PHYSICAL REVIEW D **88,** 124026 (2013)

Towards a definition of locality in a manifoldlike causal set

Lisa Glaser[1] and Sumati Surya[2]

[1]*Niels Bohr Institute, Blegdamsvej 17, 2100 Copenhagen, Denmark*

[2]*Raman Research Institute, C. V. Raman Avenue, Sadashivanagar, Bangalore 560 080, India*

(Received 27 September 2013; published 9 December 2013)

It is a common misconception that spacetime discreteness necessarily implies a violation of local Lorentz invariance. In fact, in the causal set approach to quantum gravity, Lorentz invariance follows from the specific implementation of the discreteness hypothesis. However, this comes at the cost of locality. In particular, it is difficult to define a ''local'' region in a manifoldlike causal set, i.e., one that corresponds to an approximately flat spacetime region. Following up on suggestions from previous work, we bridge this lacuna by proposing a definition of locality based on the abundance of m-element order-intervals as a function of m in a causal set. We obtain analytic expressions for the expectation value of this function for an ensemble of causal set that faithfully embeds into an Alexandrov interval in d-dimensional Minkowski spacetime and use it to define local regions in a manifoldlike causal set. We use this to argue that evidence of local regions is a necessary condition for manifoldlikeness in a causal set. This in addition provides a new continuum dimension estimator. We perform extensive simulations which support our claims.

DOI: 10.1103/PhysRevD.88.124026 PACS numbers: 04.60.Nc

I. INTRODUCTION

Causal set theory is a candidate for quantum gravity where the spacetime continuum is replaced by a discrete substructure which is a locally finite partially ordered set [1–4]. It is often assumed that Lorentz violation is an inevitable consequence of spacetime discreteness. This is explicitly false in a causal set discretization of a spacetime—on the contrary, as shown in [5], the causal set hypothesis instead *implies* Lorentz invariance. However, this comes at the cost of locality. In a causal set, the nearest neighbors of an element are the links or irreducible relations. For example, in infinite causal set that is approximated by Minkowski spacetime, every element has an infinite number of nearest neighbors, both to the past and to the future. The resulting graph is therefore of infinite valency, in stark contrast to other types of spacetime discreteness in which the graphs are of finite valency. Indeed, it is this very feature of a causal set which captures the essence of Lorentz-invariant discreteness, since there are noncompact invariant hyperbolas associated with every Lorentz boost about a spacetime event in Minkowski spacetime. This feature of causal set discretization in turn is due to the requirement of a uniform distribution which preserves the number to volume correspondence, crucial to the recovery of the Lorentzian spacetime geometry in the continuum approximation [5].

While Lorentz invariance is a great asset to causal set theory, the resulting nonlocality of the causal set graph impedes a straightforward reconstruction of continuum information from the discrete substructure. Unlike a simplicial decomposition, for example, where the discrete scalar curvature has a simple *local* geometric interpretation, there is no analogous local construction in a causal set. Indeed, it is only recently that a causal set expression for scalar curvature and hence a causal set action has been found in arbitrary dimensions [6,7]. Nevertheless, despite the difficulty in recovering local information from a causal set, over the years substantial progress has been made in understanding how topology and geometry emerges from a causal set, sometimes with the aid of fairly ingenious order-theoretic constructions. This includes the reconstruction of spacetime dimension, timelike distance, spacelike distance and spatial homology, for causal sets that are approximated by continuum spacetimes [8].

Importantly, in many of these reconstructions the causal set is assumed to be approximated by a region of curved spacetime which is small compared to the scale of flatness. In the continuum such a region has a natural interpretation of being ''local'' or approximately flat. From the continuum perspective small, or local neighborhoods are essential to several geometric constructions and are key to the conception of a manifold. However, until now there has been no purely order theoretic characterization of such local neighborhoods in a causal set. It is therefore an important step to be able to define local regions in a causal set and hence provide an appropriate context for some of the reconstruction results.

Our prescription for locality uses a well-known order theoretic definition of a spacetime region, namely an Alexandrov interval $I[x, y] := \{z | x \prec\prec z \prec\prec y\}$, where $\prec\prec$ is the chronological relation. In the continuum $I[x, y]$ is characterized both by the timelike distance $\tau(x, y)$ from x to y, as well as its volume $\mathrm{Vol}(I[x, y])$. Unlike open ball neighborhoods in a Riemannian manifold, however, even arbitrarily small choices of $\tau(x, y)$ or $\mathrm{Vol}(I[x, y])$ do not correspond to a region in which the scale of flatness is large as illustrated in Fig. 1. In the continuum, $\tau(x, y)$ or $\mathrm{Vol}(I[x, y])$ are the only Lorentz invariant quantities that characterize $I[x, y]$. However, it is clear that the

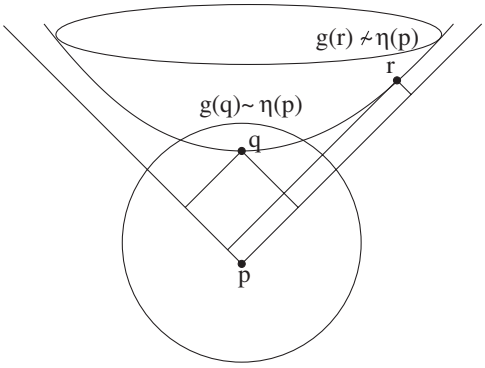


FIG. 1. The two Alexandrov intervals $I[p, q]$ and $I[p, r]$ have the same volume, but the former lies in an approximately flat neighborhood of p, while the latter does not.

corresponding discrete geometry, i.e., a causal set which faithfully embeds into $I[x, y]$, should contain more detailed geometric information.

The main proposal of our work is that there is indeed an order theoretic characterization of locality. The motivation for this arises from the work of Benincasa and Dowker [6,7] where it was found that the scalar curvature of an element e in a causal set C can be obtained from knowing the abundances N_m of order-intervals of size m that lie to the past of e. Here, N_0 is the abundance of 0 order-intervals, i.e., the number of irreducible relations or links in C, N_1 the abundance of 1-element order intervals, or irreducible 3-chains, etc. For a generic nonlocality scale the discrete Einstein-Hilbert action is constructed from all possible N_m, but when the nonlocality scale is taken to be the Planck scale, the action simplifies considerably. For example, the 2d causal set action takes the elegant form

$$\frac{1}{\hbar} S_{2D} = N - 2N_0 + 4N_1 - 2N_2, \tag{1}$$

which only involves the abundances of intervals of volume 0, 1 and 2.

In simulations of 2d quantum gravity using Markov Chain Monte Carlo methods the N_m were used as covariant observables for tracking thermalization [9]. Importantly, it was observed that N_m as a function of m has a characteristic behavior in the phase in which flat spacetime is emergent, and differs drastically from the nonmanifold phase. Simulations of 2d flat spacetime showed that this characteristic curve could indeed be used as a reliable indicator of flatness.

In this work we carry this idea forward. We begin by obtaining analytic expressions for the expectation value of the interval abundances $\langle N_m^d \rangle$ for a causal set that faithfully embeds into an Alexandrov interval in flat spacetime of arbitrary dimensions d. Our main proposal is that the characteristic curves for $\langle N_m^d \rangle$ as a function of m can be used to define a local region in a manifoldlike causal set C. The existence of a local region in a causal set is therefore a necessary condition for manifoldlikeness of C and hence a new continuum dimension estimator. Specifically, since the characteristic curves for $\langle N_m^d \rangle$ for fixed cardinality are sufficiently distinct for each d, it is possible to use them to find the continuum dimension of the local region in the causal set. This estimator therefore gives a null result for causal sets which are nonmanifoldlike. Because the $\langle N_m^d \rangle$ provide an entire family of observables, it is tempting to conjecture that the requirement on interval abundances is not only a necessary but also a sufficient requirement for manifoldlikeness of a causal set.

We test our proposals with simulations of causal sets that are approximated by spacetimes as well as those that are not. We find that our necessary condition for manifoldlikeness works extremely well even for relatively small causal sets.

Indeed, not only do the simulated interval abundances reproduce on average the characteristic curve, they follow it with reasonable precision even in a single realization.

The latter is especially important in assessing manifoldlikeness in a single causal set, as opposed to an ensemble of causal sets. Our simulations verify that apart from being able to determine the local regions of a manifoldlike causal set, our prescription is also a test for manifoldlikeness itself and thence, manifold dimension.

Our construction demonstrates clearly the geometric richness of a locally finite poset which is approximated by a spacetime. Our analysis indicates the existence of a local geometric ''rigidity'' present in manifoldlike causal sets—significant deviations from the $\langle N_m^d \rangle$ result in Alexandrov intervals that are explicitly *not* local. Thus, $\langle N_m^d \rangle$ as a function of m provides us a local, covariant, geometric measure for manifoldlikeness.

In Sec. II we give a short introduction to the main concepts of causal set theory and define quantities that we will need further on. The calculation of the interval abundances in flat spacetime are in Sec. III. We find that the ratio of these abundances is scale invariant in the large N limit, which provides a strong motivation for using the $\langle N_m^d \rangle$ as indicators of locality. In Sec. IV we present the main ideas in this work, namely, how the $\langle N_m^d \rangle$ can be used to define locality in a manifoldlike causal set. We then conjecture that the $\langle N_m^d \rangle$ provide a rigidity criterion for manifoldlikeness for a causal set that faithfully embeds into an Alexandrov interval in flat spacetime. In Sec. V we present results from extensive numerical simulations that support these ideas. In particular, we examine the interval abundances of causal sets that we know a priori to be either manifoldlike or not and find that our test works extremely well. Our tests include flat spacetimes of different dimensions, the 2-d cut-trousers topology, as well as the flat geometries on $S^1 \times \mathbb{R}$ and $T^2 \times \mathbb{R}$, 4-d FRW spacetimes, including de Sitter spacetime, and some examples of nonmanifoldlike causal sets. In particular, we examine the claim from [10] that causal sets grown with transitive

percolation are manifoldlike. We find that while macroscopic indicators may suggest manifoldlikeness, it fails our microscopic test. We end with a discussion on open questions and future directions in Sec. VI.

II. PRELIMINARIES

Studies of Lorentzian geometry have long stressed the importance of the causal structure [11]. For causal spacetimes, the causal structure provides a partial order on the set of spacetime events. This partial order is a unique characteristic of a Lorentzian signature $((-, +, + \cdots +))$ spacetime, a feature absent in all other signature spacetimes. It was shown by Malament, Hawking and others in [12] that a bijection between two past and future distinguishing spacetimes which preserves the causal structure is a conformal isomorphism. Thus, knowing the causal relations between all points in a spacetime is enough to define its geometry up to a conformal factor. The causal set hypothesis of a fundamental discreteness adds to this classical result by providing a discrete volume element to help recover the conformal factor. Roughly, every discrete event comes with approximately one unit of spacetime volume so that the number of events in a region corresponds to the volume of that region. In other words, in causal set theory, an appropriately discretized partially ordered set replaces continuum Lorentzian geometry, summarized in the slogan: Order + Number $\approx$ Spacetime.

Formally a causal set C is defined to be a locally finite partially ordered set, namely a countable set C with an order relation $\preceq$ on its elements which is

(i) Reflexive: for all $x \in \mathcal{C}$, $x \preceq x$
(ii) Transitive: for all $x, y, z \in \mathcal{C}$ and $x \preceq y$ and $y \preceq z$ then $x \preceq z$,
(iii) Acyclic: for all $x, y \in \mathcal{C}$, $x \preceq y \preceq x \Rightarrow x = y$
(iv) Locally Finite: for all $x, y \in \mathcal{C}$ $|I(x, y)| \equiv |\{z | x \preceq z \preceq y\}| < \infty$.

This last condition is equivalent to the assumption of a fundamental discreteness. The first figure in Fig. 2 shows the Hesse diagram of a small causal set where the elements are numbered and the links are denoted by arrows.

Causal set quantum gravity is thus a quantum theory of causal sets with the continuum existing only as an approximation to a fundamentally discrete substructure. In particular, the ensemble of causal sets that are approximated by a given spacetime (M, g) is obtained via a Poisson process for a given discreteness scale ρ^{-1}. The probability of assigning m-elements of a causal set C in a spacetime region of volume V is given by

$$P_V(m) = \frac{(\rho V)^m}{m!} e^{-\rho V}. \tag{2}$$

The causal set is then recovered via the induced causal relations on the set of elements. This Poisson "sprinkling" is a key feature of causal set discretization of the continuum. The second figure in Fig. 2 shows a sprinkling of 100 elements into flat space. Conversely, given a causal set, C, we say that it is approximated by a spacetime (M, g) if there exists a *faithful embedding* Φ: $C \rightarrow (M, g)$ such that $\Phi(C) \subset (M, g)$ corresponds to a high probability Poisson sprinkling into (M, g) with the order relations in C being the same as those induced by the causal relations in (M, g) onto $\Phi(C)$. An important conjecture in the theory is that a given causal set cannot faithfully embed into two distinct spacetimes, namely, those which differ on scales larger than the discreteness scale. In other words continuum structures below the discreteness scale are irrelevant to the theory. We refer the reader to the literature for more details on the basics of the causal set hypothesis [1–4].

As discussed in the introduction, the key focus of this work is to be able to define a local region in a manifoldlike causal set. In the continuum, an Alexandrov interval $I[x, y] \equiv \{z | x \prec\prec z \prec\prec y\}$, where $\prec\prec$ is the chronological relation. In a causal set, there is no a priori distinction

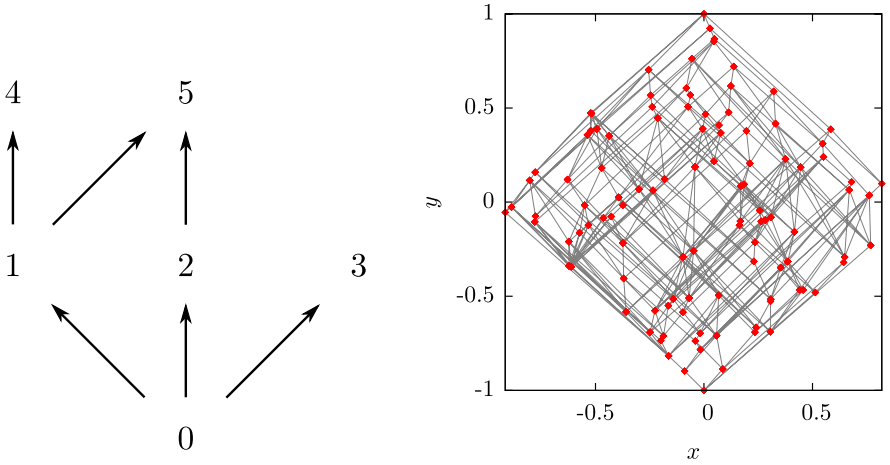


FIG. 2 (color online). On the left is the Hesse diagram of a simple causal set where only the links are shown by arrows. The figure on the right is a sprinkling of 100 elements into an interval of flat spacetime, where all the causal relations are depicted by lines.

between causal and chronological relations and hence we define an order interval $I[x, y] \subset C$ as $I[x, y] = \{z | x \prec z \prec y\}$. A natural characterization of $I[x, y]$ is its cardinality or discrete volume. However, this information does not suffice to distinguish an $I[x, y]$ which is local and one that is not. Since the discrete geometry should include all relevant information about the continuum, we expect that there must exist other observables in $I[x, y]$ which can be used to characterize locality. In this work, we find that the abundance of m-element order-intervals in $I[x, y]$ is indeed such a family of observables. Namely, for every m we count the number or abundance of order intervals of size m in $I[x, y]$. For us, an $m = 0$ order interval is a link, namely, an order interval which contains only its endpoints, an $m = 1$ order interval is one with a single element that lies between the endpoints, or an irreducible 3 chain, and $m = 2$ can be an irreducible 4 chain or an irreducible diamond poset, i.e., an order interval with two elements between the endpoints.

For a causal set that faithfully embeds into flat spacetime, we find that the interval abundances follow a characteristic, monotonically decreasing curve as m increases. It is this characteristic curve that we will use as a ''ruler'' to determine the locality of an order interval in a more general manifoldlike causal set. However, in order to do so, we would like to ensure that the scale of flatness in every region of the spacetime is much larger than the discreteness scale, so that the manifold approximation of the causal set is well-defined everywhere. By this we mean the following. Consider a causal set that faithfully embeds into an ''approximately flat'' spacetime region in which Riemann normal coordinates are valid. Such a region is characterized by a dimensionless size which we will refer to as the scale of flatness $\zeta^{-1} \gg 1$ where $\zeta = R\tau^2$ with R denoting any component of the Riemann tensor at an event in the region and τ the proper time between any two events in the region. In flat spacetime $\zeta = 0$ and hence the size $\zeta^{-1} \rightarrow \infty$ as expected. For a generic spacetime we will refer to such regions as ''small'': for a given R, the size of the region τ must be small enough for $\zeta \ll 1$. Let C be a causal set that faithfully embeds at density ρ into an Alexandrov interval $I[x, y]$ of volume V which lies in a region for which $\zeta^{-1} \gg 1$. If $N \sim \rho V \sim 1$ then the continuum approximation of C breaks down. Thus, in order for the $I[x, y]$ to be adequately represented by the causal set, we require that $N \gg 1$, i.e., the discreteness scale N^{-1} must be small with respect to the scale of flatness ζ^{-1}.

III. THE ABUNDANCES OF ORDER INTERVALS IN FLAT SPACETIME

We now find closed form expressions for the abundance of order-intervals in a causal set C that is faithfully embedded into an Alexandrov interval $I[p, q]$ in flat spacetime. To begin with we find the abundance of links $\langle N_0^d \rangle$. Lorentz invariance then allows us to generalize this expression to that for $\langle N_m^d \rangle$ in a straightforward way. Although we use series expansions to evaluate the integrals, the final expressions take relatively simple closed forms.

Consider the interval $I[p, q]$ in Fig. 3 with volume V and proper time τ from p to q. The probability that there is a link from an element x to an element y in this region is given by

$$P_{xy} = e^{-\rho V_{xy}}, \tag{3}$$

where V_{xy} is the spacetime volume of the Alexandrov interval $I[x, y] \subset I[p, q]$ in the embedding spacetime. Moreover, *given* x, y lies to its future, i.e., $y \in I[x, q]$, while x can lie anywhere in $I[p, q]$. Thus, the expectation value for the number of links in $I[p, q]$ is given by

$$\langle N_0^d(\rho, V) \rangle = \rho^2 \int_{\diamond} dV_x \int_{\diamond_x} dV_y e^{-\rho V_{xy}}, \tag{4}$$

where the symbol $\diamond$ denotes $I[p, q]$ and the symbol $\diamond_x$ denotes $I[x, q]$. Since this expression is Lorentz invariant, it can only depend on the proper time τ or volume V of $I[p, q]$. Thus we may choose p at the origin $p = (0, \ldots, 0)$ and q on the time axis $q = (\sqrt{a}, 0, \ldots, 0)$, where $\tau = \sqrt{2}a$.

Lorentz covariance also implies that the integration over $\diamond_x$ depends only the proper time $\tau(x, q)$ and the volume V_{xq} of $I[x, q]$. We may thus again calculate this integral in convenient coordinates and then recast it in terms of $\tau(x, q)$ and V_{xq}. We take x to lie at the origin $x = (0, \ldots, 0)$ and q to lie on the time axis $q = (\sqrt{2}a', 0, \ldots, 0)$, where $\tau(x, q) = \sqrt{2}a' \leq \sqrt{2}a$. Using light-cone coordinates

$$v = \frac{1}{\sqrt{2}}(t + r) \qquad u = \frac{1}{\sqrt{2}}(t - r), \tag{5}$$

the integration measure in flat space for $y = (u_y, v_y, \vec{\Omega}_y)$ is then

$$\int dV_y = \frac{1}{2^{\frac{d}{2}-1}} \int d\Omega_y \int_0^{a'} dv_y \int_0^{v} du_y (v - u)^{d-2}. \tag{6}$$

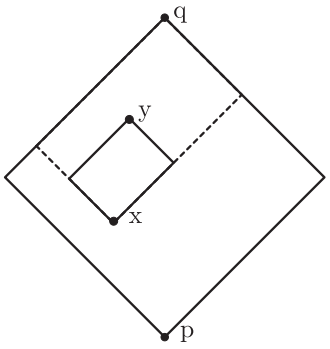


FIG. 3. The Alexandrov neighborhood $I[p, q]$ in the integration Eq. (4): x lies anywhere in $I[p, q]$ while y lies in the intersection of $I[p, q]$ with the chronological future of x.

The function we will integrate over does not involve the angular coordinates of y, since $\tau(x, y) = \sqrt{2u_y v_y}$. We can then rewrite

$$\rho V_{xy} = \rho \frac{S_{d-2} 2^{-\frac{d}{2}+1}}{d(d-1)} (u_y v_y)^{\frac{d}{2}} = \rho \zeta_d (u_y v_y)^{\frac{d}{2}}, \tag{7}$$

where S_{d-2} is the volume of the $d-2$ sphere $S_{d-2} = \int d\Omega_y = \frac{2\pi^{\frac{d-1}{2}}}{\Gamma(\frac{d-1}{2})}$, and we define the dimension-dependent constant $\zeta_d \equiv \frac{S_{d-2} 2^{-\frac{d}{2}+1}}{d(d-1)}$. Thus the angular integration over $\diamondsuit_x$ factors out so that, in these coordinates, the integral $I'_{\diamondsuit_x} \equiv \int_{\diamondsuit_x} dV_y e^{-\rho V_{xy}}$ reduces to

$$I'_{\diamondsuit_x} = d(d-1)\zeta_d \int_0^{a'} dv \int_0^v du (v-u)^{d-2} e^{-\rho \zeta_d (uv)^{d/2}}, \tag{8}$$

where we have suppressed the subscripts in (u_y, v_y). Expanding $(v-u)^{d-2}$ in terms of binomial coefficients and $e^{-\rho\zeta_d (uv)^{d/2}}$ as a power series simplifies the integration considerably

$$I'_{\diamondsuit_x} = d(d-1)\zeta_d \sum_{n=0}^{\infty} \frac{(-\rho\zeta_d)^n}{n!} \sum_{k=0}^{d-2} \binom{d-2}{k} (-1)^k \times \int_0^{a'} dv \int_0^v du v^{d(\frac{n}{2}+1)-2-k} u^{\frac{dn}{2}+k} \tag{9}$$

$$= d(d-1)\zeta_d \sum_{n=0}^{\infty} \frac{(-\rho\zeta_d)^n}{n!} \frac{a'^{d(n+1)}}{d(n+1)} \sum_{k=0}^{d-2} \binom{d-2}{k} \frac{(-1)^k}{\frac{dn}{2}+k+1}. \tag{10}$$

Rewriting

$$\sum_{k=0}^{d-2} \binom{d-2}{k} \frac{(-1)^k}{\frac{dn}{2}+k+1} = \frac{\Gamma(d-1)\Gamma(\frac{dn}{2}+1)}{\Gamma(\frac{d}{2}(n+2))} \tag{11}$$

we find that

$$I'_{\diamondsuit_x} = d(d-1)\zeta_d \Gamma(d-1) \sum_{n=0}^{\infty} \frac{(-\rho\zeta_d)^n}{n!} \frac{a'^{d(n+1)}}{d(n+1)} \frac{\Gamma(\frac{dn}{2}+1)}{\Gamma(\frac{d}{2}(n+2))}. \tag{12}$$

We can now convert the above expression into a Lorentz covariant form by substituting a' for the proper time $\tau(x, q) = \sqrt{2}a'$ of $I[x, q]$. In the original coordinates adapted for $I[p, q]$ this is $\tau(x, q)^2 = 2(a - v_x)(a - u_x)$ Thus, to complete the calculation of $\langle N_0^d \rangle$ we must evaluate the integral

$$I_\diamond = \int_\diamond dV_x ((a-u_x)(a-v_x))^{\frac{d(n+1)}{2}} = \zeta_d d(d-1) \int_0^a dv_x \int_0^{v_x} du_x (v_x - u_x)^{d-2} ((a-u_x) \times (a-v_x))^{\frac{d(n+1)}{2}}. \tag{13}$$

To shorten the notation we defined this integral without the sum over n which we will have to restore in the final expression for $\langle N_0^d \rangle$. Substituting $u = a - u_x$ and $v = a - v_x$ and again using the binomial expansion

$$I_\diamond = \zeta_d d(d-1) \sum_{k=0}^{d-2} \binom{d-2}{k} (-1)^k \times \int_0^a dv \int_0^v du u^{\frac{d}{2}(n+1)+k} v^{\frac{d}{2}(n+3)-2-k} \tag{14}$$

$$= \zeta_d d(d-1) \frac{a^{d(n+2)}}{d(n+2)} \sum_{k=0}^{d-2} \binom{d-2}{k} (-1)^k \times \frac{1}{\frac{d}{2}(n+1)+k+1}. \tag{15}$$

Using the identity (11), and the fact that $V = \zeta_d a^d$, we find the Lorentz covariant expression for the average number of links in an interval $I[p, q]$ of volume V to be

$$\langle N_0^d \rangle(\rho, V) = \Gamma(d)^2 \sum_{n=0}^{\infty} \frac{(-\rho V)^{n+2}}{(n+2)!} \frac{\Gamma(\frac{dn}{2}+1)}{\Gamma(\frac{d}{2}(n+2))} \times \frac{\Gamma(\frac{d}{2}(n+1)+1)}{\Gamma(\frac{d}{2}(n+3))}. \tag{16}$$

This expression can now be used to find the m element interval abundances in flat space for general m by observing that (2) can be rewritten as

$$P(m, V, \rho) = \frac{(-\rho)^m}{m!} \frac{\partial^m}{\partial \rho^m} e^{-\rho V}. \tag{17}$$

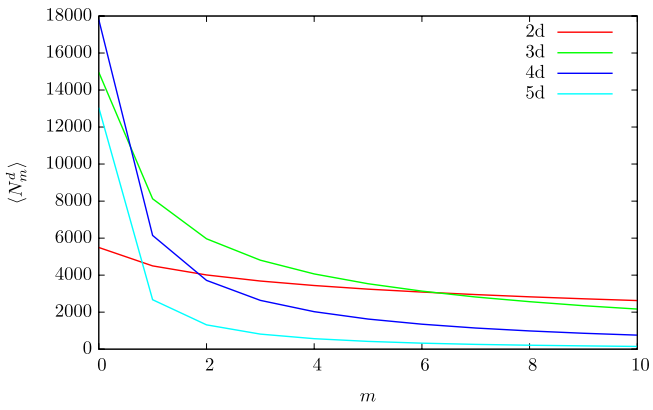


FIG. 4 (color online). The function $\langle N_m^d \rangle$ v/s m for $N = 1000$ and $d = 1, \ldots, 5$.

Thus, the average number of m-element intervals in a volume V is simply given as

$$\langle N_m^d\rangle(\rho,V)=\rho^2\int_{\diamond}\mathrm{d}v_y\int_{\diamond_y}\mathrm{d}V_x\frac{(\rho V)^m}{m!}e^{-\rho V}=\frac{(-\rho)^{m+2}}{m!}\frac{\partial^m}{\partial\rho^m}\int_{\diamond}\mathrm{d}V_x\int_{\diamond_y}\mathrm{d}v_y e^{-\rho V}=\frac{(-\rho)^{m+2}}{m!}\frac{\partial^m}{\partial\rho^m}\rho^{-2}\langle N_0^d\rangle(\rho,V) \tag{18}$$

which evaluates to:

$$\langle N_m^d\rangle(\rho,V)=\frac{\Gamma(d)^2}{m!}(-\rho V)^{m+2}\sum_{n=0}^{\infty}\frac{(-\rho V)^n}{n!}\frac{1}{(n+m+1)(n+m+2)}\times\frac{\Gamma(\frac{d}{2}(n+m)+1)}{\Gamma(\frac{d}{2}(n+m+2))}\frac{\Gamma(\frac{d}{2}(n+m+1)+1)}{\Gamma(\frac{d}{2}(n+m+3))}. \tag{19}$$

This can then be recast as a closed form expression in terms of generalized hypergeometric functions:

$$\langle N_m^d\rangle(\rho,V)=\frac{(\rho V)^{m+2}}{(m+2)!}\frac{\Gamma(d)^2}{(\frac{d}{2}(m+1)+1)_{d-1}}\frac{1}{(\frac{d}{2}m+1)_{d-1}}{}_dF_d\left(\begin{matrix}1+m,\frac{2}{d}+m,\frac{4}{d}+m,\ldots,\frac{2(d-1)}{d}+m\\ 3+m,\frac{2}{d}+m+2,\frac{4}{d}+m+2,\ldots,\frac{2(d-1)}{d}+m+2\end{matrix}\middle|-\rho V\right), \tag{20}$$

where ${}_pF_q(\{a_1,\ldots,a_p\},\{b_1,\ldots,b_q\}|-z)$ is a generalized hypergeometric function and $(a)_n$ is the Pochhammer symbol. This expression is convergent because, as is well known, generalized hypergeometric functions converge for all z values if $p\leq q$. The details of obtaining this form for the $\langle N_m^d\rangle$ are given in Appendix A.

In Fig. 4 we plot the function $\langle N_m^d\rangle$ for different values of d. $\langle N_m^d\rangle$ rapidly and monotonically decreases as m increases thus providing a clear characteristic signature for the flat spacetime case, which we will use to define locality and thence a continuum dimension estimator.

A. The asymptotic limit

We now examine the behavior of these expressions in the large $N=\rho V$ limit. Although the continuum limit is not physically relevant per se to causal set theory, it is nevertheless an interesting limit of the theory. Clearly, $\langle N_m^d\rangle$ will diverge with N, but it is not clear a priori what the behavior will be after normalization, say with respect to the abundance of links, $\langle N_0^d\rangle$. If there were a leading order N-dependence, then either this would diverge, or go to zero in the limit, thus drastically changing the nature of the $\langle N_m^d\rangle$ characteristic curve.

What we find is that the ratio is in fact independent of N to leading order and yields a surprisingly simple expression in the asymptotic limit

$$\mathcal{S}_m^d\equiv\lim_{\rho\to\infty}\frac{\langle N_m^d\rangle(\rho,V)}{\langle N_0^d\rangle(\rho,V)}=\frac{\Gamma(\frac{2}{d}+m)}{\Gamma(\frac{2}{d})\Gamma(m+1)}. \tag{21}$$

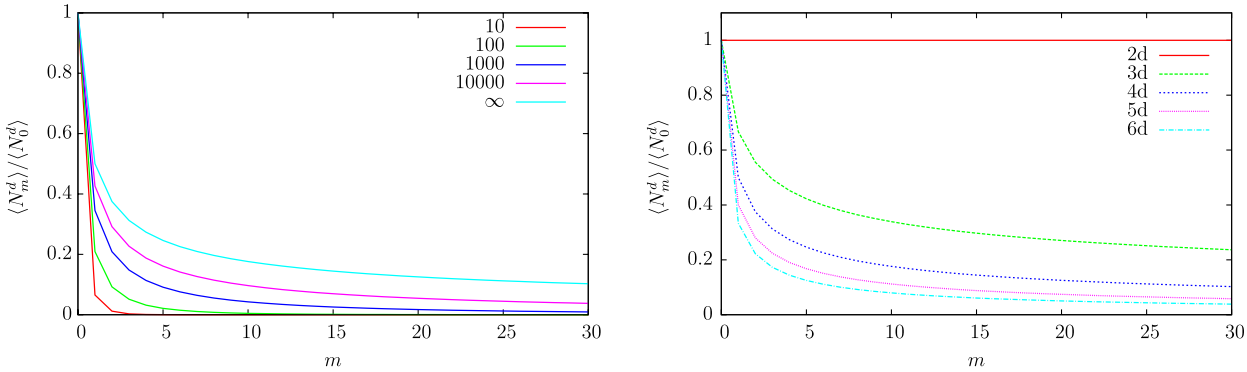
This scale invariance seems to echo that of Minkowski spacetime suggesting that the $\langle N_m^d\rangle$ captures an essential and perhaps even defining ingredient of flat spacetime geometry. We discuss this in some detail in the following section.

Here we give a quick sketch of how this limit is obtained, leaving details to Appendix C. The $N=\rho V$ dependence in Eq. (20) comes from the overall N^{m+2} factor as well as the hypergeometric function ${}_dF_d$ which when appropriately rearranged is of the form

$${}_dF_d\left(\begin{matrix}a_1,\ldots,a_d\\ a_1+2,\ldots,a_d+2\end{matrix}\middle|-N\right),$$
$$a_i=\frac{2i}{d}+m,\quad i=1,\ldots d-1,\quad a_d=1+m. \tag{22}$$

Thus, to investigate the $N\to\infty$ limit of the $\langle N_m^d\rangle$ we need a large N expansion of this function. We make repeated use of the following identity [13]

$$\begin{aligned}&{}_pF_q\left(\begin{matrix}a_1,\ldots,a_p\\ a_1+m_1,\ldots,a_n+m_n,b_{n+1},\ldots,b_q\end{matrix}\middle|z\right)\\ &=\prod_{j=1}^{n}\frac{(a_j)_{m_j}}{(m_j-1)!}\sum_{k=1}^{n}\sum_{j_1=0}^{m_1-1}\cdots\sum_{j_n=0}^{m_n-1}\frac{1}{a_k+j_k}\prod_{l=1}^{n}\frac{(1-m_l)_{j_l}}{j_l!}\\ &\quad\times\prod_{\substack{i=1\\ i\neq k}}^{n}\frac{1}{a_i+j_i-a_k-j_k}\,{}_{p-n+1}F_{q-n+1}\left(\begin{matrix}a_k+j_k,a_{n+1},\ldots,a_p\\ a_k+j_k+1,b_{n+1},\ldots,b_q\end{matrix}\middle|z\right),\\ &m_n\in\mathbb{Z}\wedge m_n>0\wedge n\leq q\wedge a_i+j_i\neq a_k+j_k,\quad\forall\ j_i=0,\ldots m_i,\,1\leq i,j\leq n.\end{aligned} \tag{23}$$



(a) Convergence in 4d (b) Varying Dimension

FIG. 5 (color online). These plots illustrate some properties of the $N \to \infty$ limit. (a) The left-hand plot shows the convergence of $\frac{\langle N_m^d\rangle}{\langle N_0^d\rangle}$ to S_m^d in 4d for $N = 10, 100, 1000, 10000$. (b) The right-hand plot shows how S_m^d changes with dimension.

This can be used to reduce the ${}_dF_d$ of the form in Eq. (22) to (i) a sum over ${}_1F_1$ in odd dimensions, (ii) a sum over ${}_3F_3$ in even dimensions $d > 2$. In $d = 2$ ${}_dF_d$ is simply ${}_2F_2$ which can be examined directly. We demonstrate these results explicitly in Appendix C.

Specifically, in odd dimensions there is a sum over

$${}_1F_1\left(\begin{matrix} a_k + j_k \\ a_k + j_k + 1 \end{matrix}\middle| -N\right),$$

which using

$${}_1F_1\left(\begin{matrix} a \\ a+1 \end{matrix}\middle| -z\right) = a(z)^{-a}(\Gamma(a) - \Gamma(a,z)) \quad (24)$$

$$\Gamma(a,z) \propto e^{-z}z^{a-1} \times\left(\frac{(2-a)(1-a)}{z^2} - \frac{1-a}{z} + \cdots + 1\right)\Big/;(|z|\to\infty). \quad (25)$$

gives a leading order dependence of $N^{-a_k-j_k}$, for the smallest values of $a_k + j_k$ which is $k = 1$, $j_1 = 0$, which makes it $N^{-\frac{2}{d}-m}$. In even dimensions for $d > 2$, the dependence on N appears in a sum over

$${}_3F_3\left(\begin{matrix} \frac{2}{d}k + m + j_k, \frac{2}{d}l + m + 1, m + 1 \\ \frac{2}{d}k + m + j_k + 1, \frac{2}{d}l + m + 2, m + 3 \end{matrix}\middle| -N\right). \quad (26)$$

We obtain an asymptotic expansion of this using MATHEMATICA and find that the leading order contribution is again $N^{-a_k-j_k}$ and hence comes from the $k = 1$, $j_1 = 0$ term which makes it $\sim N^{-\frac{2}{d}-m}$. In $d = 2$, the hypergeometric function in Eq. (20) is simply

$${}_2F_2\left(\begin{matrix} m+1, m+1 \\ m+3, m+3 \end{matrix}\middle| -N\right)$$

whose leading order contribution is of the form $N^{-1-m}\log N$. Combining these we find that $\langle N_m^d\rangle \sim N^{2-2/d}$ to leading order for $d > 2$ and $\langle N_m^d\rangle \sim N\log N$ for $d = 2$.

What our detailed calculations show, moreover, is that in *all* dimensions the coefficient of the leading order term takes the simple form:

$$\frac{1}{m!}\Gamma\left(\frac{2}{d} + m\right)\frac{\Gamma(d)}{(\frac{d}{2}-1)(\frac{d}{2}+1)_{d-2}} \quad (27)$$

which implies Eq. (21). The subleading contributions however vary from dimension to dimension as

$$\langle N_m^d\rangle(N) = \frac{N^{2-\frac{2}{d}}}{m!}\Gamma\left(\frac{2}{d} + m\right)\frac{\Gamma(d)}{(\frac{d}{2}-1)(\frac{d}{2}+1)_{d-2}} + \begin{cases} \mathcal{O}(N) & \text{for } d = 3 \\ \mathcal{O}(N\log N) & \text{for } d = 4 \\ \mathcal{O}(N^{2-\frac{4}{d}}) & \text{for } d > 4 \end{cases} \quad (28)$$

for all $d > 2$ and

$$\langle N_m^2\rangle(N) = N\log N + \mathcal{O}(N), \quad (29)$$

for $d = 2$. We refer the reader to Appendix C for the details of the calculation.

In particular we note that all contributions are slower than N^2 and that the convergence toward the limit happens polynomially, and hence is quite slow. In Fig. 5 we plot $\langle N_m^d\rangle(N)$ for $d = 4$ for a range of N-values, as well as the asymptotic limit. The slow convergence makes it clear that it will not be possible to test this limit computationally. In Fig. 5 we plot the asymptotic limits for various d. Notably, for $d = 2$ $S_m^2 = 1$ and therefore independent of m.

IV. DEFINING LOCAL REGIONS IN A CAUSAL SET

The main goal of this work is to show that $\langle N_m^d\rangle$ can be used as a definition of locality in a causal set which faithfully embeds into a continuum spacetime. Conversely it can be used as a test for manifoldlikeness as well as a continuum dimension estimator.

Consider an N-element causal set C which faithfully embeds into an Alexandrov interval $I[p, q]$ in d-dimensional Minkowski spacetime at a given density ρ. As we have just shown above, if one considers the ensemble of causal sets obtained via a Poisson sprinkling into $I[p, q]$, then the average $\langle N_m^d \rangle$ has a characteristic behavior with m. For large enough ρ, the interval abundances $N_m(C)$ for a single "typical" realization will with high probability "track" $\langle N_m^d \rangle$, i.e., $N_m(C) \sim \langle N_m^d \rangle \times (N \pm \sqrt{N})$ for all m. This is what we would expect from a Poisson distribution. As we will show in the following section, this expectation is confirmed by simulations. Simulations, moreover, show that the distribution of the N_m^d for any given m for an ensemble of causal sets obtained via a Poisson sprinkling into $I[p, q]$ is nearly Gaussian with a standard deviation of $\sim\sqrt{N}$.

Importantly, the closeness of a typical $N_m(C)$ to $\langle N_m^d \rangle$ can be used as a characterization of locality. Namely, if C is such that $N_m(C) \sim \langle N_m^d \rangle (N \pm \sqrt{N})$ for all m, and for a fixed d we will refer to it as a "local" causal set. For a causal set $\tilde{C}$ which faithfully embeds into an Alexandrov interval $I[p, q]$ in an arbitrary curved d-dimensional spacetime, one expects that because of the deviation from flatness, $N_m(\tilde{C})$ will differ significantly from the $\langle N_m^d \rangle$. Again, this is borne out by simulations. Thus, $\tilde{C}$ is "nonlocal" in this sense. However, as long as the scale of flatness everywhere in $I[p, q]$ is much larger than the discreteness scale as discussed in Sec. II,[1] $\tilde{C}$ will contain N-element subcausal sets C which lie in an approximately flat Alexandrov interval $I[p', q'] \subset I[p, q]$. If N is large enough, then C will be local in the above sense. Thus, the $\langle N_m^d \rangle$ provide a strong characterization of local regions in a causal set. Again, this is borne out by simulations on a class of curved spacetimes as well as those with nontrivial topology.

Thus, the function $\langle N_m^d \rangle$ suggests a criterion for "rigidity" of C in the sense used by mathematicians. Namely, if $N_m(C) \sim \langle N_m^d \rangle (N \pm \sqrt{N})$ for all m, then it suggests that C must faithfully embed into Minkowski spacetime of dimension d at large enough embedding density. We now formalize these ideas as best as we can, leaving a more detailed study to future work.

Definition 1 We will say that an N-element causal set C is *strongly d-rigid* if $\exists$ a d for which $N_m(C) \sim \langle N_m^d \rangle \times (N \pm \sqrt{N})$. If C possesses an N' element sub causal set C' which is strongly d-rigid, then C is said to be *weakly d-rigid* with respect to C'.

Clearly, strong d-rigidity is a necessary condition for an N-element causal set C to faithfully embed into an Alexandrov interval of flat d-dimensional spacetime of volume V as long as $N \gg 1$. On the other hand, weak d-rigidity is a rather weak necessary condition for C to faithfully embed into a d-dimensional curved spacetime, since the only requirement is that there exist a local or strongly d-rigid sub-causal set C' in C. Indeed, in this case, one should expect a whole family of strongly d-rigid subcausal sets $\{C_i'\}$ in C for fixed d. However, a straightforward analysis of this case is far from clear at the moment and we leave this for future investigations.

We summarize the above in the following Claim:

Claim 1 Let C be an N-element causal set that faithfully embeds into an Alexandrov interval $I[p, q]$ in a d-dimensional spacetime such that the discreteness scale is much smaller than the scale of flatness everywhere. Then there exists a subcausal set $C' \subset C$ of cardinality $N' \gg 1$ such that C' is strongly d-rigid. Moreover, if $I[p, q]$ is an Alexandrov interval in d-dimensional Minkowski spacetime then for large enough N, C is itself strongly d-rigid.

While the above arguments require that N and N' be arbitrarily large in order to suppress fluctuations, the simulations that we will present in the next section show that the necessary condition works extremely well even for N' values as low as 100 for a *single* "typical" realization of C'. This is true both for the flat spacetime case as well as for regions where the scale of flatness is large. Of course, for a generic curved spacetime, one does need to go to higher densities, but here too, there is strong evidence that the numbers can be relatively small.

Could this condition also be sufficient for manifoldlikeness? As discussed above, in the general case, it clearly is insufficient since one needs requirements on an appropriately chosen family of strongly d-rigid sub-causal sets in C. On the other hand, it is a plausible sufficiency condition for a causal set to be faithfully embeddable into an interval in d-dimensional Minkowski spacetime. There are several hints that support this. We first note that the interval abundance profile for generic causal sets or Kleitman Rothschild posets [14] which dominate the class of posets for large N differs vastly from $\langle N_m^d \rangle$. This difference in profile is easy to understand: these posets have a large number of links but almost no two or three element intervals. Thus, even at relatively small m, the interval abundances differ drastically from $\langle N_m^d \rangle$. Similar arguments can be made for the multiple layered class of causal sets studied in [15,16] which are subdominant but are also largely devoid of small intervals with $m > 1$. In Sec. V E we show the interval abundances for a chain and a Kleitman-Rothschild (KR) poset. Another example of a nonmanifoldlike causal set is the $2D$ orders corresponding to the crystalline phase of [9]. These are again layered, much like the KR posets, but here too, there is a large deviation from the flat spacetime $\langle N_m^d \rangle$. Of course such examples cannot suffice since the space of causal sets is littered with those that have no simple characterization. Hence we cannot at the moment prove that strong d-rigidity is violated for *all* causal sets which do not faithfully embed into an interval in flat spacetime.

On the other hand, as shown in Sec. III A the ratio $\langle N_m^d \rangle / \langle N_0^d \rangle$ is scale invariant in the limit $N \to \infty$.

[1] We will henceforth always assume that this condition is met.

In particular, this mimics the scale invariance of flat spacetime. Prompted by discussions with Sorkin we conjecture:

Conjecture 1 If the interval abundances $N_m(C)$ for an N-element causal set C are such that that $N_m(C) \sim \langle N_m^d \rangle \times (N \pm \sqrt{N})$ for some d in the large N limit, then C faithfully embeds into an Alexandrov interval in d-dimensional Minkowski spacetime.

In other words, we suggest that $\langle N_m^d \rangle$ provides a rigidity condition for a causal set to be approximated by an Alexandrov interval in Minkowski spacetime. A continuum version of this would require C to, moreover, be scale invariant or homogeneous, and it would be interesting to explore whether there are examples of homogeneous orders like the Box spaces [17] which could provide counter-examples to the conjecture.[2]

V. SIMULATIONS

We now show evidence for the above results and conjectures using simulations of relatively small causal sets. We consider causal sets that are sprinkled into flat and curved spacetime as well as nonmanifoldlike causal sets, using the CACTUS code causal set framework [18,19]. In particular, we perform our test on causal sets discretizations of flat spacetime for $d = 2, \ldots, 4$ as well as on the 2-d cut-trousers and the flat geometries on $S^1 \times \mathbb{R}$ and $T^2 \times \mathbb{R}$. As examples of curved spacetime we consider Friedmann-Robertson-Walker (FRW) spacetimes for $d = 4$ including de Sitter spacetime, both for small and large scales of flatness, and find significant deviations from the flat spacetime curves in the case of a small scale of flatness. All these examples provide ample support for Claim 1 even for relatively small N. Next, we consider simulations of causal sets generated by transitive percolation for the specific cases studied in [10] and show that they do not pass our test for manifoldlikeness. Finally, as support for our conjecture, we examine distinctly nonmanifoldlike causal sets, a chain and the class of Kleitman-Rothschild causal sets and show that, as expected, they fail our test of manifoldlikeness.

Once the causal set C is simulated, the interval abundances can be obtained within an appropriately chosen order interval $I[p, q] \subseteq C$. We employ two different procedures for this purpose. The first procedure is a test of locality of an entire causal set. Here we consider sprinklings into a large interval in flat spacetime and "cap" C to the past and the future by adding a pair of extra elements p, q so that $I[p, q] = C$. This allows us to measure the interval abundance for the entire causal set. It is especially useful when comparing the results from simulations into flat spacetime with the analytic plots. Thus, we do not look for local regions in a given causal set, but test for the locality of an entire causal set, or in the language of the previous section whether it is strongly d-rigid for some d.

The second procedure is for finding local regions in a causal set C which may not itself be local. Here we pick out an element $p \in C$ and then examine the set of order intervals to which p belongs. By comparing with the $\langle N_m^d \rangle$ curves, one can then identify which of these order intervals might serve as a local neighborhood of p. This method has two hurdles we must overcome.

First, we cannot control the location of p in the embedding spacetime. This makes it hard to find intervals that sample a specific feature of a spacetime, say a singularity in the cut-trousers topology of 5.2 or a point close to the origin of the FRW spacetime. A little control can be exerted using the fact that the current CACTUS code uses a natural labeling of the causal set, namely if $p \prec q$ then the labels satisfy $l(p) < l(q)$. Thus, picking a point with a low/high labeling allows us to choose the lower/upper area of the region we sprinkled in.

Second, the number of intervals that contain an element can be very large even for moderately sized causal sets. This can be ameliorated by only examining intervals within a certain size range.

Thus our second procedure will be to pick an appropriate element in the causal set and then examine all intervals, within a certain size range, that contain this element.

A. Flat spacetime simulations

We first consider the class of causal sets obtained via a Poisson sprinkling into flat spacetime intervals with $\langle N \rangle = 10^d$ elements using the existing CACTUS code and calculate N_m^d for each realization of a sprinkled causal set, for $d = 2, \ldots, 4$. We consider 1000 realizations in each case and calculate the standard deviation for the interval abundances. We find a remarkable agreement with the analytic curve for N_m^d as shown in Fig. 6, where we have also plotted the analytic curves for $N \pm \sqrt{N}$.

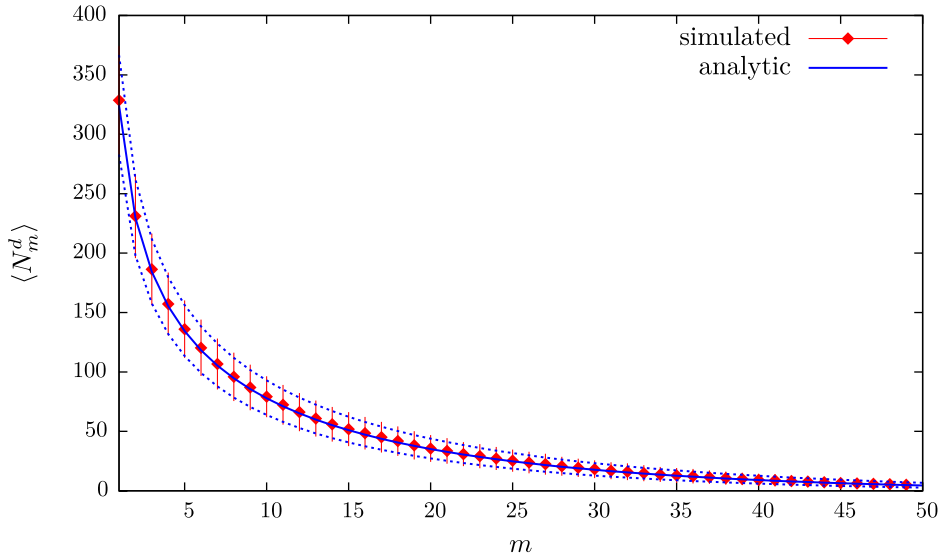
We also find that for single realizations of a sprinkled causal set, the distribution of N_m^d lies well within these curves as shown in 7. This plot also shows that the abundance can be used as a continuum dimension estimator. For causal sets that are nonmanifoldlike this will give a null result since the profile of $N_m(C)$ will not match that of the continuum for any d.

B. Examining other topologies

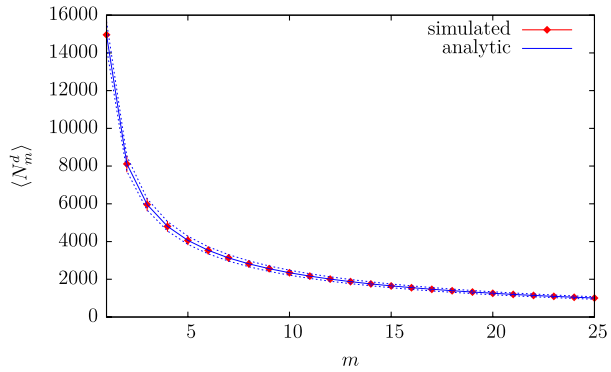
As the simplest generalization of flat spacetime intervals, we consider causal sets that are sprinkled into flat geometries with nontrivial spatial topology. An example of this is a "cut-trousers" topology in 2-d, with two disjoint spatial intervals $I \cup I$ joining up to give a single spatial interval I, as depicted in Fig. 8.

For the plot in Fig. 9 we obtained 100 realizations of a 1000-element causal set. For each of these, we picked large

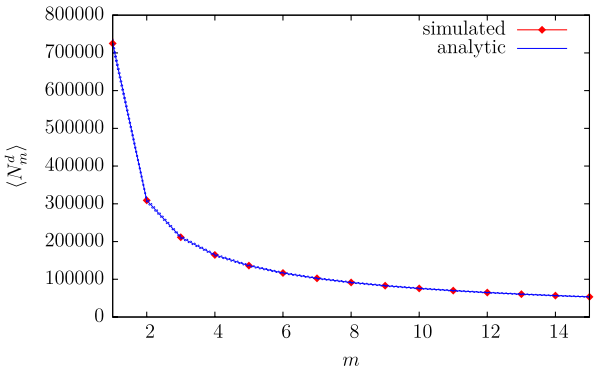
[2]We thank Rafael Sorkin for discussions on the continuum limit.



(a) 2d -100 Points



(b) 3d - 1000 Points



(c) 4d - 10000 Points

FIG. 6 (color online). Simulations of the expectation value of interval abundances $\langle N_m(C) \rangle$ in flat space for $N = 10^d$ element causal sets obtained by sprinkling 1000 times into an interval in flat spacetime. The red dots depict the simulations along with error bars. The solid blue line is $\langle N_m^d \rangle(N)$ while the dotted blue lines are $\langle N_m^d \rangle(N \pm \sqrt{N})$. (a) Depicts 100 Points in 2d. (b) 1000 Points in 3d and (c) 10000 Points in 4d.

intervals by choosing a minimal element p and a maximal element q such that $|I[p, q]|$ is the largest interval. Because of the nature of the topology we are considering, these intervals are "incomplete" if taken to be embedded in flat spacetime, as shown in Fig. 8. We find that this size fluctuates by 380.68 ± 14.01 and thus, we can average over the 100 realizations to obtain the expectation value of the interval abundances for 380 element causal sets. This is within the standard deviation or fluctuation $\sqrt{380} \approx 19.5$ expected. As shown in Fig. 9, the curves $\langle N_m(C') \rangle$ exhibit a clear deviation from flat spacetime.

We also test causal sets sprinkled into $d = 2$ and $d = 3$ flat spacetimes with toroidal spatial topologies, i.e., with $M \equiv \mathbb{R} \times S^1$ (the cylinder) and $M \equiv \mathbb{R} \times T^2$, respectively. For the large intervals in $d = 2$, we generate 100 realizations of 100-element causal sets via sprinkling and for $d = 3$, 100 we generate 100 realizations of 1000-element causal sets. The results are shown in Fig. 10.

For the small intervals we look at single realizations of 10000-element causal sets and examines intervals of size 100 in both cases. We take more points to obtain a higher density causal set, which allows us to find 100-element intervals that do not probe the topology. As expected, for both $d = 2$ and $d = 3$ the large intervals, which wrap around the compact spatial topology, have a distribution of intervals which has large deviations from the flat spacetime curve, but most of the small intervals do not. That some of the small intervals probe the topology of the spacetime is due to the nonlocality of the causal set. There will always be some small intervals that are almost lightlike, and thus probe the topology of the torus. We illustrate both the "nonlocal" and the "local" intervals in Fig. 11.

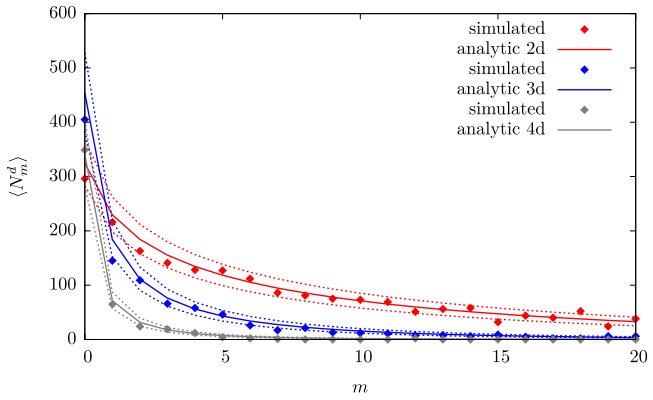


FIG. 7 (color online). Simulations in flat space for single realizations of $N = 100$ element causal sets obtained via Poisson sprinkling into flat spacetime intervals. The dots represent the simulations for a single causal set while the solid and dotted lines are $\langle N_m^d \rangle(N)$ and $\langle N_m^d \rangle(N \pm \sqrt{N})$, respectively. The agreement is striking.

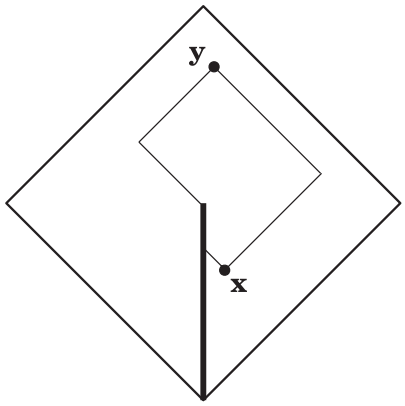


FIG. 8. A sketch of the cut-trousers topology. The Alexandrov neighborhood between the points x and y is modified by the cut.

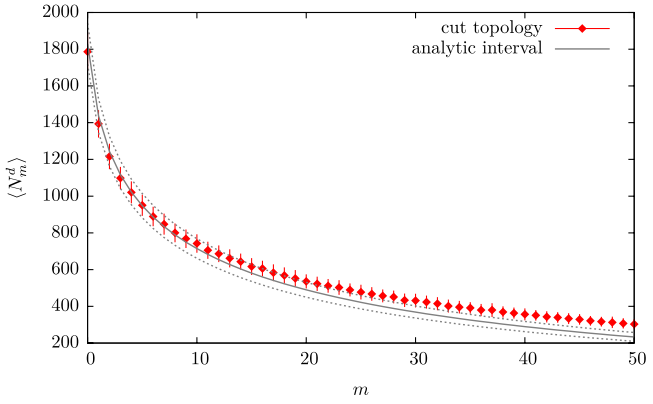


FIG. 9 (color online). $\langle N_m(C) \rangle$ for the largest intervals contained in 100 realizations of $N = 1000$ element causal sets obtained from sprinklings into the cut-trousers topology.

C. Curved spacetime: FRW and de Sitter spacetimes

Next, we consider causal sets which are sprinkled into 4-d $k = 0$ FRW spacetime with metric

$$ds^2 = -dt^2 + a(t)^2\left(\sum_{i=1}^{4}(dx^i)^2\right), \tag{30}$$

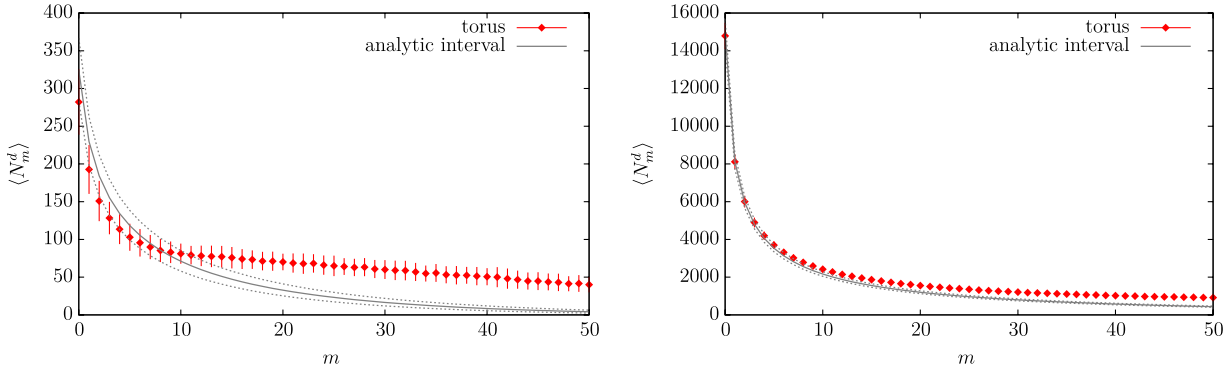
where

$$a(t) = a_0 t^q \quad \text{with:} \quad q = \frac{2}{3(1+w)} \tag{31}$$

with equation of state $p = w\rho$. We show our results from simulations for $w = 0, 1/3$ and the de Sitter case $w = -1$ (resp. matter, radiation and cosmological constant dominated) as examples. In the de Sitter case the de Sitter radius arises as a new free parameter, we chose a radius of $\ell = 1.3$. For each of these we examine 100 realizations of $\langle N \rangle = 1000$-element causal sets and find the average interval abundances. We expect significant deviations from the flat spacetime case, due to the nontrivial spacetime curvature and our simulations do confirm this expectation.

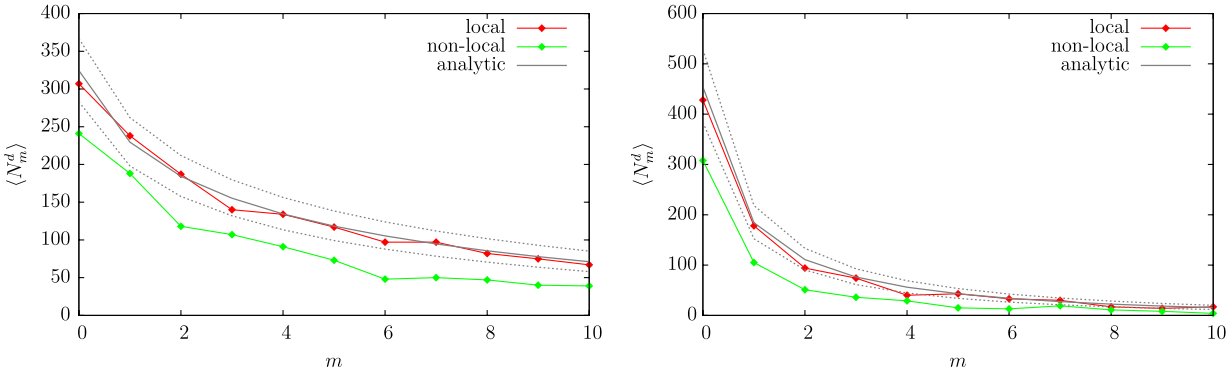
We first show that the large intervals do not follow the flat spacetime characteristic curve. As for the intervals with nontrivial topology, the size of the intervals varies as $N_x \pm \sqrt{N}x$. The results for all three choices clearly show the effect of curvature on the interval abundances. (c.f. Fig. 12).

To test manifoldlikeness of the causal set, we examine the intervals of size 100 containing a randomly chosen element in a single realization of an 10000 element causal set. For $w = \frac{1}{3}$ and $w = 0$ we found that there are intervals for which the abundances follow the flat spacetime curve and those which demonstrate significant deviations. As in the case of nontrivial topology, these latter interval neighborhoods must sample a region in which the scale of flatness ζ^{-1} is small, i.e., they are elongated intervals. The result is shown in Fig. 13. The colored boxes indicate the plotted intervals in comoving coordinates, while the shading indicates the scale factor $a(t)$ which needs to be taken into consideration when comparing the intervals. Although the elongated intervals do not fit the curve for flat 4d space, they are still not in agreement with higher or lower dimensional spacetime. If the same type of test is done on flat sprinkled causal sets, sprinkled sufficiently densely, there are no such stark differences between elongated and flat intervals. The same is true for intervals in de Sitter spacetime. In Fig. 14 we show intervals of size 100 and 2000 picked out of a 10000 element de Sitter sprinkling. They are as similar to each other as would be the case for flat space. This is because de Sitter space is maximally symmetric. While the 100 intervals are in agreement with flat 4d spacetime, the 2000 element sets all show a significant deviation from flatness.



(a) 2-d 100 elements (b) 3-d 1000 elements

FIG. 10 (color online). $\langle N_m(C) \rangle$ for large intervals in causal sets that are obtained from 100 sprinklings into flat spacetimes with toroidal spatial slices for (a) $d = 2$, (b) $d = 3$.



(a) 2-d local and non-local small intervals (b) 3-d local and non-local small intervals

FIG. 11 (color online). $N_m(C')$ for a small 100-element interval causal set C' contained in a single 10000-element causal set obtained by sprinkling into flat spacetime with toroidal spatial slices in (a) $d = 2$, (b) $d = 3$.

D. Causal sets obtained from transitive percolation

Causal sets should ideally not be created by sprinkling but grow naturally from some form of process. One such process, put forward by Rideout and Sorkin, is transitive percolation [10,20].

In transitive percolation the causal set grows iteratively, one element at a time. Each element gets added and then connected to the older elements with a certain probability. The probability for an element at step n to be connected to an element at step $n - 1$ is denoted as p, which is the only free parameter. In [10,20] causal sets of this type have been observed to have some manifoldlike characteristics. One thing that was examined was the functional relation between the proper time distance of two points and the volume that lies causally between them. It was found that for a variety of parameter combinations this curve can be well fit with the corresponding volume of a de Sitter spacetime, using the de Sitter radius ℓ and a proportionality factor between the length of the longest chain and the proper time τ as free parameters. To examine if percolated causal sets also appear manifoldlike under our new test we picked some of the possible parameter combinations, summarized in Table I.

In the paper they fit the curve to only those intervals which for a given proper time, had the largest volume. We followed this up in finding those intervals and measuring their interval abundances.

We created 100 percolated causal sets for each of the parameter combinations stated in Table I and calculated the

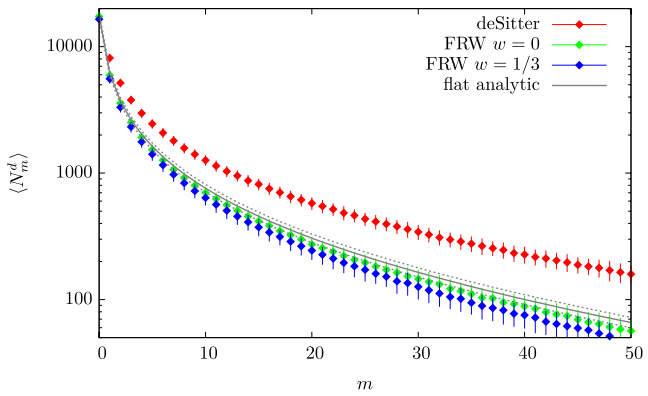
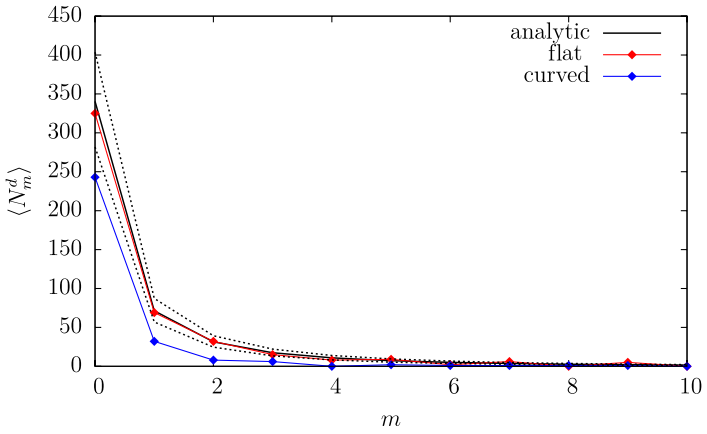
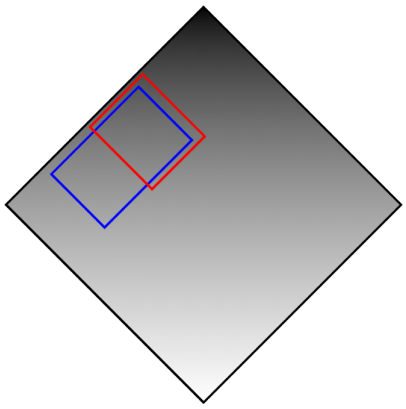
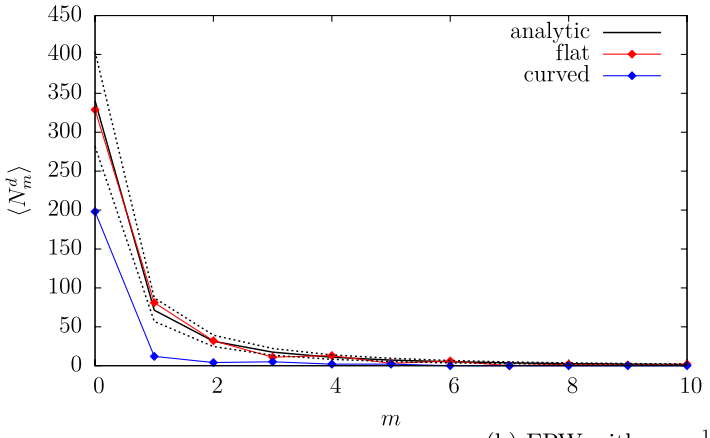


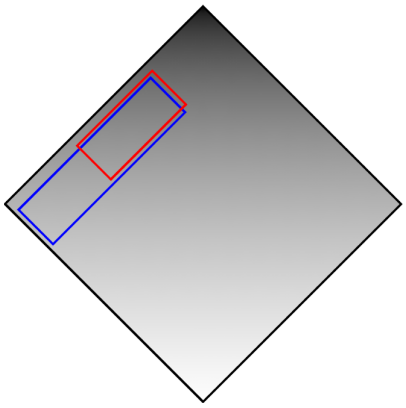
FIG. 12 (color online). $\langle N_m(C) \rangle$ for $N = 1000$ element causal sets obtained from sprinkling 100 times into $4d$ FRW spacetimes which are Λ, matter or radiation dominated.





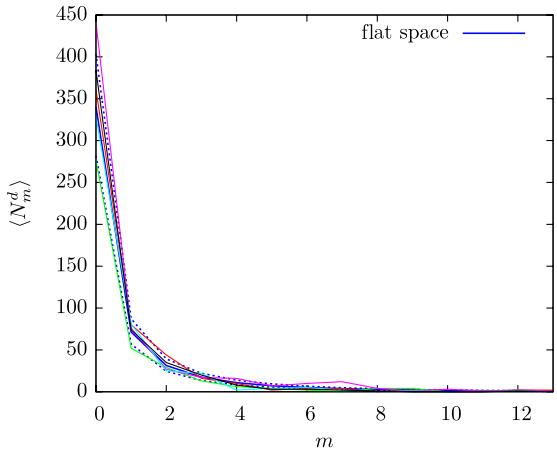
(a) FRW with $w = 0$



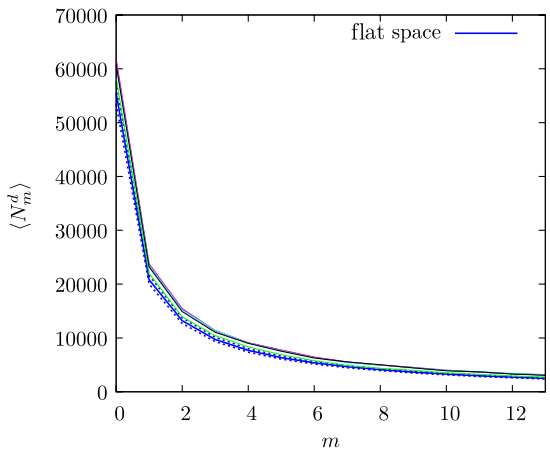


(b) FRW with $w = \frac{1}{3}$

FIG. 13 (color online). Single realizations of small interval causal sets C' contained in an $N = 10000$ element causal set C obtained from a sprinkling into $4d$ FRW spacetimes which are matter (a) or radiation (b) dominated. The sketches on the right-hand side show which intervals are local and which nonlocal, while the shading indicates the scale factor of the universe.



(a) 100 element deSitter intervals



(b) 2000 element deSitter intervals

FIG. 14 (color online). The above plots depict 5 realizations of smaller interval causal sets C' with $N' = 100$ (a) and 2000 (b), respectively, contained in an $N = 10000$ element causal set C approximated by $4d$ de Sitter spacetime. In the former, there is a strong agreement with the flat spacetime curve and with the latter, a strong deviation.

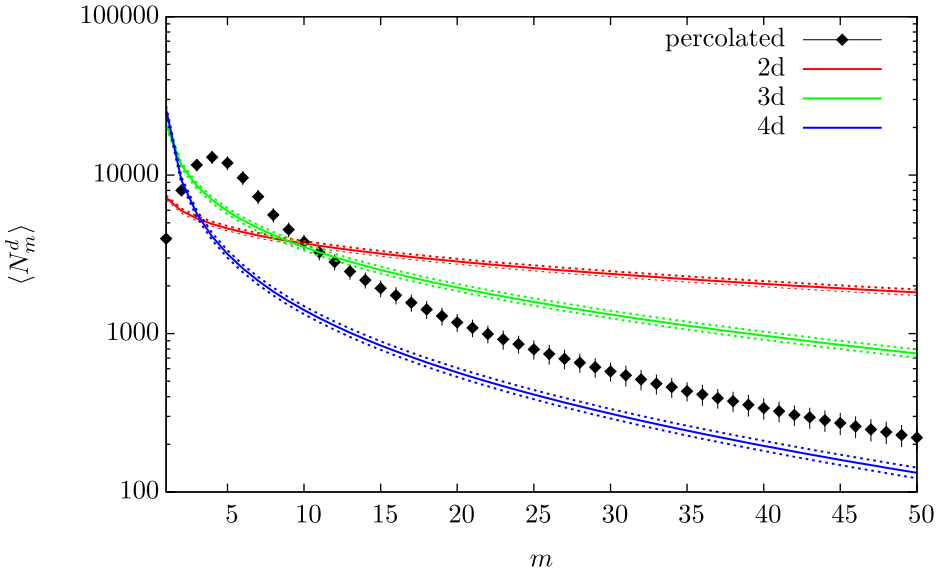
TABLE I. Three sets of parameter values from [10] which we have examined.

p	N	d	ℓ	m
0.03	1000	3d	2.331 ± 0.011	1.046 ± 0.006
0.01	2000	3d	4.086 ± 0.028	1.136 ± 0.006
0.005	15000	4d	6.20 ± 0.12	1.710 ± 0.013

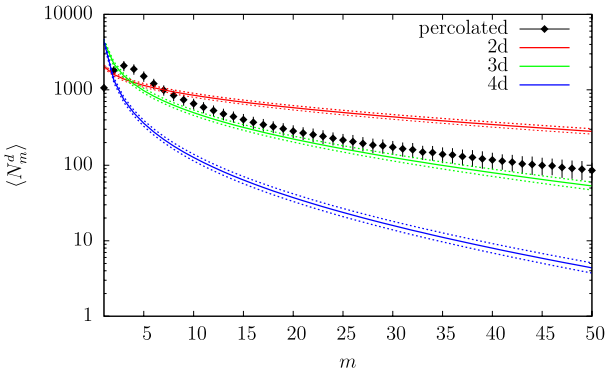
average interval abundances for intervals of different proper times. In Fig. 15 we plot this for the intervals of height 20. The behavior of the interval abundance is similar for heights between around 10–50 which is roughly the range of heights to which the de Sitter volume profile was fitted in [10]. Indeed there is a striking dissimilarity with flat spacetime: the abundances $N_m(C')$ show a maximum at some $m > 0$, unlike in flat spacetime for which $N_m(C')$ is a maximum for the links, i.e., $m = 0$. This maximum shifts to larger m as one examines larger Alexandrov intervals. One could perhaps argue that the different shape of the intervals could arise from curvature. However, the difference in shape persists even for very small intervals, which should look flat. This gives a strong indication that percolated causal sets are not manifoldlike.

Interestingly, while the interval abundance is clearly not that of flat spacetime it does converge towards the interval abundance of the dimension measured in [10] for large intervals. While the abundance of links and 1-element intervals for the percolated causal sets is very different from the analytic prediction, it falls off monotonically after the maximum and appears to get closer to the analytic prediction for manifoldlikeness.

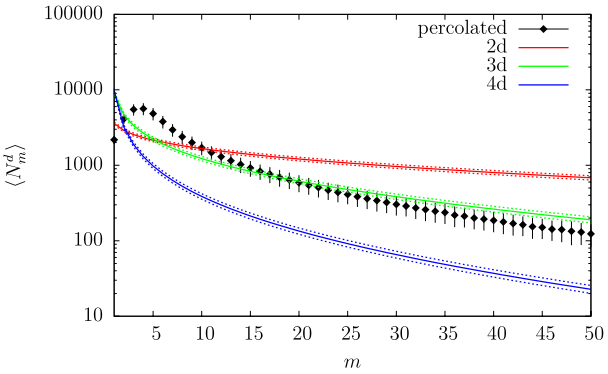
This suggests that perhaps the percolated causal sets, while not manifoldlike in the small, might be manifoldlike at a coarse grained level and hence satisfy our test. The coarse graining procedure involves keeping each element of C with a certain probability P. In Fig. 16 we



(a) p=0.005 N=15000, average size $\langle N \rangle = 1259.1 \pm 47.8$

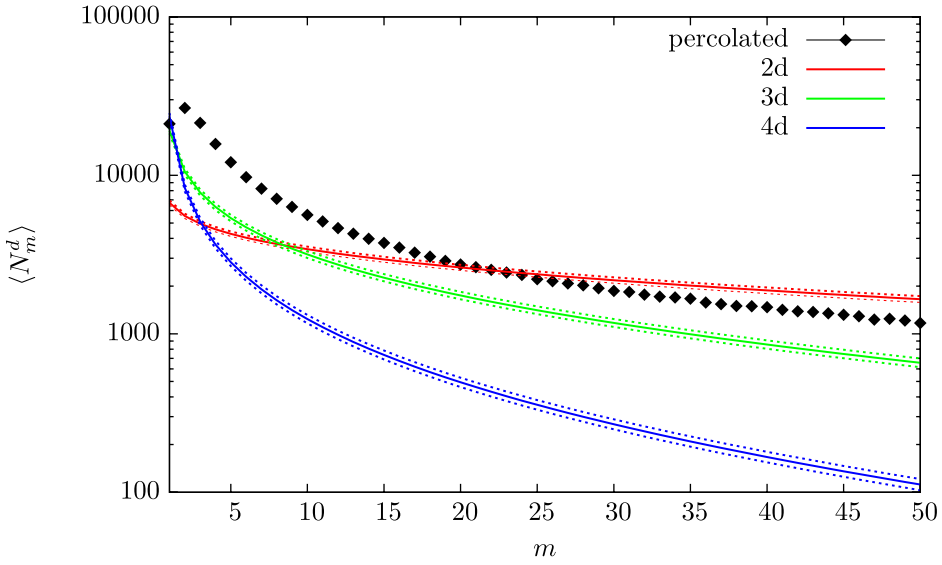


(b) p=0.03 N=1000, average size $\langle N \rangle = 335.2 \pm 25.9$

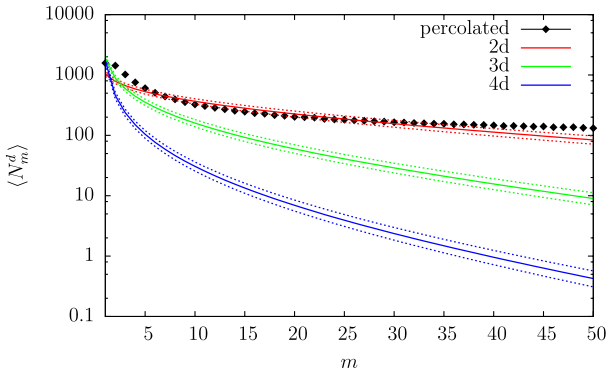


(c) p=0.01 N=2000, average size $\langle N \rangle = 695.7 \pm 53.5$

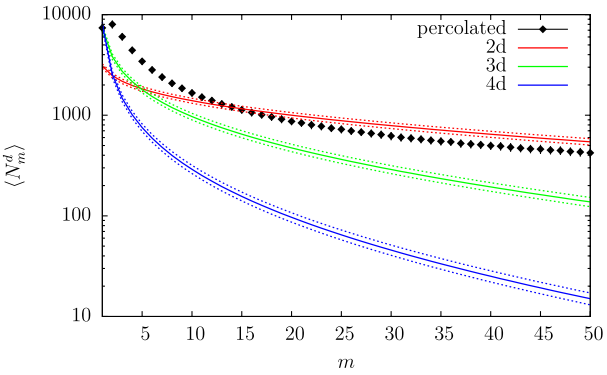
FIG. 15 (color online). The $\langle N_m(C) \rangle$ for percolated causal sets whose longest chain has 20 elements, are shown in black and compared with the $\langle N_m^d \rangle$ for (a) $d = 2$, (b) $d = 3$, (c) $d = 4$.



(a) p=0.005 N=15000, average size $\langle N \rangle = 1185.8 \pm 23.16$



(b) p=0.03 N=1000, average size $\langle N \rangle = 249.5 \pm 11.7$



(c) p=0.01 N=2000, average size $\langle N \rangle = 615.5 \pm 30.3$

FIG. 16 (color online). The $\langle N_m(C) \rangle$ for coarse grained percolated causal sets, whose longest chain is 20 elements long, is shown in black and compared with the $\langle N_m^d \rangle$ for (a) $d = 2$, (b) $d = 3$, (c) $d = 4$.

show the interval abundance for coarse grained transitive percolated causal sets. We choose $P = 0.25$, used the same values of p as before, and fixed the number of elements such that it would agree with Table I after coarse graining. Figure 16 shows that coarse graining does change the structure of the interval abundances significantly, since the maximal abundance moves to smaller m and the abundances become a monotonically decreasing function of m. However, despite this promising behavior, the detailed curve differs strongly from the $\langle N_m^d \rangle$. Of course the size of the parameter space makes it hard to make a stronger claim, but for coarse grainings where three-quarters, half, one quarter or one tenth of the points were kept we did not find agreement with the analytic curve for flat spacetime. Further study to compare to the interval abundance for de Sitter space might be useful, but first attempts at it do not indicate a substantial change in the results.

E. Nonmanifoldlike causal sets

There are several types of nonmanifoldlike causal sets that can be examined in this manner. The first that comes to mind is the totally ordered poset or chain and the totally unordered poset or antichain. The $N_m(C)$ for the former has a simple linearly decreasing behavior with m as depicted by the left-hand plot of Fig. 18, while all the $N_m(C)$ for the latter are simply zero. However, apart from such exotic

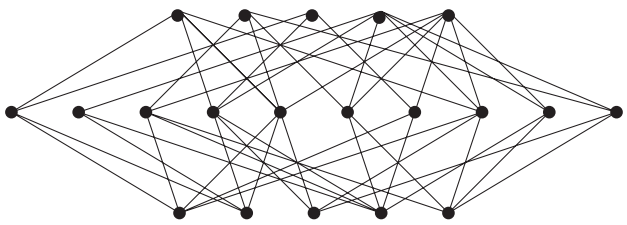
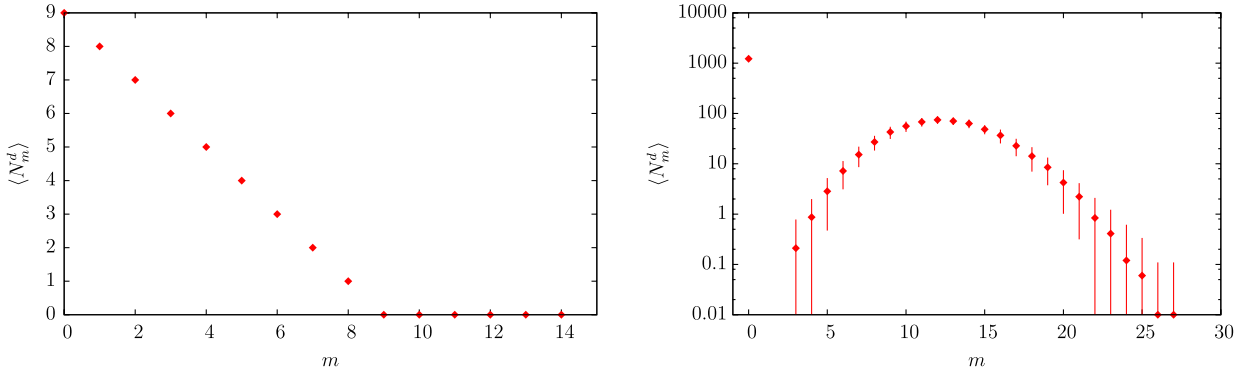


FIG. 17. An example of a small KR order.



(a) chain (b) KR-order

FIG. 18 (color online). The $N_m(C)$ are plotted for two different nonmanifoldlike causal sets. (a) The left-hand plot is for a chain of 10 elements and (b) the right-hand plot an average over 100 realizations of an 100-element KR order.

causal sets, one is interested in what $N_m(C)$ looks like for a more typical causal set. Here, we are aided by analytic results which tell us that as N becomes large, the set of causal sets is dominated by those that are of the Kleitman-Rothschild or KR form. A sketch of a small KR order is shown in Fig. 17. These are distinctly nonmanifoldlike since they possess only three "moments of time." A typical KR order has three layers with roughly $N/4$ minimal and maximal elements each and $N/2$-elements in the middle layer. Each minimal and maximal element are linked to roughly half the elements in the middle layer and every minimal element is related to every maximal element. Using David Rideout's CACTUS thorn to generate KR orders, we perform 100 realizations of $N = 100$ element KR orders to obtain $\langle N_m(C) \rangle$, which we show in Fig. 18.

VI. CONCLUSIONS AND OUTLOOK

In this work we demonstrated that interval abundances $N_m(C)$ in a causal set C provide an important class of observables for causal set theory. In particular, by comparing with the expectation value of the interval abundance $\langle N_m^d \rangle$ for an ensemble of causal sets obtained via a Poisson sprinkling into a flat spacetime interval, we showed that the $N_m(C)$ can be used to obtain "local" regions or subcausal sets $C' \subset C$ for a C which faithfully embeds into a general curved spacetime. Conversely, the existence of local regions in a causal set is a necessary test for manifoldlikeness of C and as a new continuum dimension estimator.

We began by obtaining closed form expressions for $\langle N_m^d \rangle$, and showed that the ratio $\langle N_m^d \rangle / \langle N_0^d \rangle$ is independent of the size N of the causal set to leading order. This scale invariance reflects that of flat spacetime and suggests a rigidity condition encoded by this class of observables. This prompted us to conjecture that knowing the $N_m(C)$ for $N \gg 1$ is sufficient to determine if C faithfully embeds into a flat spacetime interval of a given dimension. We tested these ideas with extensive simulations. We found that even for a relatively small ensemble of causal sets obtained via a Poisson sprinkling into an interval in Minkowski spacetime, the expectation value of the interval abundances matches very well with our analytic curves for $\langle N_m^d \rangle$. In addition, the agreement is very good even for a single causal set, up to the expected Poisson fluctuations in the size of the causal set, $N \pm \sqrt{N}$, as shown in Fig. 7. This suggests a prescription for extracting the continuum spacetime dimension from a causal set and thus a necessary condition for it to be faithfully embeddable into Minkowski spacetime.

In curved spacetime we considered both FRW and de Sitter spacetimes. The simulations agree with the $\langle N_m^d \rangle$ up to fluctuations as long as the scale of flatness is large, but deviate strongly from it when the scale of flatness is small. In the former case we found that the causal set represents a local or approximately flat spacetime region while in the latter case it is distinctly not local. We also examined the effect of topology on the $\langle N_m^d \rangle$ and found that again there is agreement with $\langle N_m^d \rangle$ only if the region explored in the spacetime is local.

We then examined a class of causal sets generated via transitive percolation to see if they passed our test for manifoldlikeness. In [10] it was claimed that a class of such causal sets possesses manifoldlike properties and has a Myrheim-Myer manifold dimension of 3 or 4. We tested several of these examples and found that they fail our test since the $N_m(C)$ do not agree with the $\langle N_m^d \rangle$ for *any* d. Hence we concluded that these causal sets are almost certainly *not* manifoldlike. However, it is possible that manifoldlikeness emerges after coarse graining. Preliminary tests showed that this is still not the case, but a more detailed study is currently underway [21].

Our simulations thus provide strong support for the rigidity conjecture in Sec. IV, namely, that knowing $N_m(C)$ is sufficient to determine whether C faithfully embeds into a flat spacetime interval or not. However, the

question arises whether and why this class of observables is more special than others. For one, it does provide an entire class of observables, and this itself is useful. But this is also the case for the abundances of chains $\langle C_m \rangle$, where an m-chain is a totally ordered m-element subset of C. Thus, a similar analysis may be possible using the $\langle C_m \rangle$. However, what distinguishes m-intervals from m-chains is that the former do in fact encode Lorentz-invariant *local* information while this is not true of the latter. In particular, the set of m-element intervals with fixed future endpoint q in a causal set that faithfully embeds into flat spacetime "layer" the past light cone of q along its past invariant hyperbolas, for each m. For example, elements which are linked to q lie within a volume $\sim\rho^{-1}$ to the past of q. Thus, as m increases, one explores regions further and further from the past null cone boundary of q. However, a chain lacks the same local information. For example a relation or 2-chain $p < q$ could either be a link with $|I[p, q]| = 0$ or separated by a very large interval size $|I[p, q]| \gg 1$. Thus, the number of relations to the future or the past of p can lie arbitrarily far from the boundary of the light cones from p – they are *not* nearest neighbors even in the Lorentzian sense. We believe it is this Lorentz-invariant locality of the $N_m(C)$ which makes them useful in defining locality. There are no other obvious candidates for families of observables and it is therefore tempting to conclude that the $N_m(C)$ are unique in this sense.

It is relatively straightforward to extend these calculations to a region of small curvature using Riemann normal coordinates and the techniques of [22,23]. However, the expressions for the $\langle N_m^d \rangle$ are far more complex, and extracting even an analytic curve from them requires more computationally intensive tools than in the flat spacetime case. We leave such investigations to future work where effects of curvature can be studied in greater detail than in the present work.

Finally, local regions in a manifoldlike causal set C could in principle be used to define a covering $\{C_i\}$ of $C = \cup_i C_i$, from which a nerve simplicial complex can be constructed. In the continuum, given a manifold M, a nerve simplicial complex can be obtained via a locally finite convex cover $\mathcal{O} = \{O_i\}$, $M = \cup_i O_i$, i.e., a cover in which (i) each set is convex, so that there exists a unique geodesic between any two points in the set, (ii) every $x \in M$ is contained in a finite number of elements of $\mathcal{O}$. The nerve simplicial complex is obtained from $\mathcal{O}$ by mapping each k-wise intersection of sets in $\mathcal{O}$ to a k-simplex. This simplicial complex is then homotopic to M as shown in [24]. In [25] a nerve simplicial complex was constructed to obtain the homology of spatial slices in both the continuum and in a causal set. One of the main obstructions to extending it to the full continuum spacetime or the full causal set is that a locally finite cover built out of Alexandrov intervals does not have an obvious local characterization. In particular, there is no way of distinguishing a convex from a nonconvex Alexandrov interval purely order theoretically. Our prescription for locality in the *discrete* case however overcomes this difficulty and it would be interesting to see if the spacetime homology could in fact be recovered from such a local covering of a causal set.

Our work opens up several new arenas in the study of discrete causal structure, some of which may lead us closer to answering fundamental questions in causal set theory.

ACKNOWLEDGMENTS

We thank Rafael Sorkin for discussions and David Rideout for help with the CACTUS code, which we used extensively. The authors would like to thank the ICTP in Trieste for their kind hospitality during parts of this work, L. G. would also like to thank the Raman Research Institute in Bangalore for a very productive visit. L. G. acknowledges support from the ERC-Advance grant 291092, "Exploring the Quantum Universe" (EQU) and of FNU, the Free Danish Research Council, from the grant "quantum gravity and the role of black holes".

APPENDIX A: A GENERAL FORMULA FOR HYPERGEOMETRIC FUNCTIONS

After solving the integration for the interval abundance in Sec. III it is necessary to find a closed form expression for sums of the type

$$\sum_{n=0}^{\infty} \frac{(-\rho V)^n}{n!} \left(\prod_{i=1}^{q} \frac{1}{n + a_i} \right) \frac{\Gamma(x(n + c))}{\Gamma(x(n + c) + m_1)} \frac{\Gamma(x(n + e))}{\Gamma(x(n + e) + m_2)}. \tag{A1}$$

In this expression it is assumed that m_1, m_2 are integers. We can then rewrite it as

$$\sum_{n=0}^{\infty} \frac{(-\rho V)^n}{n!} \left(\prod_{i=1}^{q} \frac{1}{n + a_i} \right) \left(\prod_{k=0}^{m_1 - 1} \frac{1}{x(n + c) + k} \right) \times \left(\prod_{l=0}^{m_2 - 1} \frac{1}{x(n + e) + l} \right). \tag{A2}$$

To express this in a closed form we factorize out the x and rewrite the products as gamma functions which gives

$$\sum_{n=0}^{\infty} \frac{(-\rho V)^n}{n!} x^{m_1 + m_2} \left(\prod_{i=1}^{q} \frac{\Gamma(n + a_i)}{\Gamma(n + a_i + 1)} \right) \times \left(\prod_{k=0}^{m_1 - 1} \frac{\Gamma(n + c + \frac{k}{x})}{\Gamma(n + c + \frac{k}{x} + 1)} \right) \left(\prod_{l=0}^{m_2 - 1} \frac{\Gamma(n + e + \frac{l}{x})}{\Gamma(n + e + \frac{l}{x} + 1)} \right). \tag{A3}$$

This can be rewritten using Pochhammer symbols $(a)_n = \Gamma(n + a)/\Gamma(a)$. Taking all the factors independent of n out of the sum leads to

$$x^{m_1+m_2}\Big(\prod_{i=1}^{q}\frac{\Gamma(a_i)}{\Gamma(a_i+1)}\Big)\Big(\prod_{k=0}^{m_1-1}\frac{\Gamma(c+\frac{k}{x})}{\Gamma(c+\frac{k}{x}+1)}\Big)\Big(\prod_{l=0}^{m_2-1}\frac{\Gamma(e+\frac{l}{x})}{\Gamma(e+\frac{l}{x}+1)}\Big)\times\sum_{n=0}^{\infty}\frac{(-\rho V)^n}{n!}\Big(\prod_{i=1}^{q}\frac{(a_i)_n}{(a_i+1)_n}\Big)$$
$$\times\Big(\prod_{k=0}^{m_1-1}\frac{(c+\frac{k}{x})_n}{(c+\frac{k}{x}+1)_n}\Big)\Big(\prod_{l=0}^{m_2-1}\frac{(e+\frac{l}{x})_n}{(e+\frac{l}{x}+1)_n}\Big), \tag{A4}$$

which can then be reexpressed in terms of the ${}_{l+m_1+m_2}F_{l+m_1+m_2}$ hypergeometric function

$$\Big(\prod_{i=1}^{q}\frac{1}{a_i}\Big)\Big(\prod_{k=0}^{m_1-1}\frac{1}{xc+k}\Big)\Big(\prod_{l=0}^{m_2-1}\frac{1}{xe+l}\Big)\times{}_{l+m_1+m_2}F_{l+m_1+m_2}\left(\begin{matrix}[a_i],\left[c+\frac{k}{x}\right],\left[e+\frac{l}{x}\right]\\ [a_i+1],\left[c+\frac{k}{x}+1\right],\left[e+\frac{l}{x}+1\right]\end{matrix}\;\middle|\;-\rho V\right), \tag{A5}$$

where the $[a_i]$ stand for $a_1,\ldots,a_q$ and $[c+\frac{k}{x}]$ (respective $[e+\frac{l}{x}]$) do stand for m_1 (m_2) terms in which k (l) varies from 0 to m_1-1 (0 to m_2-1). One last simplification allows us to write

$$\Big(\prod_{i=1}^{q}\frac{1}{a_i}\Big)\frac{\Gamma(xc)\Gamma(xe)}{\Gamma(xc+m_1)\Gamma(xe+m_2)}\times{}_{l+m_1+m_2}F_{l+m_1+m_2}\left(\begin{matrix}[a_i],\left[c+\frac{k}{x}\right],\left[e+\frac{l}{x}\right]\\ [a_i+1],\left[c+\frac{k}{x}+1\right],\left[e+\frac{l}{x}+1\right]\end{matrix}\;\middle|\;-\rho V\right). \tag{A6}$$

If c, e, a_i have an integer distance smaller than m_1 (m_2) it is possible to simplify this further since arguments of the hypergeometric function that arise on both sides cancel each other. In our calculation these simplifications will indeed take place but details are specific to each case.

APPENDIX B: DERIVING THE $m \neq 0$ CASE FROM THE $m = 0$ CASE USING HYPERGEOMETRIC FUNCTION IDENTITIES

We can derive $\langle N_m^d\rangle$ from $\langle N_0^d\rangle$ by taking derivatives. The expression for N_0^d is of the form

$$\chi(\rho V)^2{}_pF_p\left(\begin{matrix}a_1,\ldots,a_p\\ b_1,\ldots,b_p\end{matrix}\;\middle|\;-\rho V\right), \tag{B1}$$

where we have lumped some of the dimension dependent constants into the term χ. Using (17) the expression for the $\langle N_m^d\rangle$ is

$$\chi\frac{(-\rho)^{m+2}}{m!}\frac{\partial^m}{\partial\rho^m}V^2{}_pF_p\left(\begin{matrix}a_1,\ldots,a_p\\ b_1,\ldots,b_p\end{matrix}\;\middle|\;-\rho V\right). \tag{B2}$$

We use the identity [26]

$$\frac{\partial^m}{\partial z^m}{}_pF_p\left(\begin{matrix}a_1,\ldots,a_p\\ b_1,\ldots,b_p\end{matrix}\;\middle|\;z\right)$$
$$=\frac{\prod_{j=1}^{p}(a_j)_m}{\prod_{j=1}^{p}(b_j)_m}{}_pF_p\left(\begin{matrix}m+a_1,\ldots,m+a_p\\ m+b_1,\ldots,m+b_q\end{matrix}\;\middle|\;z\right) \tag{B3}$$

to simplify Eq. (B2) to

$$\chi\frac{(\rho V)^{m+2}}{m!}\frac{\prod_{j=1}^{p}(a_j)_m}{\prod_{j=1}^{p}(b_j)_m}{}_pF_p\left(\begin{matrix}a_1+m,\ldots,a_p+m\\ b_1+m,\ldots,b_p+m\end{matrix}\;\middle|\;-\rho V\right). \tag{B4}$$

This expression allows for further simplifications, depending on the a_j, b_j. These can be done for each individual case.

APPENDIX C: DERIVATION OF THE CONTINUUM LIMIT

The quantity we calculate is

$$S_m^d\equiv\lim_{\rho\to\infty}\frac{\langle N_m^d\rangle(\rho,V)}{\langle N_0^d\rangle(\rho,V)}, \tag{C1}$$

where $\langle N_m^d\rangle$ is given by Eq. (20). To investigate the $N\to\infty$ limit of the $\langle N_m^d\rangle$ we need a large $N=\rho V$ expansion of the hypergeometric functions that appear in Eq. (20), which when appropriately rearranged are of the form

$${}_dF_d\left(\begin{matrix}a_1,\ldots,a_d\\ a_1+2,\ldots,a_d+2\end{matrix}\;\middle|\;-N\right),$$
$$a_i=\frac{2i}{d}+m,\quad i=1,\ldots d-1,\quad a_d=1+m. \tag{C2}$$

We make repeated use of the identity Eq. (23) as well as the identity Eq. (24). Since the first identity cannot be used if two of the a_i are equal, or equal up to an integer whose absolute value is smaller than m_i, we will need to be careful in even dimensions. For $d=2$, in particular we need a different approach. We will thus treat odd and even dimensions separately and $d=2$ as a separate case.

1. Odd dimensions

This involves the most straightforward application of Eq. (23) to expand (C2):

$$
{}_dF_d\left(\begin{matrix} a_1,\ldots,a_d \\ a_1+2,\ldots,a_d+2 \end{matrix}\middle| -z\right)
= \prod_{j=1}^{d}(a_j)_2 \sum_{k=1}^{d}\sum_{j_1=0}^{1}\cdots\sum_{j_d=0}^{1}\frac{1}{a_k+j_k}\prod_{l=1}^{d}(-1)^{j_l}
\times \prod_{\substack{i=1\\ i\neq k}}^{d}\frac{1}{a_i+j_i-a_k-j_k}\,{}_1F_1\left(\begin{matrix} a_k+j_k \\ a_k+j_k+1 \end{matrix}\middle| -z\right). \tag{C3}
$$

For us $m_l = 2j_l = 0,\ 1$ so that $(1-m_l)_{j_l} = (-1)^{j_l}$. Using Eq. (24) in (C3).

$$
= \sum_{k=1}^{d}\sum_{j_1=0}^{1}\cdots\sum_{j_d=0}^{1}(z)^{-a_k-j_k}(\Gamma(a_k+j_k)-\Gamma(a_k+j_k,z))
\times \prod_{l=1}^{d}(a_l)_2(-1)^{j_l}\prod_{\substack{i=1\\ i\neq k}}^{d}\frac{1}{a_i+j_i-a_k-j_k}. \tag{C4}
$$

For $z\to\infty$ the terms containing $\Gamma(a,z)$ fallofflike $e^{-z}z^{a-1}$, c.f. Eq. (25) and do not contribute in the large z limit. Thus, to leading order the hypergeometric function is a power series with terms $z^{-a_i-j_i}$. The leading order term is therefore $a_1 = \frac{2}{d}+m$, $j_1 = 0$, while the next-to-leading order is $a_d = 1+m$ for $d=3$ and $a_2 = \frac{4}{d}+m$ for $d\geq 5$. We then only need to calculate the case $k=1$, $j_k = 0$.

Combining the products in Eq. (C4), we then sum over the j_i,

$$
\sum_{j_i=0}^{1}\frac{(a_i)_2(-1)^{j_i}}{a_i+j_i-a_1} = (a_i)_2\left(\frac{1}{a_i-a_1}-\frac{1}{a_i+1-a_1}\right)
= \frac{(a_i)_2}{(a_i-a_1)(a_i+1-a_1)} \tag{C5}
$$

after which we take the product over i to obtain

$$
(a_k)_2\prod_{\substack{i=1\\ i\neq k}}^{d}\frac{(a_i)_2}{(a_i-a_k)(a_i+1-a_k)}
= \left(\frac{2}{d}+m\right)_2\frac{(m+1)_2}{(1-\frac{2}{d})(2-\frac{2}{d})}\prod_{i=2}^{d-1}\frac{(\frac{2}{d}i+1)_2}{\frac{2}{d}(i-1)(\frac{2}{d}(i-1)+1)}. \tag{C6}
$$

We do the product for the different parts separately:

$$
\left(\frac{2}{d}+m\right)_2\prod_{i=2}^{d-1}\left(\frac{2}{d}i+1\right)_2 = \prod_{i=1}^{d-1}\left(\frac{2}{d}i+m\right)\left(\frac{2}{d}i+m+1\right) = \left(\frac{2}{d}\right)^{2d-2}\left(\frac{d}{2}m+1\right)_{d-1}\left(\frac{d}{2}(m+1)+1\right)_{d-1},
$$
$$
\prod_{i=2}^{d-1}\frac{1}{\frac{2}{d}(i-1)+1}\frac{d}{2(i-1)} = \left(\frac{2}{d}\right)^{-2d+2}\frac{1}{\Gamma(d-1)(\frac{d}{2}-1)_{d-2}}. \tag{C7}
$$

These are combined to find

$$
z^{-\frac{2}{d}-m}\frac{2\Gamma(\frac{2}{d}+m)(m+1)(m+2)}{\Gamma(d)(d-2))(\frac{d}{2}+1)_{d-2}}\left(\frac{d}{2}(m+1)+1\right)_{d-1}\left(\frac{d}{2}m+1\right)_{d-1} + \begin{cases}\mathcal{O}(z^{-m-1}) & \text{if } d=3\\ \mathcal{O}(z^{-\frac{4}{d}-m}) & \text{if } d\geq 5.\end{cases} \tag{C8}
$$

Inserting this into Eq. (20) for large N gives

$$
N_m^d(N) = \frac{N^{2-\frac{2}{d}}}{m!}\Gamma\left(\frac{2}{d}+m\right)\frac{\Gamma(d)}{(\frac{d}{2}-1)(\frac{d}{2}+1)_{d-2}} + \begin{cases}\mathcal{O}(N) & \text{if } d=3\\ \mathcal{O}(N^{2-\frac{4}{d}}) & \text{if } d\geq 5\end{cases}, \tag{C9}
$$

which gives us Eq. (21).

2. Even dimensions

In even dimensions it is possible for two of the a_i to be equal, or equal up to an integer whose absolute value is less than m_i. This therefore requires more care. The arguments of the hypergeometric functions in Eq. (20) however do admit a nondegenerate split:

$$
{}_dF_d\left(\begin{matrix} \frac{2}{d}+m,\ldots,1-\frac{2}{d}+m,\, m+1,\, \frac{2}{d}+m+1,\ldots,2-\frac{2}{d}+m,\, m+1 \\ \underbrace{\frac{2}{d}+m+2,\ldots,3-\frac{2}{d}+m,}_{\text{the first }\frac{d}{2}-1\text{ terms}}\ \underbrace{m+3}_{\text{the }\frac{d}{2}\text{th term}}\ ,\ \underbrace{\frac{2}{d}+m+3,\ldots,4-\frac{2}{d}+m,\, m+3}_{\text{the }\frac{d}{2}-1\text{ terms before the last}} \end{matrix}\middle| -z\right), \tag{C10}
$$

which can be shuffled to simplify the calculation. Exchanging the first $\frac{d}{2}-1$ terms and the $\frac{d}{2}-1$ terms before the last in the upper row changes the relationship between the top and bottom row. Instead of a hypergeometric function of the form (C2) we now have one of the form

$$ {}_dF_d\left(\begin{matrix} a_1,\ldots,a_{\frac{d}{2}-1},a_{\frac{d}{2}},a_{\frac{d}{2}+1},\ldots,a_{d-1},a_d \\ a_1+1,\ldots,a_{\frac{d}{2}-1}+1,a_{\frac{d}{2}}+2,a_{\frac{d}{2}+1}+3,\ldots,a_{d-1}+3,a_d+2 \end{matrix}\middle| -z\right). \tag{C11}$$

We now proceed in two steps. The first is to use (23) on the first $\frac{d}{2}-1$ terms. Here $n=d/2-1$, $m_i=1 \quad \forall\ i\in[1,\ldots,d/2-1]$ and hence $j_i=0 \quad \forall\ i$. Thus

$$\begin{aligned} &{}_dF_d\left(\begin{matrix} \frac{2}{d}+m+1,\ldots,2-\frac{2}{d}+m,m+1,\frac{2}{d}+m,\ldots,1-\frac{2}{d}+m,m+1 \\ \frac{2}{d}+m+2,\ldots,3-\frac{2}{d}+m,m+3,\frac{2}{d}+m+3,\ldots,4-\frac{2}{d}+m,m+3 \end{matrix}\middle| -z\right) \\ &\quad=\sum_{\alpha=1}^{\frac{d}{2}-1}{}_{\frac{d}{2}+2}F_{\frac{d}{2}+2}\left(\begin{matrix} \frac{2}{d}\alpha+m+1,\frac{2}{d}+m,\ldots,1-\frac{2}{d}+m,m+1,m+1 \\ \frac{2}{d}\alpha+m+2,\frac{2}{d}+m+3,\ldots,4-\frac{2}{d}+m,m+3,m+3 \end{matrix}\middle| -z\right)\prod_{\substack{j=1\\ j\neq\alpha}}^{\frac{d}{2}-1}\frac{j+\frac{d}{2}(m+1)}{j-\alpha}. \end{aligned} \tag{C12}$$

Next we apply Eq. (23) for a second time, with $i=2,\ldots d/2+1$. Now $n=d/2$ and $m_i=3$ for $i\in[2,\ldots,d/2]$ while $m_{d/2+1}=2$. Thus,

$$\begin{aligned} &{}_{\frac{d}{2}+2}F_{\frac{d}{2}+2}\left(\begin{matrix} \frac{2}{d}\alpha+m+1,\frac{2}{d}+m,\ldots,1-\frac{2}{d}+m,m+1,m+1 \\ \frac{2}{d}\alpha+m+2,\frac{2}{d}+m+3,\ldots,4-\frac{2}{d}+m,m+3,m+3 \end{matrix}\middle| -z\right) \\ &\quad=\left(\prod_{j=1}^{\frac{d}{2}}\frac{(\frac{2}{d}j+m)_3}{2}\right)\cdot\frac{2}{m+3}\sum_{k=1}^{\frac{d}{2}}\sum_{j_1}^{2}\cdots\sum_{j_{\frac{d}{2}-1}}^{2}\sum_{j_{\frac{d}{2}}}^{1}\frac{(-1)^{j_{\frac{d}{2}}}}{\frac{2}{d}k+j_k+m}\left(\prod_{l=1}^{\frac{d}{2}-1}\frac{(-2)_{j_l}}{j_l!}\right) \\ &\quad\times\left(\prod_{\substack{i=1\\ i\neq k}}^{\frac{d}{2}}\frac{1}{\frac{2}{d}(i-k)+j_i-j_k}\right){}_3F_3\left(\begin{matrix} \frac{2}{d}k+m+j_k,\frac{2}{d}\alpha+m+1,m+1 \\ \frac{2}{d}k+m+j_k+1,\frac{2}{d}\alpha+m+2,m+3 \end{matrix}\middle| -z\right). \end{aligned} \tag{C13}$$

To take the limit $z\to\infty$ we need to expand this ${}_3F_3$ for large z. We cannot do the entire expansion because of the special cases where $\frac{2}{d}k+m+j_k=\frac{2}{d}\alpha+m+1$ and $\frac{2}{d}k+m+j_k=m+1$. We thus need to make the expansion for three different possibilities:

(i) $\frac{2}{d}k+m+j_k\neq\frac{2}{d}\alpha+m+1\neq m+1$ in this case we can find an exact expansion for all z

(ii) $\frac{2}{d}k+m+j_k=\frac{2}{d}\alpha+m+1$ for this case we can take an expansion in large z and find that its leading order contributions in that limit are $z^{-1-m}\log z$.

(iii) $\frac{2}{d}k+m+j_k=m+1$ we can again take a large z expansion and find terms proportional to $z^{-m-1}\log z$ in leading order.

The expansions for the first case, $\frac{2}{d}k+m+j_k\neq\frac{2}{d}\alpha+m+1\neq m+1$, can be found using Eq. (23)

$$\begin{aligned} &{}_3F_3\left(\begin{matrix} \frac{2}{d}k+m+j_k,\frac{2}{d}\alpha+m+1,m+1 \\ \frac{2}{d}k+m+j_k+1,\frac{2}{d}\alpha+m+2,m+3 \end{matrix}\middle| -z\right) \\ &\quad=e^{-z}\frac{(2+m)(\frac{2}{d}\alpha+m+1)(\frac{2}{d}k+m+j_k)}{\frac{2}{d}\alpha(\frac{2}{d}k+j_k-1)}-\frac{(1+m)(2+m)(\frac{2}{d}\alpha+m+1)(\frac{2}{d}k+m+j_k)}{(\frac{2}{d}\alpha-1)(\frac{2}{d}k+j_k-2)}z^{-2-m}(\Gamma(2+m) \\ &\quad-\Gamma(2+m,z))+\frac{(2+m)(\frac{2}{d}\alpha+m+1)(\frac{2}{d}k+m+j_k)}{\frac{2}{d}\alpha(\frac{2}{d}k+j_k-1)}z^{-1-m}(\Gamma(2+m)-\Gamma(2+m,z)) \\ &\quad-\frac{(2+m)(1+m)(\frac{2}{d}\alpha+m+1)(\frac{2}{d}k+m+j_k)}{\frac{2}{d}\alpha(\frac{2}{d}\alpha-1)(\frac{2}{d}(\alpha-k)+1-j_k)}z^{-1-m-\frac{2}{d}\alpha}\left(\Gamma\left(\frac{2}{d}\alpha+1+m\right)-\Gamma\left(\frac{2}{d}\alpha+1+m,z\right)\right) \\ &\quad-\frac{(2+m)(1+m)(\frac{2}{d}\alpha+m+1)(\frac{2}{d}k+m+j_k)}{(\frac{2}{d}k+j_k-2)(\frac{2}{d}k+j_k-1)(\frac{2}{d}(k-\alpha)+j_k-1)}z^{-j_k-m-\frac{2}{d}k}\left(\Gamma\left(j_k+m+\frac{2}{d}k\right)-\Gamma\left(j_k+m+\frac{2}{d}k,z\right)\right). \end{aligned} \tag{C14}$$

For the second and third case we use MATHEMATICA to obtain the expansion. In the second case, $\frac{2}{d}k + m + j_k = \frac{2}{d}\alpha + m + 1$,

$$
\begin{aligned}
&{}_3F_3\left(\begin{matrix}\frac{2}{d}\alpha+m+1,\frac{2}{d}\alpha+m+1,m+1\\ \frac{2}{d}\alpha+m+2,\frac{2}{d}\alpha+m+2,m+3\end{matrix}\,\middle|\,-z\right)\\
&=\frac{(m+2)(\frac{2}{d}\alpha+m+1)^2}{(\frac{2}{d}\alpha)^2}z^{-m-1}\Gamma(m+2)-\frac{(m+1)(\frac{2}{d}\alpha+m+1)^2}{(\frac{2}{d}\alpha-1)^2}z^{-m-2}\Gamma(m+3)\\
&\quad+\frac{(m+1)(m+2)(\frac{2}{d}\alpha+m+1)}{(\frac{2}{d}\alpha-1)\frac{2}{d}\alpha}\Gamma\left(m+\frac{2}{d}\alpha+2\right)z^{-\frac{2}{d}\alpha-m-2}\left(\frac{2\frac{2}{d}\alpha-1}{(\frac{2}{d}\alpha-1)\frac{2}{d}\alpha}+\psi^{(0)}\left(m+\frac{2}{d}\alpha+1\right)+\log(z)\right)+\cdots
\end{aligned}
\tag{C15}
$$

and in the third case, $\frac{2}{d}k + m + j_k = m + 1$,

$$
\begin{aligned}
{}_3F_3\left(\begin{matrix}m+1,\frac{2}{d}\alpha+m+1,m+1\\ m+2,\frac{2}{d}\alpha+m+2,m+3\end{matrix}\,\middle|\,-z\right)&=-\frac{(1+m)^2(2+m)(1+m+\frac{2}{d}\alpha)}{(\frac{2}{d}\alpha)^2(\frac{2}{d}\alpha-1)}z^{-1-m-\frac{2}{d}\alpha}\Gamma\left(1+m+\frac{2}{d}\alpha\right)\\
&\quad-\frac{(m+1)(1+m+\frac{2}{d}\alpha)}{(\frac{2}{d}\alpha)^2}z^{-1-m}\Gamma(3+m)(2-\log z+\Psi^{(0)}(m+1))+\cdots
\end{aligned}
\tag{C16}
$$

The leading order terms are $z^{-\frac{2}{d}-m}$, for $k = 1$, $j_k = 0$ from Eq. (C14). The next-to-leading order term is $z^{-\frac{4}{d}-m}$ if $d > 4$, while for $d = 4$ it is $z^{-m-1}\log z$, from (C15) and (C16). Thus to leading order we find

$$
\begin{aligned}
&\left(\prod_{j=1}^{\frac{d}{2}}\frac{(\frac{2}{d}j+m)_3}{2}\right)\cdot\frac{2}{m+3}\sum_{j_1}^{2}\cdots\sum_{j_{\frac{d}{2}-1}}^{2}\sum_{j_{\frac{d}{2}}}^{1}\frac{1}{\frac{2}{d}1+m}\left(\prod_{l=1}^{\frac{d}{2}-1}\frac{(-2)_{j_l}}{j_l!}\prod_{i=2}^{\frac{d}{2}-1}\frac{1}{\frac{2}{d}(i-1)+j_i}\right)\\
&\quad\times\frac{(-1)^{j_{\frac{d}{2}}}}{1+j_{\frac{d}{2}}-\frac{2}{d}}\frac{(2+m)(1+m)(\frac{2}{d}\alpha+m+1)(\frac{2}{d}+m)}{(\frac{2}{d}-2)(\frac{2}{d}-1)(\frac{2}{d}(\alpha-1)+1)}z^{-m-\frac{2}{d}}\Gamma\left(m+\frac{2}{d}\right).
\end{aligned}
\tag{C17}
$$

The products in (C17) simplify as

$$
\begin{aligned}
\prod_{j=1}^{\frac{d}{2}}\frac{(\frac{2}{d}j+m)_3}{2}&=2^d d^{-\frac{3}{2}d}\frac{\Gamma(\frac{d}{2}(m+3)+1)}{\Gamma(\frac{d}{2}m+1)},\\
\sum_{j_l=0}^{2}\frac{(-2)_{j_l}}{2j_l!(\frac{2}{d}(l-1)+j_l)}&=\frac{d^3}{4(l-1)(l-1+d)((l-1)+\frac{d}{2})},\\
\prod_{l=2}^{\frac{d}{2}-1}\frac{d^3}{4(l-1)(l-1+d)((l-1)+\frac{d}{2})}&=\frac{2^{2-d}(d-2)(d-1)d^{-4+\frac{3}{2}d}}{\Gamma(\frac{3}{2}d-1)},
\end{aligned}
\tag{C18}
$$

so that (C17) simplifies to

$$
\frac{d\Gamma(\frac{d}{2}(m+3))}{(d-2)(d-1)}\frac{\Gamma(\frac{2}{d}+m)(m+1)(m+2)}{\Gamma(\frac{d}{2}m+1)\Gamma(\frac{3}{2}d-1)}\frac{\frac{d}{2}(m+1)+\alpha}{\frac{d}{2}+\alpha-1}z^{-\frac{2}{d}-m}.
\tag{C19}
$$

Inserting this into Eq. (C12) we can perform the summation over α:

$$
\sum_{\alpha=1}^{\frac{d}{2}-1}\frac{\frac{d}{2}(m+1)+\alpha}{\frac{d}{2}+\alpha-1}\prod_{\substack{j=1\\ j\neq\alpha}}^{\frac{d}{2}-1}\frac{j+\frac{d}{2}(m+1)}{j-\alpha}.
\tag{C20}
$$

The product in the above expression gives

$$
\begin{aligned}
&\left(\frac{d}{2}(m+1)+\alpha\right)\prod_{\substack{j=1\\ j\neq\alpha}}^{\frac{d}{2}-1}\frac{j+\frac{d}{2}(m+1)}{j-\alpha}\\
&\quad=\frac{(-1)^{\alpha-1}\Gamma(\frac{d}{2}(m+2))}{\Gamma(\alpha)\Gamma(\frac{d}{2}-\alpha)\Gamma(\frac{d}{2}(m+1)+1)}
\end{aligned}
\tag{C21}
$$

so that the sum reduces to

$$\sum_{\alpha=1}^{\frac{d}{2}-1} \frac{(-1)^{\alpha-1}\Gamma(\frac{d}{2}(m+2))}{\Gamma(\alpha)\Gamma(\frac{d}{2}-\alpha)\Gamma(\frac{d}{2}(m+1)+1)} \frac{1}{\alpha-1+\frac{d}{2}} = \frac{\Gamma(\frac{d}{2})(d-1)\Gamma(\frac{d}{2}(m+2))}{\Gamma(d)\Gamma(\frac{d}{2}(m+1)+1)}. \tag{C22}$$

Combining this with Eq. (C19) gives

$$z^{-m-\frac{2}{d}} \frac{d\Gamma(\frac{d}{2})}{(d-2)\Gamma(d)\Gamma(\frac{3}{2}d-1)} \Gamma\left(\frac{2}{d}+m\right)(m+1)(m+2) \times \left(\frac{d}{2}(m+1)+1\right)_{d-1} \left(\frac{d}{2}m+1\right)_{d-1}. \tag{C23}$$

Thus, we find

$$\langle N_m^d\rangle(N) = \frac{N^{2-\frac{2}{d}}}{m!} \frac{\Gamma(d+1)\Gamma(\frac{d}{2})}{(d-2)\Gamma(\frac{3}{2}d-1)} \Gamma\left(\frac{2}{d}+m\right) + \begin{cases} \mathcal{O}(N\log N) & \text{for } d=4 \\ \mathcal{O}(N^{2-\frac{4}{d}}) & \text{for } d>4. \end{cases}$$

Realizing that

$$\frac{d\Gamma(\frac{d}{2})}{\Gamma(\frac{3}{2}d-1)} = \frac{2}{(\frac{d}{2}+1)_{d-2}}, \tag{C24}$$

we note that this agrees with the expression for odd dimensions, and we can thus write

$$\langle N_m^d\rangle(N) = \frac{N^{2-\frac{2}{d}}}{m!} \Gamma\left(\frac{2}{d}+m\right) \frac{\Gamma(d)}{(\frac{d}{2}-1)(\frac{d}{2}+1)_{d-2}} + \begin{cases} \mathcal{O}(N) & \text{for } d=3 \\ \mathcal{O}(N\log N) & \text{for } d=4 \\ \mathcal{O}(N^{2-\frac{4}{d}}) & \text{for } d>4 \end{cases}$$

for all $d > 2$, which gives Eq. (21).

3. The case $d = 2$

$d = 2$ is a special case for which ${}_2F_2$ can be expanded for $z \to \infty$ using MATHEMATICA,

$${}_2F_2\left(\begin{matrix} m+1, m+1 \\ m+3, m+3 \end{matrix}\middle| -z\right) = z^{-1-m}\Gamma(3+m)(m+1)(m+2)\log z + \mathcal{O}(z^{-1-m}). \tag{C25}$$

Inserting this in (20) leads to

$$\langle N_m^2\rangle(N) = N\log N + \mathcal{O}(N), \tag{C26}$$

which gives

$$\lim_{\rho\to\infty} \frac{\langle N_m^2\rangle(\rho, V)}{\langle N_0^2\rangle(\rho, V)} = 1 \tag{C27}$$

and hence to leading order agrees with the expression for $d > 2$, so that we have finally recovered Eq. (21) for all $d \geq 2$

[1] L. Bombelli, J. Lee, D. Meyer, and R. D. Sorkin, Phys. Rev. Lett. **59**, 521 (1987).
[2] R. D. Sorkin, in *Lectures on Quantum Gravity. Proceedings, School of Quantum Gravity, Valdivia, Chile, January 4-14, 2002*, edited by A. Gomberoff and D. Marolf (Springer, Berlin, 2003), p. 305.
[3] F. Dowker, in *100 Years of Relativity Space-time Structure: Einstein and Beyond*, edited by Abhay Ashtekar (World Scientific, Singapore, 2005).
[4] S. Surya, arXiv:1103.6272.
[5] L. Bombelli, J. Henson, and R. D. Sorkin, Mod. Phys. Lett. A **24**, 2579 (2009).
[6] D. M. T. Benincasa and F. Dowker, Phys. Rev. Lett. **104**, 181301 (2010).
[7] F. Dowker and L. Glaser, Classical Quantum Gravity **30**, 195016 (2013).
[8] D. A. Meyer, Ph.D. thesis, M.I.T., 1988; G. Brightwell and R. Gregory, Phys. Rev. Lett. **66**, 260 (1991); S. Major, D. P. Rideout, and S. Surya, J. Math. Phys. (N.Y.) **48**, 032501 (2007); D. Rideout and P. Wallden, Classical Quantum Gravity **26**, 155013 (2009); S. Johnston, Classical Quantum Gravity **25**, 202001 (2008); Phys. Rev. Lett. **103**, 180401 (2009).
[9] S. Surya, Classical Quantum Gravity **29**, 132001 (2012).
[10] M. Ahmed and D. Rideout, Phys. Rev. D **81**, 083528 (2010).
[11] E. H. Kronheimer and R. Penrose, Proc. Cambridge Philos. Soc. **63**, 481 (1967); E. C. Zeeman, J. Math. Phys. (N.Y.) **5**, 490 (1964); R. P. Geroch, E. H. Kronheimer, and R. Penrose, Proc. R. Soc. A **327**, 545 (1972).
[12] S. W. Hawking, A. R. King, and P. J. Mccarthy, J. Math. Phys. (N.Y.) **17**, 174 (1976); D. B. Malament, J. Math. Phys. (N.Y.) **18**, 1399 (1977).
[13] Wolfram alpha functions. http://functions.wolfram.com/07.31.03.0010.01, 2013.
[14] D. Kleitman, and B. L. Rothschild, Trans. Am. Math. Soc. **205**, 205 (1975).
[15] D. Dhar, J. Math. Phys. (N.Y.) **19**, 1711 (1978).
[16] D. Dhar, Pac. J. Math. **90**, 299 (1980).
[17] B. Bollobas and G. Brightwell, SIAM J. Discrete Math. **10**, 318 (1997).
[18] G. Allen, T. Goodale, F. Löffler, D. Rideout, E. Schnetter, and E. L. Seidel, arXiv:1009.1341.
[19] T. Goodale, G. Allen, G. Lanfermann, J. Mass, T. Radke, E. Seidel, and O. Shalf, Vector and

Parallel Processing - VECPAR'2002, 5th International Conference, Lecture Notes in Computer Science, 2003, The CACTUS framework and toolkit: Design and applications (to be published).

[20] D. P. Rideout and R. D. Sorkin, Phys. Rev. D **61**, 024002 (1999).

[21] L. Glaser (to be published).

[22] S. Khetrapal and S. Surya, Classical Quantum Gravity **30**, 065005 (2013).

[23] M. Roy, D. Sinha, and Sumati Surya, Phys. Rev. D **87**, 044046 (2013).

[24] A. Weil, Commentarii mathematici Helvetici **26**, 119 (1952); G. De Rham, Annales de l'institut Fourier, **2**, 51 (1950).

[25] S. Major, D. P. Rideout, and S. Surya, J. Math. Phys. (N.Y.) **48**, 032501 (2007).

[26] Wolfram alpha functions. http://functions.wolfram.com/07.31.20.0011.01, 2013.