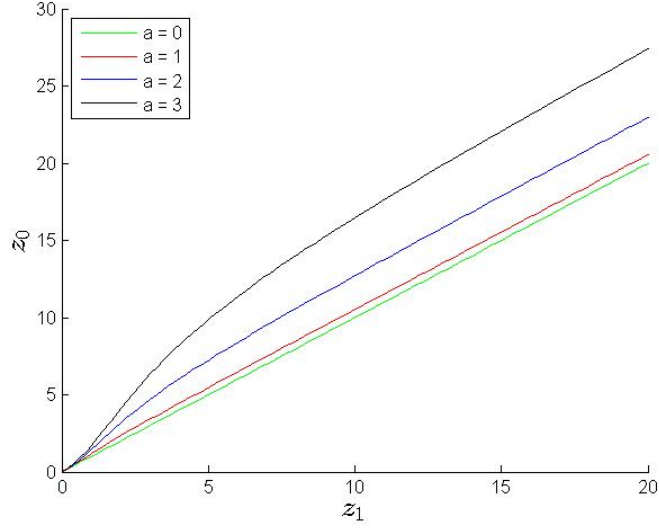
FIG. 10. Embedding function $z_0 = f(z_1)$ for the ΛCDM spacetime. [**wc:** What is α? Is $\alpha = a^{2/3}$? -kz-] [**kz:** Yes -wc-]

of the transformed coordinates is described by [**wc:** Why suddenly τ instead of t? -kz-] [**kz:** Fixed. -wc-]

$$z_1' = R(t) \tag{119}$$

$$z_0' = \int_0^{z_1'} \sqrt{1 + \frac{a^2 x \alpha^2}{\alpha^3 + x^3}} \, \mathrm{d}x \tag{120}$$

$$\langle \vec{z}_a', \vec{z}_b' \rangle = -\left(z_0'\right)_a \left(z_0'\right)_b + \left(z_1'\right)_a \left(z_1'\right)_b \tag{121}$$

When rotated into the spatial planes by the angles $(\theta_1, \theta_2, \theta_3)$, we find the scalar product becomes

$$z_0 = z_0' \tag{122}$$

$$z_1 = z_1' \cos\theta_1 \tag{123}$$

$$z_2 = z_1' \sin\theta_1 \cos\theta_2 \tag{124}$$

$$z_3 = z_1' \sin\theta_1 \sin\theta_2 \cos\theta_3 \tag{125}$$

$$z_4 = z_1' \sin\theta_1 \sin\theta_2 \sin\theta_3 \tag{126}$$

$$\langle \vec{z}_a, \vec{z}_b \rangle = -\left(z_0\right)_a \left(z_0\right)_b + \left(z_1\right)_a \left(z_1\right)_b + \left(z_2\right)_a \left(z_2\right)_b + \left(z_3\right)_a \left(z_3\right)_b + \left(z_4\right)_a \left(z_4\right)_b \tag{127}$$

The condition for calculating the distance is given by Definition 2.6 in [1]. If the vector $\vec{x} - \vec{y}$ is timelike, this means [**wc:** `Please introduce the notation used in (101) and` `(102). Also are` x `and` $\vec{x}$ `the same or different?` -*kz*-] [**kz:** `Fixed.` -*wc*-]

$$\langle \vec{x} - \vec{y}, \vec{x} - \vec{y} \rangle < 0 \tag{128}$$

$$\vec{x}^2 - 2\vec{x} \cdot \vec{y} + \vec{y}^2 = 2\left(1 - \vec{x} \cdot \vec{y}\right) \tag{129}$$

$$\vec{x} \cdot \vec{y} > 1 \tag{130}$$

and then the distance is given by

$$d(\vec{x}, \vec{y}) = \cosh^{-1}(\langle \vec{x}, \vec{y} \rangle) \tag{131}$$

Geometrically this is the case where points lie on a great hyperbola. If the vector $\vec{x} - \vec{y}$ is lightlike (which does not occur for randomly distributed points) the distance is identically zero. If the inner product $\langle \vec{x}, \vec{y} \rangle \leq -1$ then the distance is infinite since the plane defined by the two vectors intersects the de Sitter space is timelike, with the two points on disconnected components. If the vector $\vec{x} - \vec{y}$ is spacelike on the condition that $\langle \vec{x}, \vec{y} \rangle \geq -1$ then

$$\vec{x} \cdot \vec{y} < 1 \tag{132}$$

and the distance is given by

$$d(\vec{x}, \vec{y}) = \cos^{-1}(\langle \vec{x}, \vec{y} \rangle) \tag{133}$$

This represents the case where the points are connected by a great ellipse.

[**wc:** `After all explanations, can we write down a ready-to-use formula for` `computing distance between two nodes` $n_a = (t_a, (\theta_1)_a, (\theta_2)_a, (\theta_3)_a)$ `and` $n_b = (t_b, (\theta_1)_b, (\theta_2)_b, (\theta_3)_b)$? `(I still prefer to use` ϕ, θ, `and` χ`)` -*kz*-] [**kz:** `I'll add this later.` -*wc*-]

2. *Validation Test for Manifold Embedding*

One method of validating this approach [**wc:** `I tried to make sense of the beginning of the sentence, but failed :) -kz-`] [**kz:** `Fixed. -wc-`] is to calculate geodesic distances is to identify if time-like and space-like intervals in 3+1 de Sitter space are also classified as such in the embedded 4+1 manifold. If two nodes are connected when the network is constructed, the distance between the two points should be time-like. If not, it should be classified as space-like. There should be no light-like vectors because nodes are randomly distributed. We can characterize the success of the embedding by defining the confusion matrix as follows:

$$C = \begin{pmatrix} \frac{S_4 \bigcap S_5}{S_4} & \frac{T_4 \bigcap S_5}{T_4} \\ \frac{S_4 \bigcap T_5}{S_4} & \frac{T_4 \bigcap T_5}{T_4} \end{pmatrix} \tag{134}$$

where S and T indicate space-like and time-like classifications respectively, and the subscripts indicate the number of dimensions used in the method with which they were classified. For a universe with no matter (pure de Sitter) we expect this to be a unit matrix, and that is exactly what we find in simulation. For the parameters $N = 10240$, $\tau_0 = 4$, $a = 1$, $\delta = 1$ and a random seed -1406627360 in the ΛCDM spacetime we find the confusion matrix is

$$C = \begin{pmatrix} 1.000 & 0.934 \\ 0.000 & 0.066 \end{pmatrix} \tag{135}$$

[**wc:** `It seems that` $T_5 \approx 0$ `and` $S_5 \approx 1$ `and the network is essentially disconnected. What is` $\bar{k}$ `in this case, maybe it is too small? We need to see how entries of this matrix change with` $\bar{k}$`. -kz-`] [**kz:** `This is something I will add. -wc-`]

This indicates when the 4-D manifold is embedded in the fifth dimension using this set of constraints, 93.4% of timelike pairs are misclassified as spacelike. Figure X [**wc:** `Where is it? -kz-`] [**kz:** `Not ready yet, I'll introduce it for a range of parameters. -wc-`] indicates how misclassified distances vary as a function of $\Delta\theta$, $\Delta\eta$, and the pseudoradius a.

Because we are using the pseudoradius as our variable parameter in the greedy routing experiment, it is useful to know the analytic expression for the asymptotic behavior of the matter and energy densities as t_0 (and therefore a) tends to zero and ∞. We begin with the following definitions which were derived earlier:

$$\rho_\Lambda = \frac{3}{8\pi G}\frac{1}{a^2} \tag{136}$$

$$\rho_M = \frac{3}{8\pi G}\left(a\sinh\frac{3\tau_0}{2}\right)^{-2} \tag{137}$$

$$t_0 = \left(\frac{\bar{k}}{\delta\bar{\kappa}}\right)^{1/4}\tau_0 \tag{138}$$

For the two densities we can drop the numerical coefficients since we are interested in asymptotic behavior only.

Statement:

1.

$$\lim_{a\to 0} t_0 = \infty \tag{139}$$

$$\lim_{a\to\infty} t_0 = \hat{t}_0 \in (0,\infty) \tag{140}$$

2.

$$\lim_{a\to 0}\rho_M = 0 \tag{141}$$

$$\lim_{a\to\infty}\rho_M = \hat{\rho}_M \in (0,\infty) \tag{142}$$

To show this we need to find the asymptotic behavior of $\bar{\kappa}(\tau_0)$ as $\tau_0 \to 0$ and $\tau_0 \to \infty$.

Lemma:

$$\bar{\kappa}(\tau_0) = \begin{cases} \mathcal{O}(\tau_0^4) & \text{as} \quad \tau_0 \to 0 \\ \mathcal{O}(\tau_0) & \text{as} \quad \tau_0 \to \infty \end{cases} \tag{143}$$

Using this fact, we can find the dependence on the pseudoradius for the remaining two quantities. For t_0 we find

40

$$t_0 = \left(\frac{\bar{k}}{\delta}\right)^{1/4} \frac{\tau_0}{\bar{\kappa}\left(\tau_0\right)^{1/4}} \tag{144}$$

$$\lim_{a \to 0} t_0 = \lim_{\tau_0 \to \infty} t_0 = \left(\frac{\bar{k}}{\delta}\right)^{1/4} \lim_{\tau_0 \to \infty} \frac{\tau_0}{\mathcal{O}\left(\tau_0^{1/4}\right)} = \infty \tag{145}$$

$$\lim_{a \to \infty} t_0 = \lim_{\tau_0 \to 0} t_0 = \left(\frac{\bar{k}}{\delta}\right)^{1/4} \lim_{\tau_0 \to \infty} \frac{\tau_0}{\mathcal{O}\left(\tau_0\right)} = \hat{t}_0 \tag{146}$$

and for the density of matter we find

$$\rho'_M = \frac{8\pi G}{3}\rho_M = \left(a \sinh \frac{3\tau_0}{2}\right)^{-2} = \left(\frac{\bar{k}}{\delta\bar{\kappa}\left(\tau_0\right)}\right)^{-1/2} \left(\sinh \frac{3\tau_0}{2}\right)^{-2}$$
$$= \left(\frac{\delta}{\bar{k}}\right)^{1/2} \frac{\bar{\kappa}\left(\tau_0\right)^{1/2}}{\left(\sinh \frac{3\tau_0}{2}\right)^2} \tag{147}$$

$$\lim_{a \to 0} \rho'_M = \lim_{\tau_0 \to \infty} \rho'_M = \left(\frac{\delta}{\bar{k}}\right)^{1/2} \lim_{\tau_0 \to \infty} \frac{\mathcal{O}\left(\tau_0\right)}{\mathcal{O}\left(e^{3\tau_0}\right)} = 0 \tag{148}$$

$$\lim_{a \to \infty} \rho'_M = \lim_{\tau_0 \to 0} \rho'_M = \left(\frac{\delta}{\bar{k}}\right)^{1/2} \lim_{\tau_0 \to 0} \frac{\mathcal{O}\left(\tau_0^2\right)}{\mathcal{O}\left(\tau_0^2\right)} = \hat{\rho}'_M \tag{149}$$

[wc: Before we start experiments with greedy routing, we need to fix 2 things: 1) find for which values of parameters we can control $\bar{k}$ in simulated graphs, i.e. generate graphs with observed $\bar{k}$ close to the target one. First step is to fix N, δ, and a, and to plot $\langle \bar{k} \rangle$ and $\langle \mathcal{G} \rangle / N$ as functions of $\bar{k}$; expectations can be estimated using, say 1000 realizations. 2) Fix the entries of confusion matrix. It must be close to the identity matrix, otherwise, our embedding and the corresponding computation of the distance do not make sense. -kz-]

4. Greedy Routing Algorithm

The following describes the greedy routing algorithm in terms of pseudocode

INPUT:

- N - List of 4-D node coordinates

- N_{tar} - Target number of nodes in connected component

- N_{sr} - Percent of node pairs to traverse (up to $\frac{N(N-1)}{2}$)

START:

    for each pair (i, j):

        If $i == j$ **or** either node has no edges, **continue**

        If nodes in separate components, **continue**

        $loc = i$

        while $loc \mathrel{!}= j$

            $d_{min} = \text{INF}$

            $next = loc$

            for each past neighbor pn:

                If pn == j **then** loc = j; **break while**

                If distance(N_{pn}, N_j) $< d_{min}$ **then** save new d_{min} and $next$

            for each future neighbor fn:

                If fn == j **then** loc = j; **break while**

                If distance(N_{fn}, N_j) $< d_{min}$ **then** save new d_{min} and $next$

            If $next$ has not been traversed **then** $loc = next$

            Else break while loop

      If $loc == j$ **then** $R_s = R_s + 1.0$

      $N_{tr} = N_{tr} + 1$

    $R_s = R_s \mathbin{/} N_{tr}$

OUTPUT:

- R_s - Success ratio

- N_{tr} - Number of traversed pairs

5. *Success Ratio*

Here we have tested several combinations of parameters:

V. SUMMARY OF RESULTS

VI. LATEST CHALLENGES

Appendices

Appendix A: Usage of the Causal Set Software

[**wc:** `I would put this sectiona at the end, since many commands become clear only after reading the consequent sections.` *-kz-*] [**kz:** `Moved.` *-wc-*] The tasks in the program are broken into modular subroutines so that it can be used for diverse purposes. Functions are divided in a hierarchical structure so that the main method only accesses high-level functions (described in *Causets.cu*) with the entire *Network* data structure. In a similar fashion, these high-level functions strictly access low-level functions (in *NetworkCreator.cu* and *Measurements.cu*), and the low-level functions access mathematical (*Operations.h*) and algorithmic (*Subroutines.cu*) functions. Debugging and other validation tests are located in the *Validation.cu* file. When the GPU is used, entire low-level functions are replaced (e.g. in *NetworkCreator_GPU.cu*) because the flow of information is entirely different.

All tasks and parameters may be specified as command-line parameters - these may be queried using the **-h** or **–help** flags. They are shown here alphabetically for the purposes of documentation.

-A, –alpha: Set unphysical free parameter used in the universe's causal set, described in (11) and (12) of [3].

-a, –radius: Set pseudoradius of spacetime.

43

-C, –clustering:

    Calculate clustering coefficients and average clustering.

-c, –core: Set fraction of nodes used in adjacency matrix. A hybrid method is used which minimizes both memory allocation and lookup times. Typically this value is $\mathcal{O}(0.01)$ of all nodes.

-D, –delta: Set node density.

-d, –dim: Set number of spatial dimensions (1 or 3).

-E, –embedding:

    Perform embedding validation test. Space-like and time-like intervals are compared in the 3+1 and 4+1 universe's causal set.

-G, –components:

    Calcualate size of the giant connected component and the number of smaller connected components.

-g, –graph: Set graph ID. This is used to indicate a unique ID of an existing graph which the program reads and analyzes.

-h, –help: Display this information.

-k, –degrees: Set the average degrees of the graph.

-l, –lambda: Set the cosmological constant. Related to the pseudoradius via the equation $\Lambda = \frac{3}{a^2}$.

-m, –manifold:

    Indicate a [d]e Sitter or [h]yperbolic manifold. Used to indicate the type of graph being read when the **-g** flag is used.

-n, –nodes: Set the number of nodes to generate and attempt to link.

-o, –energy: Set the dark energy density Ω_Λ.

-r, –ratio: Set the ratio of dark energy to matter $\frac{\Omega_\Lambda}{\Omega_M}$.

-S, –success: Set the fraction of node pairs to use to study navigability and calculate the success ratio.

-s, –seed: Set the random seed.

-T, –test: Test parameters in the universe causal set to identify dependent constraints.

-t, –age: Rescaled age τ_0.

-u, –universe: Indicate a causal set with dust matter (and Big Bang).

-v, –verbose: Print verbose output when the program is run.

-y: Suppress user queries during the verbose output.

-z, –zeta: Set the curvature of a hyperbolic manifold.

–autocorr: Calculate autocorrelation of clustering coefficients.

–benchmark: Benchmark subroutines and numerical algorithms.

–conflicts: Show parameter conflicts for universe causal set.

–gpu: Use GPU accelerated subroutines.

–print: Print results to file in *dat* folder.

Appendix B: Alternative Expressions for the Conformal Time

One of the most important numerical steps in the simulation involves identifying causally connected points, which necessarily uses the conformal time. In the causal set representing the universe with matter, this requires solving

$$\eta\left(t\right) = \frac{1}{\alpha} \int_0^t \sinh^{-2/3}\left(\frac{3t'}{2a}\right) \, \mathrm{d}t' \tag{B1}$$

for every spacetime node. An integral of this form can easily be expressed in terms of the Gauss Hypergeometric function by equating a binomial series expansion to the hypergeometric series.

$$\int \sinh^n x \, \mathrm{d}x = \int \sinh x \, \sinh^{n-1} x \, \mathrm{d}x$$

$$= \int \sinh x \left(\cosh^2 x - 1 \right)^{(n-1)/2} \mathrm{d}x$$

$$= (-1)^{(1-n)/2} \int \sinh x \left(1 - \cosh^2 x \right)^{(n-1)/2} \mathrm{d}x \quad \text{(B2)}$$

$$= (-1)^{(1-n)/2} \cosh x \, {}_2F_1 \left(\frac{1}{2}, \frac{1-n}{2}; \frac{3}{2}; \cosh^2 x \right)$$

In the present case, we have $n = -2/3$, which yields the following intermediate result

$$\eta(t) = \frac{2a}{3\alpha} (-1)^{5/6} \left[\cosh \left(\frac{3t}{2a} \right) {}_2F_1 \left(\frac{1}{2}, \frac{5}{6}; \frac{3}{2}; \cosh^2 \left(\frac{3t}{2a} \right) \right) - \frac{\sqrt{\pi} \, \Gamma(1/6)}{2\Gamma(2/3)} \right] \quad \text{(B3)}$$

We can easily approximate the Hypergeometric function here by a power series, but only if the magnitude of its kernel function is less than one. Since we know $\cosh^2 x$ is always greater than one, we can implement the transformation (86) to get (after some algebra)

$$\eta(t) = \frac{a}{9\alpha\Gamma(2/3)} \left[\frac{4\sqrt{3}\pi^{3/2}}{\Gamma(5/6)} + 3\Gamma(-1/3) \operatorname{sech}^{2/3} \left(\frac{3t}{2a} \right) {}_2F_1 \left(\frac{1}{3}, \frac{5}{6}; \frac{4}{3}; \operatorname{sech}^2 \left(\frac{3t}{2a} \right) \right) \right] \quad \text{(B4)}$$

This expression may now be evaluated numerically using the Hypergeometric series

$$_2F_1(a, b \, c \, z) = \sum_{j=0}^{\infty} \frac{(a)_j \, (b)_j}{(c)_j} \frac{z^j}{j!} \quad \text{(B5)}$$

[1] ASMUS, I. Duality between hyperbolic and de sitter geometry. *Journal of Geometry 96* (2009).

[2] KRIOUKOV, D., KITSAK, M., SINKOVITS, R. S., RIDEOUT, D., MEYER, D., AND BOGUNA, M. Network cosmology. *Scientific Reports 2*, 793 (2012).

[3] KRIOUKOV, D., KITSAK, M., SINKOVITS, R. S., RIDEOUT, D., MEYER, D., AND BOGUNA, M. Supplemental information for network cosmology. *Scientific Reports 2* (2012).