Living Reviews in Relativity (2019) 22:5
https://doi.org/10.1007/s41114-019-0023-1

REVIEW ARTICLE

The causal set approach to quantum gravity

Sumati Surya[1]

Received: 28 February 2019 / Accepted: 28 August 2019 / Published online: 27 September 2019

Abstract

The causal set theory (CST) approach to quantum gravity postulates that at the most fundamental level, spacetime is discrete, with the spacetime continuum replaced by locally finite posets or "causal sets". The partial order on a causal set represents a proto-causality relation while local finiteness encodes an intrinsic discreteness. In the continuum approximation the former corresponds to the spacetime causality relation and the latter to a fundamental spacetime atomicity, so that finite volume regions in the continuum contain only a finite number of causal set elements. CST is deeply rooted in the Lorentzian character of spacetime, where a primary role is played by the causal structure poset. Importantly, the assumption of a fundamental discreteness in CST does not violate local Lorentz invariance in the continuum approximation. On the other hand, the combination of discreteness and Lorentz invariance gives rise to a characteristic non-locality which distinguishes CST from most other approaches to quantum gravity. In this review we give a broad, semi-pedagogical introduction to CST, highlighting key results as well as some of the key open questions. This review is intended both for the beginner student in quantum gravity as well as more seasoned researchers in the field.

Keywords Causal set theory · Quantum gravity · Discreteness · Causality · Poset theory

Contents

✉ Sumati Surya
ssurya@rri.res.in

[1] Raman Research Institute, CV Raman Ave, Sadashivanagar, Bangalore 560080, India

1 Overview

In this review, causal set theory (CST) refers to the specific proposal made by Bombelli, Lee, Meyer and Sorkin (BLMS) in their 1987 paper (Bombelli et al. 1987). In CST, the space of Lorentzian geometries is replaced by the set of locally finite posets, or *causal sets*. These causal sets encode the twin principles of causality and discreteness. In the continuum approximation of CST, where elements of the causal set set represent spacetime events, the order relation on the causal set corresponds to the spacetime causal order and the cardinality of an "order interval" to the spacetime volume of the associated causal interval.

This review is intended as a semi-pedagogical introduction to CST. The aim is to give a broad survey of the main results and open questions and to direct the reader to some of the many interesting open research problems in CST, some of which are accessible even to the beginner.

We begin in Sect. 2 with a historical perspective on the ideas behind CST. The twin principles of discreteness and causality at the heart of CST have both been proposed—sometimes independently and sometimes together—starting with Riemann (1873) and Robb (1914, 1936), and somewhat later by Zeeman (1964), Kronheimer and Penrose (1967), Finkelstein (1969), Hemion (1988) and Myrheim (1978), culminating in the CST proposal of BLMS (Bombelli et al. 1987). The continuum approximation of CST is an implementation of a deep result in Lorentzian geometry due to Hawking et al. (1976) and its generalisation by Malament (1977), which states that the causal structure determines the conformal geometry of a future and past distinguishing causal spacetime. In following this history, the discussion will be necessarily somewhat technical.

For those unfamiliar with the terminology of causal structure we point to standard texts (Hawking and Ellis 1973; Beem et al. 1996; Wald 1984; Penrose 1972).

In Sect. 3, we state the CST proposal and describe its continuum approximation, in which spacetime causality is equivalent to the order relation and finite spacetime volumes to cardinality. Not all causal sets have a continuum approximation—in fact we will see that most do not. Those that do are referred to as *manifold-like*. Important to CST is its "Hauptvermutung" or fundamental conjecture, which roughly states that a manifold-like causal set is equivalent to the continuum spacetime, modulo differences up to the discreteness scale. Much of the discussion on the Hauptvermutung is centered on the question of how to estimate the closeness of Lorentzian manifolds or more generally, causal sets. While there is no full proof of the conjecture, there is a growing body of evidence in its favour as we will see in Sect. 4. An important outcome of CST discreteness in the continuum approximation is that it does not violate Lorentz invariance as shown in an elegant theorem by Bombelli et al. (2009). Because of the centrality of this result we review this construction in some detail. The combination of discreteness and Lorentz invariance moreover gives rise to an inherent and characteristic non-locality, which distinguishes CST from other discrete approaches. Following Sorkin (1997), we then discuss how the twin principles behind CST force us to take certain "forks in the road" to quantum gravity.

We present some of the key developments in CST in Sects. 4, 5 and 6. We begin with the kinematical structure of theory and the program of "geometric reconstruction" in Sect. 4. Here, the aim is to reconstruct manifold invariants from *order invariants* in a manifold-like causal set. These are functions on the causal set that are independent of the labelling or ordering of the elements in the causal set. Finding the appropriate order invariants that correspond to manifold invariants can be challenging, since there is little in the mathematics literature which correlates order theory to Lorentzian geometry via the CST continuum approximation. Extracting such invariants requires new technical tools and insights, sometimes requiring a rethink of familiar aspects of continuum Lorentzian geometry. We will describe some of the progress made in this direction over the years (Myrheim 1978; Brightwell and Gregory 1991; Meyer 1988; Bombelli and Meyer 1989; Bombelli 1987; Reid 2003; Major et al. 2007; Rideout and Wallden 2009; Sorkin 2007b; Benincasa and Dowker 2010; Benincasa 2013; Benincasa et al. 2011; Glaser and Surya 2013; Roy et al. 2013; Buck et al. 2015; Cunningham 2018a; Aghili et al. 2019; Eichhorn et al. 2019a). The correlation between order invariants and manifold invariants in the continuum approximation lends support for the Hauptvermutung and simultaneously establishes weaker, observable-dependent versions of the conjecture.

Somewhere between dynamics and kinematics is the study of quantum fields on manifold-like causal sets, which we describe in Sect. 5. The simplest system is free scalar field theory on a causal set approximated by d-dimensional Minkowski spacetime $\mathbb{M}^d$. Because causal sets do not admit a natural Hamiltonian framework, a fully covariant construction is required to obtain the quantum field theory vacuum. A natural starting point is the advanced and retarded Green functions for a free scalar field theory since it is defined using the causal structure of the spacetime. The explicit form for these Green functions were found for causal sets approximated by $\mathbb{M}^d$ for $d = 2, 4$ (Daughton 1993; Johnston 2008, 2010) as well as de Sitter spacetime (Dowker et al.

2017). In trying to find a quantisation scheme on the causal set without reference to the continuum, Johnston (2009) found a novel covariant definition of the discrete scalar field vacuum, starting from the covariantly defined Peierls' bracket formulation of quantum field theory. Subsequently Sorkin (2011a) showed that the construction is also valid in the continuum, and can be used to give an alternative definition of the quantum field theory vacuum. This *Sorkin–Johnston (SJ) vacuum* provides a new insight into quantum field theory and has stimulated the interest of the algebraic field theory community (Fewster and Verch 2012; Brum and Fredenhagen 2014; Fewster 2018). The SJ vacuum has also been used to calculate Sorkin's spacetime entanglement entropy (SSEE) (Bombelli et al. 1986; Sorkin 2014) in a causal set (Saravani et al. 2014; Sorkin and Yazdi 2018). The calculation in $d = 2$ is surprising since it gives rise to a volume law rather than an area law. What this means for causal set entanglement entropy is still an open question.

In Sect. 6, we describe the CST approach to quantum dynamics, which roughly follows two directions. The first, is based on "first principles", where one starts with a general set of axioms which respect microscopic covariance and causality. An important class of such theories is the set of Markovian *classical sequential growth* (CSG) models of Rideout and Sorkin (Rideout and Sorkin 2000a, 2001; Martin et al. 2001; Rideout 2001; Varadarajan and Rideout 2006), which we will describe in some detail. The dynamical framework finds a natural expression in terms of measure theory, with the classical covariant observables represented by a covariant event algebra $\mathfrak{A}$ over the sample space Ω_g of past finite causal sets (Brightwell et al. 2003; Dowker and Surya 2006). One of the main questions in CST dynamics is whether the overwhelming entropic presence of the *Kleitman–Rothschild (KR)* posets in Ω_g can be overcome by the dynamics (Kleitman and Rothschild 1975). These KR posets are highly non-manifold-like and "static", with just three "moments of time". Hence, if the continuum approximation is to emerge in the classical limit of the theory, then the entropic contribution from the KR posets should be suppressed by the dynamics in this limit. In the CSG models, the typical causal sets generated are very "tall" with countable rather than finite moments of time and, though not quite manifold-like, are very unlike the KR posets or even the subleading entropic contributions from non-manifold-like causal sets (Dhar 1978, 1980). The CSG models have generated some interest in the mathematics community, and new mathematical tools are now being used to study the asymptotic structure of the theory (Brightwell and Georgiou 2010; Brightwell and Luczak 2011, 2012, 2015).

In CST, the appropriate route to quantisation is via the quantum measure or decoherence functional defined in the double-path integral formulation (Sorkin 1994, 1995, 2007d). In the quantum versions of the CSG (*quantum sequential growth* or QSG) models the transition probabilities of CSG are replaced by the decoherence functional. While covariance can be easily imposed, a quantum version of microscopic causality is still missing (Henson 2005). Another indication of the non-triviality of quantisation comes from a prosaic generalisation of transitive percolation, which is the simplest of the CSG models. In this "complex percolation" dynamics, however, the quantum measure does not extend to the full algebra of observables which is an impediment to the construction of covariant quantum observables (Dowker et al. 2010c). This can be somewhat alleviated by taking a physically motivated approach to measure theory

(Sorkin 2011b), but the search is on to find a quantum dynamics for which the measure does extend. An important future direction is to construct covariant observables in a wider class of quantum dynamics and look for a quantum version of coupling constant renormalisation.

Whatever the ultimate quantum dynamics however, little sense can be made of the theory without a fully developed quantum interpretation for closed systems, essential to quantum gravity. Sorkin's co-event interpretation (Sorkin 2007a; Dowker and Ghazi-Tabatabai 2008) provides a promising avenue based on the quantum measure approach. While a discussion of this is outside of the scope of the present work, one can use the broader "principle of preclusion", i.e., that sets of zero quantum measure do not occur (Sorkin 2007a; Dowker and Ghazi-Tabatabai 2008), to make a limited set of predictions in complex percolation (*Sorkin and Surya, work in progress*).

The second approach to quantisation is more pragmatic, and uses the continuum inspired path integral formulation of quantum gravity for causal sets. Here, the path integral is replaced by a sum over the sample space Ω of causal sets, using the *Benincasa–Dowker* (BD) action, which goes over to the Einstein–Hilbert action (Benincasa and Dowker 2010) in the continuum limit. This can be viewed as an effective, continuum-like dynamics, arising from the more fundamental dynamics described above. A recent analytic calculation in Loomis and Carlip (2018) showed that a subdominant class of non-manifold-like causal sets, the bilayer posets, are suppressed in the path integral when using the BD action, under certain dimension dependent conditions satisfied by the parameter space. This gives hope that such an effective dynamics might be able to overcome the entropy of the non-manifold-like causal sets.

In Surya (2012), Glaser and Surya (2016), and Glaser et al. (2018), Markov Chain Monte Carlo (MCMC) methods were used for a dimensionally restricted sample space Ω_{2d} of 2-orders, which corresponds to topologically trivial $d = 2$ causal set quantum gravity. The quantum partition function over causal sets can be rendered into a statistical partition function via an analytic continuation of a "temperature" parameter, while retaining the Lorentzian character of the theory. This theory exhibits a first order phase transition (Surya 2012; Glaser et al. 2018) between a manifold-like phase and a layered, non-manifold-like one. MCMC methods have also been used to examine the sample space Ω_n of n-element causal sets and to estimate the onset of asymptotia, characterised by the dominance of the KR posets (Henson et al. 2017). These techniques have recently been extended to topologically non-trivial $d = 2$ and $d = 3$ CST (Cunningham and Surya 2019). While this approach gives us expectation values of covariant observables which allows for a straightforward interpretation, relating it to the complex or quantum partition function is non-trivial and an open problem.

In Sect. 7, we describe in brief some of the exciting phenomenology that comes out of the *kinematical* structure of causal sets. This includes the momentum space diffusion coming from CST discreteness ("swerves") (Dowker et al. 2004) and the effects of non-locality on quantum field theory (Sorkin 2007b), which includes a recent proposal for dark matter (Saravani and Afshordi 2017). Of these, the most striking is the 1987 prediction of Sorkin for the value of the cosmological constant Λ (Sorkin 1991, 1997). While the original argument was a kinematic estimate, dynamical models of fluctuating Λ were subsequently examined (Ahmed et al. 2004; Ahmed and Sorkin 2013; Zwane et al. 2018) and have been compared with recent observations (Zwane et al. 2018).

This is an exciting future direction of research in CST which interfaces intimately with observation. We conclude with a brief outlook for CST in Sect. 8.

Finally, since this is an extensive review, in order to assist the reader we have made a list of some of the key definitions, as well as the abbreviations in Appendix A.

As is true of all other approaches to quantum gravity, CST is not as yet a complete theory. Some of the challenges faced are universal to quantum gravity as a whole, while others are specific to the approach. Although we have developed a reasonably good grasp of the kinematical structure of CST and some progress has been made in the construction of effective quantum dynamics, CST still lacks a satisfactory quantum dynamics built from first principles. Progress in this direction is therefore very important for the future of the program. From a broader perspective, it is the opinion of this author that a deeper understanding of CST will help provide key insights into the nature of quantum gravity from a fully Lorentzian, causal perspective, whatever ultimate shape the final theory takes.

It is not possible for this review to be truly complete. The hope is that the interested reader will use it as a springboard to the existing literature. Several older reviews exist with differing emphasis (Sorkin 1991, 2005b; Henson 2006b, 2010; Dowker 2005; Sorkin 2009; Wallden 2013), some of which have an in depth discussion of the conceptual underpinnings of CST. The focus of the current review is to provide as cohesive an account of the program as possible, so as to be useful to a starting researcher in the field. For more technical details, the reader is urged to look at the original references.

2 A historical perspective

One of the most important conceptual realisations that arose from the special and general theories of relativity in the early part of the twentieth century, was that space and time are part of a single construct, that of spacetime. At a fundamental level, one does not exist without the other. Unlike Riemannian spaces, spacetime has a Lorentzian signature $(-,+,+,+)$ which gives rise to local lightcones and an associated global causal structure (Fig. 1). The causal structure $(M, \prec)$ of a causal spacetime[1] (M, g) is a partially ordered set or poset, with $\prec$ denoting the causal ordering on the "event-set" M.

Causal set theory (CST) as proposed in Bombelli et al. (1987), takes the Lorentzian character of spacetime and the causal structure poset in particular, as a crucial starting point to quantisation. It is inspired by a long but sporadic history of investigations into Lorentzian geometry, in which the connections between $(M, \prec)$ and the conformal geometry were eventually established. This history, while not a part of the standard narrative of General Relativity, is relevant to the sequence of ideas that led to CST. In looking for a quantum theory of spacetime, $(M, \prec)$ has also been paired with discreteness, though the earliest ideas on discreteness go back to pre-quantum and pre-relativistic physics. We now give a broad review of this history.

[1] Henceforth, we will assume that spacetime is causal, i.e., without any closed causal curves.

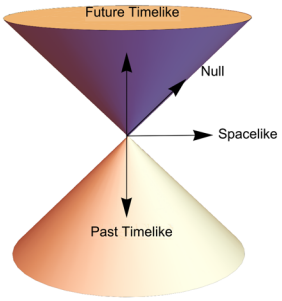

Fig. 1 The local lightcone of a Lorentzian spacetime

The first few decades after the formulation of General Relativity were dedicated to understanding the physical implications of the theory and to finding solutions to the field equations. The attitude towards Lorentzian geometry was mostly practical: it was seen as a simple, though odd, generalisation of Riemannian geometry.[2] There were however early attempts to understand this new geometry and to use causality as a starting point. Weyl and Lorentz (see Bell and Korté 2016) used light rays to attempt a reconstruction of d dimensional Minkowski spacetime $\mathbb{M}^d$, while Robb (1914, 1936) suggested an axiomatic framework for spacetime where the causal precedence on the collection of events was seen to play a critical role. It was only several decades later, however, that the mathematical structure of Lorentzian geometry began to be explored more vigorously.

In a seminal paper titled "Causality Implies the Lorentz Group", Zeeman (1964) identified the chronological poset $(\mathbb{M}^d, \ll)$ in $\mathbb{M}^d$, where $\ll$ denotes the chronological relation on the event-set $\mathbb{M}^d$. Defining a chronological automorphism[3] f_a of $\mathbb{M}^d$ as the chronological poset-preserving bijection

$$f_a : \mathbb{M}^d \to \mathbb{M}^d, \quad x \ll y \Leftrightarrow f_a(x) \ll f_a(y), \ \forall\, x, y \in \mathbb{M}^d, \tag{1}$$

Zeeman showed that the group of chronological automorphisms $\mathcal{G}_A$ is isomorphic to the group $\mathcal{G}_{Lor}$ of inhomogeneous Lorentz transformations and dilations on $\mathbb{M}^d$ when $d > 2$. While it is simple to see that the generators of $\mathcal{G}_{Lor}$ preserve the chronological structure so that $\mathcal{G}_{Lor} \subseteq \mathcal{G}_A$, the converse is not obvious. In his proof Zeeman showed that every $f_a \in \mathcal{G}_A$ maps light rays to light rays, such that parallel light rays remain parallel and moreover that the map is linear. In Minkowski spacetime every chronological automorphism is also a causal automorphism, so a Corollary to Zeeman's theorem is that the group of causal automorphisms is isomorphic to $\mathcal{G}_{Lor}$. This is a remarkable result, since it states that the physical invariants associated with $\mathbb{M}^d$ follow naturally from its *causal structure poset* $(\mathbb{M}^d, \prec)$ where $\prec$ denotes the causal relation on the event-set $\mathbb{M}^d$.

Kronheimer and Penrose (1967) subsequently generalised Zeeman's ideas to an arbitrary causal spacetime (M, g) where they identified both $(M, \prec)$ and $(M, \ll)$

[2] Hence the term "pseudo-Riemannian".

[3] Zeeman used the term "causal" instead of "chronological", but we will follow the more modern usage of these terms (Hawking and Ellis 1973; Wald 1984).

with the event-set M, devoid of the differential and topological structures associated with a spacetime. They defined an abstract causal space axiomatically, using both $(M, \prec)$ and $(M, \prec\prec)$ along with a mixed transitivity condition between the relations $\prec$ and $\prec\prec$, which mimics that in a causal spacetime.

Zeeman's result in $\mathbb{M}^d$ was then generalised to a larger class of spacetimes by Hawking et al. (1976) and Malament (1977). A chronological bijection generalises Zeeman's chronological automorphism between two spacetimes (M_1, g_1) and (M_2, g_2), and is a chronological order preserving bijection,

$$f_b : M_1 \to M_2, \quad x \prec\prec_1 y \Leftrightarrow f_b(x) \prec\prec_2 f_b(y), \ \forall\, x, y \in M_1, \tag{2}$$

where $\prec\prec_{1,2}$ refer to the chronology relations on $M_{1,2}$, respectively. The existence of a chronological bijection between two strongly causal spacetimes[4] was equated by Hawking et al. (1976) to the existence of a conformal isometry, which is a bijection $f : M_1 \to M_2$ such that f, f^{-1} are smooth (with respect to the manifold topology and differentiable structure) and $f_* g_1 = \lambda g_2$ for a real, smooth, strictly positive function λ on M_2. Malament (1977) then generalised this result to the larger class of *future and past distinguishing* spacetimes.[5] We refer to these results collectively as the Hawking–King–McCarthy–Malament theorem or *HKMM* theorem, summarised as

Theorem 1 Hawking–King–McCarthy–Malament (HKMM) *If a chronological bijection f_b exists between two d-dimensional spacetimes which are both future and past distinguishing, then these spacetimes are conformally isometric when $d > 2$.*

It was shown by Levichev (1987) that a causal bijection implies a chronological bijection and hence the above theorem can be generalised by replacing "chronological" with "causal". Subsequently Parrikar and Surya (2011) showed that the causal structure poset $(M, \prec)$ of these spacetimes also contains information about the spacetime dimension.

Thus, the causal structure poset $(M, \prec)$ of a future and past distinguishing spacetime is equivalent its conformal geometry. This means that $(M, \prec)$ is equivalent to the spacetime, except for the local volume element encoded in the conformal factor λ, which is a single scalar. As phrased by Finkelstein (1969), the causal structure in $d = 4$ is therefore (9/10) th of the metric!

En route to a theory of quantum gravity one must pause to ask: what "natural" structure of spacetime should be quantised? Is it the metric or is it the causal structure poset? The former can be defined for all signatures, but the latter is an exclusive embodiment of a causal Lorentzian spacetime. In Fig. 2, we show a 3d projection of a non-Lorentzian and non-Riemannian $d = 4$ "space-time" with signature $(-, -, +, +)$. The fact that a time-like direction can be continuously transformed into any other while still remaining time-like means that there is no order relation in the space and hence no associated

[4] A point p in a spacetime is said to be strongly causal if every neighbourhood of p contains a subneighbourhood such that no causal curve intersects it more than once. All the events in a strongly causal spacetime are strongly causal.

[5] These are spacetimes in which the chronological past and future $I^{\pm}(p)$ of each event p is unique, i.e., $I^{\pm}(p) = I^{\pm}(q) \Rightarrow p = q$.

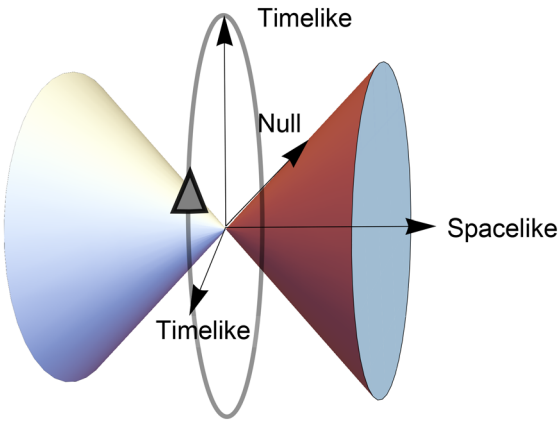

Fig. 2 An example of a signature $(-,-,+,+)$ spacetime with one spatial dimension suppressed. It is not possible to distinguish a past from a future timelike direction and hence order events, even locally

causal structure poset. We can thus view the causal structure poset as an essential embodiment of Lorentzian spacetime.

Perhaps the first explicit statement of intent to quantise the causal structure of spacetime, rather than the spacetime geometry, was by Kronheimer and Penrose (1967), who listed, as one of their motivations for axiomatising the causal structure:

> To admit structures which can be very different from a manifold. The possibility arises, for example, of a locally countable or discrete event-space equipped with causal relations macroscopically similar to those of a space-time continuum.

This brings to focus another historical thread of ideas important to CST, namely that of spacetime discreteness. The idea that the continuum is a mathematical construct which approximates an underlying physical discreteness was already present in the writings of Riemann as he ruminated on the physicality of the continuum (Riemann 1873):

> Now it seems that the empirical notions on which the metric determinations of Space are based, the concept of a solid body and that of a light ray; lose their validity in the infinitely small; it is therefore quite definitely conceivable that the metric relations of Space in the infinitely small do not conform to the hypotheses of geometry; and in fact one ought to assume this as soon as it permits a simpler way of explaining phenomena.

Many years later, in their explorations of spacetime and quantum theory, Einstein and Feynman each questioned the physicality of the continuum (Stachel 1986; Feynman 1944). These ideas were also expressed in Finkelstein's "spacetime code" (Finkelstein 1969), and most relevant to CST, in Hemion's use of *local finiteness*, to obtain discreteness in the causal structure poset (Hemion 1988). This last condition is the requirement there are only a finite number of fundamental spacetime elements in any finite volume *Alexandrov interval* $\mathbf{A}[p,q] \equiv I^+(p) \cap I^-(q)$.

Although these ideas of spacetime discreteness resonate with the appearance of discreteness in quantum theory, the latter typically manifests itself as a discrete spectrum of a continuum observable. The discreteness proposed above is different: one

is replacing the fundamental degrees of freedom, *before* quantisation, already at the kinematical level of the theory.

The most immediate motivation for discreteness however comes from the HKMM theorem itself. The missing $(1/10)$ th of the $d = 4$ metric is the volume element. A discrete causal set can supply this volume element by substituting the continuum volume with cardinality. This idea was already present in Myrheim's remarkable (unpublished) CERN preprint (Myrheim 1978), which contains many of the main ideas of CST. Here he states:

> It seems more natural to regard the metric as a statistical property of discrete spacetime. Instead we want to suggest that the concept of absolute time ordering, or causal ordering of, space-time points, events, might serve as the one and only fundamental concept of a discrete space-time geometry. In this view space-time is nothing but the causal ordering of events.

The statistical nature of the poset is a key proposal that survives into CST with the spacetime continuum emerging via a random Poisson sprinkling. We will see this explicitly in Sect. 3. Another key concept which plays a role in the dynamics is that the order relation replaces coordinate time and any evolution of spacetime takes meaning only in this intrinsic sense (Sorkin 1997).

There are of course many other motivations for spacetime discreteness. One of the expectations from a theory of quantum gravity is that the Planck scale will introduce a natural cut-off which cures both the UV divergences of quantum field theory and regulates black hole entropy. The realisation of this hope lies in the details of a given discrete theory, and CST provides us a concrete way to study this question, as we will discuss in Sect. 5.

It has been 31 years since the original CST proposal of BLMS (Bombelli et al. 1987). The early work shed considerable light on key aspects of the theory (Bombelli et al. 1987; Bombelli and Meyer 1989; Brightwell and Gregory 1991) and resulted in Sorkin's prediction of the cosmological constant Λ (Sorkin 1991). There was a seeming hiatus in the 1990s, which ended in the early 2000s with exciting results from the Rideout–Sorkin classical sequential growth models (Rideout and Sorkin 2000b, 2001; Martin et al. 2001; Rideout 2001). There have been several non-trivial results in CST in the intervening 19 odd years. In the following sections we will make a broad sketch of the theory and its key results, with this historical perspective in mind.

3 The causal set hypothesis

We begin with the definition of a causal set:

Definition A set C with an order relation $\prec$ is a *causal set* if it is

1. *Acyclic*: $x \prec y$ and $y \prec x \Rightarrow x = y, \forall x, y \in C$
2. *Transitive*: $x \prec y$ and $y \prec z \Rightarrow x \prec z, \forall x, y, z \in C$
3. *Locally finite*: $\forall x, y \in C, |\mathbf{I}[x, y]| < \infty$, where $\mathbf{I}[x, y] \equiv \text{Fut}(x) \cap \text{Past}(y)$,

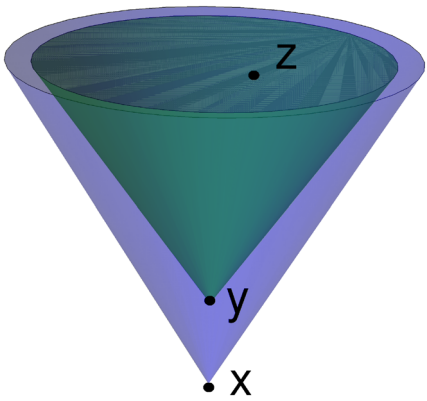

Fig. 3 The transitivity condition $x \prec y, y \prec z \Rightarrow x \prec z$ is satisfied by the causality relation $\prec$ in any Lorentzian spacetime

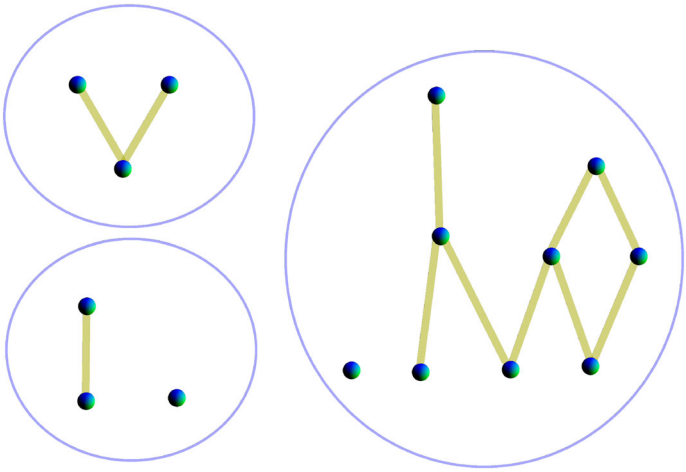

Fig. 4 The *Hasse diagrams* of some simple finite cardinality causal sets. Only the nearest neighbour relations or *links* are depicted. The remaining relations are deduced from transitivity

where $|.|$ denotes the cardinality of the set, and[6]

$$\begin{aligned} \text{Fut}(x) &\equiv \{w \in C | x \prec w, x \neq w\} \\ \text{Past}(x) &\equiv \{w \in C | w \prec x, x \neq w\}. \end{aligned} \tag{3}$$

We refer to $\mathbf{I}[x, y]$ as an *order interval*, in analogy with the Alexandrov interval in the continuum. The acyclic and transitive conditions together define a partially ordered set or poset, while the condition of local finiteness encodes discreteness (Figs. 3, 4).

The content of the HKMM theorem can be summarised in the statement:

$$\text{Causal Structure} + \text{Volume Element} = \text{Lorentzian Geometry}, \tag{4}$$

[6] These are the *exclusive* future and past sets since they do not include the element itself.

which lends itself to a discrete rendition, dubbed the "CST slogan":

$$\text{Order} + \text{Number} \sim \text{Lorentzian Geometry}. \tag{5}$$

One therefore assumes a fundamental correspondence between the number of elements in a region of the causal set and the continuum volume element that it represents. The condition of local finiteness means that all order intervals in the causal set are of finite cardinality and hence correspond in the continuum to finite volume. This CST slogan captures the essence of the (yet to be specified) continuum approximation of a manifold-like causal set, which we denote by $C \sim (M, g)$. While the continuum causal structure gives the continuum conformal geometry via the HKMM theorem, the discrete causal structure represented by the underlying causal set is conjectured to approximate the entire spacetime geometry. Thus, discreteness supplies the missing conformal factor, or the missing $(1/10)$ th of the metric, in $d = 4$.

Motivated thus, CST makes the following radical proposal (Bombelli et al. 1987):

1. Quantum gravity is a quantum theory of causal sets.
2. A continuum spacetime (M, g) is an approximation of an underlying causal set $C \sim (M, g)$, where
   (a) Order $\sim$ Causal Order
   (b) Number $\sim$ Spacetime Volume

In CST, the kinematical space of $d = 4$ continuum spacetime geometries or histories is replaced with a *sample space* Ω of causal sets. Thus, discreteness is viewed not only as a tool for regulating the continuum, but as a fundamental feature of quantum spacetime. Ω includes causal sets that have no continuum counterpart, i.e., they cannot be related via Conditions (2a) and (2b) to *any* continuum spacetime in any dimension. These non-manifold-like causal sets are expected to play an important role in the deep quantum regime. In order to make this precise we need to define what it means for a causal set to be manifold-like, i.e., to make precise the relation "$C \sim (M, g)$".

Before doing so, it is important to understand the need for a continuum approximation at all. Without it, Condition (1) yields *any* quantum theory of locally finite posets: one then has the full freedom of choosing any poset calculus to construct a quantum dynamics, without need to connect with the continuum. Examples of such poset approaches to quantum gravity include those by Finkelstein (1969) and Hemion (1988), and more recently Cortês and Smolin (2014). What distinguishes CST from these approaches is the critical role played by both causality and discrete covariance which informs the choice of the dynamics as well the physical observables. In particular, condition (2) is the requirement that in the continuum approximation these observables should correspond to appropriate continuum topological and geometric covariant observables.

What do we mean by the *continuum approximation* Condition (2)? We begin to answer this by looking for the underlying causal set of a causal spacetime (M, g). A useful analogy to keep in mind is that of a macroscopic fluid, for example a glass of water. Here, there are a multitude of molecular-level configurations corresponding to the same macroscopic state. Similarly, we expect there to be a multitude of causal sets

approximated by the same spacetime (M, g). And, just as the set of allowed microstates of the glass of water depends on the molecular size, the causal set microstate depends on the *discreteness scale* V_c, which is a fundamental spacetime volume cut-off.[7]

Since the causal set C approximating (M, g) is locally finite, it represents a proper subset of the event-set M. An *embedding* is the injective map

$$\Phi : C \hookrightarrow (M, g), \quad x \prec_C y \Leftrightarrow \Phi(x) \prec_M \Phi(y), \tag{6}$$

where $\prec_C$ and $\prec_M$ denote the order relations in C and M respectively. Not every causal set can be embedded into a given spacetime (M, g). Moreover, even if an embedding exists, this is not sufficient to ensure that $C \sim (M, g)$ since only Condition (2a) is satisfied. In addition, to correlate the cardinality of the causal set with the spacetime volume element, Condition (2b), the embeddings must also be *uniform* with respect to the spacetime volume measure of (M, g). A causal set is said to approximate a spacetime $C \sim (M, g)$ at *density* $\rho_c = V_c^{-1}$ if there exists a *faithful embedding*

$$\Phi : C \hookrightarrow M, \quad \Phi(C) \text{ is a uniform distribution in } (M, g) \text{ at density } \rho_c, \tag{7}$$

where by *uniform* we mean with respect to the spacetime volume measure of (M, g).

The uniform distribution at density ρ_c ensures that every finite spacetime volume V is represented by a finite number of elements $n \sim \rho_c V$ in the causal set. It is natural to make these finite spacetime regions causally convex, so that they can be constructed from unions of Alexandrov intervals $\mathbf{A}[p, q]$ in (M, g). However, we must ensure covariance, since the goal is to be able to recover the approximate covariant spacetime geometry. This is why $\Phi(C)$ is required to be uniformly distributed in (M, g) with respect to the spacetime volume measure. It is obvious that a "regular" lattice cannot do the job since it is not regular in all frames or coordinate systems. Hence it is not possible to consistently assign $n \sim \rho_c V$ to such lattices (see Fig. 5).

The issue of symmetry breaking is of course obvious even in Euclidean space. Any regular discretisation breaks the rotational and translational symmetry of the continuum. In the lattice calculations for QCD, these symmetries are restored only in the continuum limit, but are broken as long as the discreteness persists. In Christ et al. (1982) it was suggested that symmetry can be restored in a randomly generated lattice where there lattice points are uniformly distributed via a Poisson process. This has the advantage of not picking any preferred direction and hence not explicitly breaking symmetry, at least on average. We will discuss this point in greater detail further on.

Set in the context of spacetime, the Poisson distribution is a natural choice for $\Phi(C)$, with the probability of finding n elements in a spacetime region of volume v given by

$$P_v(n) = \frac{(\rho_c v)^n}{n!} \exp^{-\rho_c v}. \tag{8}$$

[7] The most obvious choice for V_c is the Planck volume, but we will not require it at this stage.

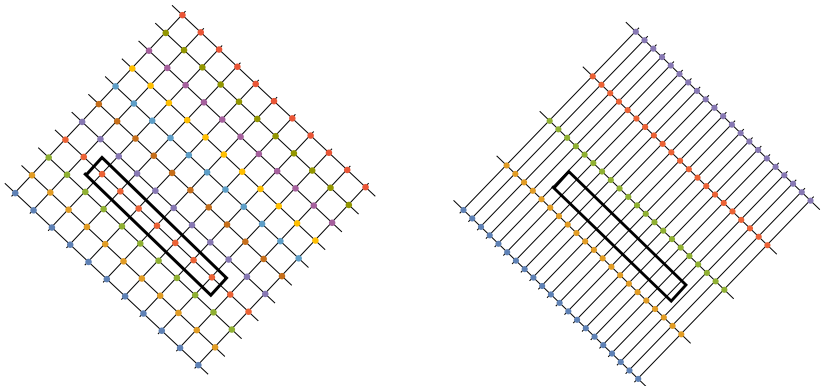

Fig. 5 The lightcone lattice in $d = 2$. The lattice on the left looks "regular" in a fixed frame but transforms into the "stretched" lattice on the right under a boost. The $n \sim \rho_c V$ correspondence cannot be implemented as seen from the example of the Alexandrov interval, which contains $n = 7$ lattice points in the lattice in the left but is empty after a boost

This means that on the *average*

$$\langle \mathbf{n} \rangle = \rho_c v, \tag{9}$$

where $\mathbf{n}$ is the random variable associated with the random causal set $\Phi(C)$. This distribution then gives us the covariantly defined $n \sim \rho_c V$ correspondence we seek.[8]

In a *Poisson sprinkling* into a spacetime (M, g) at density ρ_c one selects points in (M, g) uniformly at random and imposes a partial ordering on these elements via the induced spacetime causality relation. Starting from (M, g), we can then obtain an ensemble of "microstates" or causal sets, which we denote by $\mathcal{C}(M, \rho_c)$, via the Poisson sprinkling.[9] Each causal set thus obtained is a *realisation*, while any averaging is done over the whole ensemble.

We say that a causal set C is approximated by a spacetime (M, g) if C can be obtained from (M, g) via a high probability Poisson sprinkling. Conversely, for every $C \in \mathcal{C}(M, \rho_c)$ there is a natural embedding map

$$\Phi : C \hookrightarrow M\,, \tag{10}$$

where $\Phi(C)$ is a particular realisation in $\mathcal{C}(M, \rho_c)$. In Fig. 6, we show a causal set obtained by Poisson sprinkling into $d = 2$ de Sitter spacetime.

That there is a fundamental discrete randomness even *kinematically* is not always easy for a newcomer to CST to come to terms with. Not only does CST posit a fundamental discreteness, it also requires it to be probabilistic. Thus, even before

[8] Since $\Phi(C)$ is a random causal set, any function of $\mathbf{F} : C \to \mathbb{R}$ is therefore a random variable.

[9] $\mathcal{C}(M, \rho_c)$ explicitly depends on the spacetime metric g, which we have suppressed for brevity of notation.

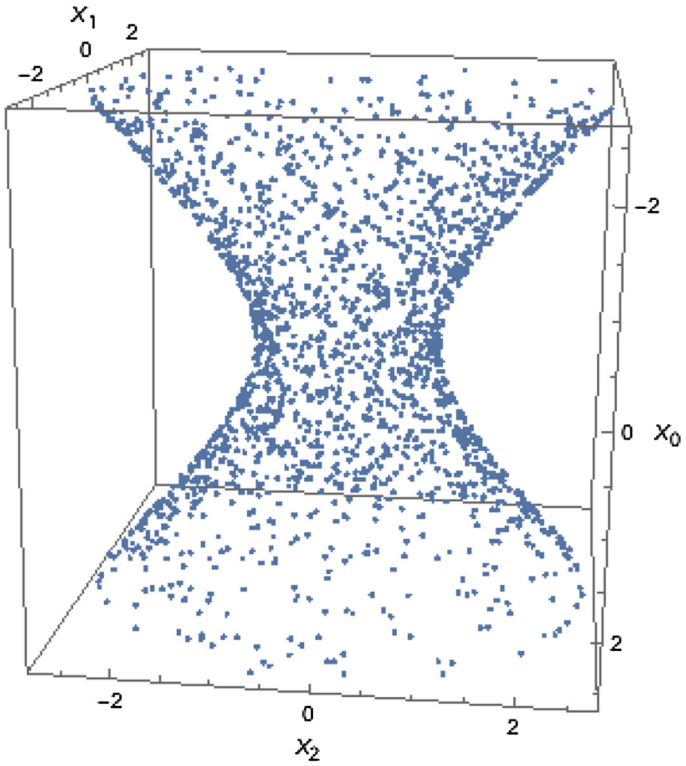

Fig. 6 A Poisson sprinkling into a portion of 2d de Sitter spacetime embedded in $\mathbb{M}^3$. The relations on the elements are deduced from the causal structure of $\mathbb{M}^3$

coming to quantum probabilities, CST makes us work with a classical, stochastic discrete geometry.

Let us state some obvious, but important aspects of Eq. (8). Let $\Phi : C \hookrightarrow (M, g)$ be a faithful embedding at density ρ_c. While the set of all finite volume regions[10] v possess on average $\langle \mathbf{n} \rangle = \rho_c v$ elements of C,[11] the Poisson fluctuations are given by $\delta n = \sqrt{n}$. Thus, it is possible that the region contains no elements at all, i.e., there is a "void". An important question to ask is how large a void is allowed, since a sufficiently large void would have an obvious effect on our macroscopic perception of a continuum. If spacetime is unbounded, as it is in Minkowski spacetime, the probability for the existence of a void of *any* size is one. Can this be compatible at all with the idea of an emergent continuum in which the classical world can exist, unperturbed by the vagaries of quantum gravity?

The presence of a macroscopic void means that the continuum approximation is not realised in this region. A prediction of CST is then that the emergent continuum regions of spacetime are bounded both spatially and temporally, even if the underlying causal set is itself "unbounded" or countable. Thus, a continuum universe is not viable

[10] We assume that these are always causally convex.

[11] Henceforth we will identify $\Phi(C)$ with C, whenever Φ is a faithful embedding.

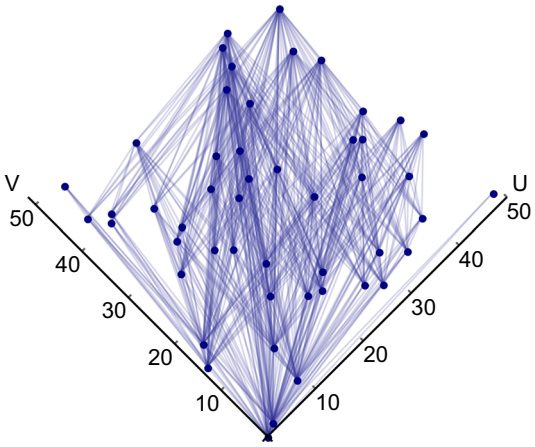

Fig. 7 A *Hasse diagram* of a causal set that faithfully embeds into a causal diamond in $\mathbb{M}^2$. In a Hasse diagram only the nearest neighbour relations or *links* are shown. The remaining relations follow by transitivity

forever. However, since the current phase of the observable universe *does* have a continuum realisation one has to ask whether this is compatible with CST discretisation. In Dowker et al. (2004) the probability for there to be at least one nuclear size void $\sim 10^{-60}m^4$ was calculated in a region of Minkowski spacetime which is the size of our present universe. Using general considerations they found that the probability is of order $10^{84} \times 10^{168} \times e^{-10^{72}}$, which is an absurdly small number! Thus, CST poses no phenomenological inconsistency in this regard.

An example of a manifold-like causal set C which is obtained via a Poisson sprinkling into a 2d causal diamond is shown in Fig. 7. A striking feature of the resulting graph is that there is a high degree of connectivity. In the *Hasse diagram* of Fig. 7 only the nearest neighbour relations or links are depicted with the remaining relations following from transitivity. $e \prec e' \in C$ is said to be a *link* if $\nexists\, e'' \in C$ such that $e'' \neq e, e'$ and $e \prec e'' \prec e'$. In a causal set that is obtained from a Poisson sprinkling, the *valency*, i.e., the number of nearest neighbours or links from any given element is typically very large. This is an important feature of continuum like causal sets and results from the fact that the elements of C are uniformly distributed in (M, g). For a given element $e \in C$, the probability of an event $x \succ e$ to be a link is equal to the probability that the Alexandrov interval $\mathbf{A}[e, x]$ does not contain any elements of C. Since

$$P_V(0) = e^{-\rho_c V}, \tag{11}$$

the probability is significant only when $V \sim V_c$. As shown in Fig. 8, in $\mathbb{M}^d$, the set of events within a proper time $\propto (V)^{1/d}$ to the future (or past) of a point p lies in the region between the future light cone and the hyperboloid $-t^2 + \Sigma_i x_i^2 \propto (V)^{2/d}$, with $t > 0$. Up to fluctuations, therefore, most of the future links to e lie within the hyperboloid with $V = V_c \pm \sqrt{V_c}$. This is a non-compact, infinite volume region and hence the number of future links to e is (almost surely) infinite. Since linked elements are the nearest neighbours of e, this means the valency of the graph C is infinite. It is this feature of manifold-like causal sets which gives rise to a characteristic "non-

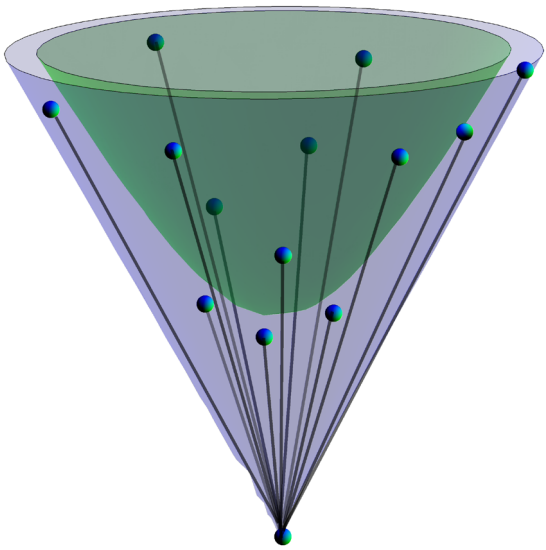

Fig. 8 The *valency* or number of nearest neighbours of an element in a causal set obtained from a Poisson sprinkling into $\mathbb{M}^2$ is infinite

locality", and plays a critical role in the continuum approximation of CST, time and again.

The Poisson distribution is not the only choice for a uniform distribution. A pertinent question is whether a different choice of distribution is possible, which would lead to a different manifestation of the continuum approximation. In Saravani and Aslanbeigi (2014), this question was addressed in some detail. We summarise this discussion. Let $C \sim (M, g)$ at density ρ_c. Consider k non-overlapping Alexandrov intervals of volume V in (M, g). Since C is uniformly distributed, $\langle \mathbf{n} \rangle = \rho_c V$. The most optimal choice of distribution, is also one in which the fluctuations $\delta\mathbf{n}/\langle \mathbf{n} \rangle = \sqrt{\langle (\mathbf{n} - \langle \mathbf{n} \rangle)^2 \rangle}/\langle \mathbf{n} \rangle$ are minimised. This ensures that C is as close to the continuum as possible. For the Poisson distribution $\delta\mathbf{n}/\langle \mathbf{n} \rangle = 1/\sqrt{\langle \mathbf{n} \rangle} = 1/\sqrt{\rho_c V}$. Is this as good as it gets? It was shown by Saravani and Aslanbeigi (2014) that for $d > 2$, and under certain further technical assumptions, the Poisson distribution indeed does the best job. Strengthening these results is important as it can improve our understanding of the continuum approximation.

3.1 The Hauptvermutung or fundamental conjecture of CST

An important question is the uniqueness of the continuum approximation associated to a causal set C. Can a given C be faithfully embedded at density ρ_c into two different spacetimes, (M, g) and (M', g')? We expect that this is the case if (M, g) and (M', g') differ on scales smaller than ρ_c, or that they are, in an appropriate sense, "close" $(M, g) \sim (M', g')$. Let us assume that a causal set *can* be identified with two macroscopically distinct spacetimes at the same density ρ_c. Should this be interpreted as a hidden duality between these spacetimes, as is the case for example for isospectral manifolds or mirror manifolds in string theory (Greene and Plesser 1991)? The

answer is clearly in the negative, since the aim of the CST continuum approximation is to ensure that C contains *all* the information in (M, g) at scales above ρ_c^{-1}. Macroscopic non-uniqueness would therefore mean that the intent of the CST continuum approximation is *not* satisfied.

We thus state the fundamental conjecture of CST:

The Hauptvermutung of CST: *C can be faithfully embedded at density ρ_c into two distinct spacetimes, (M, g) and (M', g') iff they are approximately isometric.*

By an *approximate isometry* , $(M, g) \sim (M', g')$ at density ρ_c, we mean that (M, g) and (M', g') differ only at scales smaller than ρ_c. Defining such an isometry rigorously is challenging, but concrete proposals have been made by Bombelli (2000), Noldus (2002, 2004), Bombelli and Noldus (2004) and Bombelli et al. (2012), en route to a full proof of the conjecture. Because of the technical nature of these results, we will discuss it only very briefly in the next section, and instead use the above intuitive and functional definition of closeness.

Condition (1) tells us that the kinematic space of Lorentzian geometries must be replaced by a *sample space* Ω of causal sets. Let Ω be the set of all countable causal sets and $\mathcal{H}$ the set of all possible Lorentzian geometries, in all dimensions. If $\sim$ denotes the approximate isometry at a given ρ_c, as discussed above, the quotient space $\mathcal{H}/\sim$ corresponds to the set of all continuum-like causal sets $\Omega_{\text{cont}} \subset \Omega$ at that ρ_c. Thus, causal sets in Ω correspond to Lorentzian geometries of *all* dimensions! Couched this way, we see that CST dynamics has the daunting task of not only obtaining manifold-like causal sets in the classical limit, but also ones that have dimension $d = 4$.

As mentioned in the introduction, the sample space of n element causal sets Ω_n is dominated by the KR posets depicted in Fig. 9 and are hence very non-manifold-like (Kleitman and Rothschild 1975). A KR poset has three "layers" (or abstract "moments of time"), with roughly $n/4$ elements in the bottom and top layer and such that each element in the bottom layer is related to roughly half those in the middle layer, and similarly each element in the top layer is related to roughly half those in the middle layer.

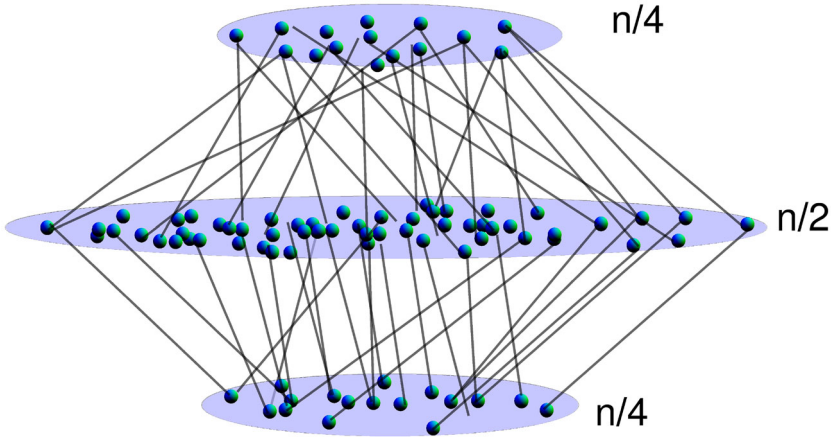

Fig. 9 A Kleitman–Rothschild or KR poset

The number of KR posets grows as $\sim 2^{n^2/4}$ and hence must play a role in the deep quantum regime. Since they are non-manifold-like they pose a challenge to the dynamics, which must overcome their entropic dominance in the classical limit of the theory. Even if the entropy from these KR posets is suppressed by an appropriate choice of dynamics, however, there is a sub-dominant hierarchy of non-manifold-like posets (also layered) which also need to be reckoned with (Dhar 1978, 1980; Promel et al. 2001).

Closely tied to the continuum approximation is the notion of "coarse graining". Given a spacetime (M, g) the set $\mathcal{C}(M, \rho_c)$ can be obtained for different values of ρ_c. Given a causal set C which faithfully embeds into (M, g) at ρ_c, one can then *coarse grain* it to a smaller subcausal set $C' \subset C$ which faithfully embeds into (M, g) at $\rho_c' < \rho_c$. A natural coarse graining would be via a random selection of elements in C such that for every n elements of C roughly $n' = (\rho_c'/\rho_c)n$ elements are chosen. Even if C itself does not faithfully embed into (M, g) at ρ_c, it is possible that a coarse graining of C can be faithfully embedded. This would be in keeping with our sense in CST that the deep quantum regime need not be manifold-like. One can also envisage manifold-like causal sets with a regular fixed lattice-like structure attached to each element similar to a "fibration", in the spirit of Kaluza–Klein theories. Instead of the coarse graining procedure, it would be more appropriate to take the quotient with respect to this fibre to obtain the continuum like causal set. Recently, the implications of coarse graining in CST, both dynamically and kinematically, were considered in Eichhorn (2018) based on renormalisation techniques.

3.2 Discreteness without Lorentz breaking

It is often assumed that a fundamental discreteness is incompatible with continuous symmetries. As was pointed out in Christ et al. (1982), in the Euclidean context, symmetry can be preserved on average in a random lattice. In Bombelli et al. (2009), it was shown that a causal set in $\mathcal{C}(\mathbb{M}^d, \rho_c)$ not only preserves Lorentz invariance on average, but in *every* realisation, with respect to the Poisson distribution. Thus, in a very specific sense a manifold-like causal set does not break Lorentz invariance. In order to see the contrast between the Lorentzian and Euclidean cases we present the arguments of Bombelli et al. (2009) starting with the easier Euclidean case.

Consider the Euclidean plane $\mathcal{P} = (\mathbb{R}^2, \delta_{ab})$, and let $\Phi : \mathcal{C}(\mathcal{P}, \rho_c) \hookrightarrow \mathcal{P}$ be the natural embedding map, where $\mathcal{C}(\mathcal{P}, \rho_c)$ denotes the ensemble of Poisson sprinklings into $\mathcal{P}$ at density ρ_c. A rotation $r \in SO(2)$ about a point $p \in \mathcal{P}$, induces a map $r^* : \mathcal{C}(\mathcal{P}, \rho_c) \to \mathcal{C}(\mathcal{P}, \rho_c)$, where $r^* = \Phi^{-1} \circ r \circ \Phi$ and similarly a translation t in $\mathcal{P}$ induces the map $t^* : \mathcal{C}(\mathcal{P}, \rho_c) \to \mathcal{C}(\mathcal{P}, \rho_c)$. The action of the Euclidean group is clearly not transitive on $\mathcal{C}(\mathcal{P}, \rho_c)$ but has non-trivial orbits which provide a fibration of $\mathcal{C}(\mathcal{P}, \rho_c)$. Thus the ensemble $\mathcal{C}(\mathcal{P}, \rho_c)$ preserves the Euclidean group on average. This is the sense in which the discussion of Christ et al. (1982) states that the random discretisation preserves the Euclidean group.

The situation is however different for a *given* realisation $P \in \mathcal{C}(\mathcal{P}, \rho_c)$. Fixing an element $e \in \Phi(P)$, we define a direction $\mathbf{d} \in S^1$, the space of unit vectors in $\mathcal{P}$ centred at e. Under a rotation r about e, $\mathbf{d} \to r^*(\mathbf{d}) \in S^1$. In general, we want a rule that assigns a natural direction to every $P \in \mathcal{C}(\mathcal{P}, \rho_c)$. One simple choice is to

find the closest element to e in $\Phi(P)$, which *is* well defined in this Euclidean context. Moreover, this element is almost surely unique, since the probability of two elements being at the same radius from e is zero in a Poisson distribution. Thus we can define a "direction map" $\mathbf{D}_e : \mathcal{C}(\mathcal{P}, \rho_c) \to S^1$ for a fixed $e \in \Phi(P)$ consistent with the rotation map, i.e., $\mathbf{D}_e$ commutes with any $r \in SO(2)$, or is *equivariant*.

Associated with $\mathcal{C}(\mathcal{P}, \rho_c)$, is a probability distribution μ arising from the Poisson sprinkling which associates with every measurable set α in $\mathcal{C}(\mathcal{P}, \rho_c)$ a probability $\mu(\alpha) \in [0, 1]$. The Poisson distribution being volume preserving (Stoyan et al. 1995), the measure on $\mathcal{C}(\mathcal{P}, \rho_c)$ moreover must be independent of the action of the Euclidean group on $\mathcal{C}(\mathcal{P}, \rho_c)$, i.e.: $\mu \circ r = \mu$.

In analogy with a continuous map, a measurable map is one whose preimage from a measurable set is itself a measurable set. The natural map $\mathbf{D}$ we have defined is a measurable map, and we can use it to define a measure on S^1: $\mu_{\mathbf{D}} \equiv \mu \circ \mathbf{D}^{-1}$. Using the invariance of μ under rotations and the equivariance of $\mathbf{D}$ under rotations

$$\mu_{\mathbf{D}} = \mu \circ r \circ \mathbf{D}^{-1} = \mu \circ \mathbf{D}^{-1} \circ r = \mu_{\mathbf{D}} \circ r \quad \forall\, r \in SO(2), \tag{12}$$

we see that $\mu_{\mathbf{D}}$ is also invariant under rotations. Because S^1 is compact, this does not lead to a contradiction. In analogy with the construction used in Bombelli et al. (2009) for the Lorentzian case, we choose a measurable set $s \equiv (0, 2\pi/n) \in S^1$. A rotation by $r(2\pi/n)$, takes $s \to s'$ which is non-overlapping, so that after n successive rotations, $r^n(2\pi/n) \circ s = s$. Since each rotation does not change $\mu_{\mathbf{D}}$ and $\mu_{\mathbf{D}}(S^1) = 1$, this means that $\mu_{\mathbf{D}}(s) = 1/n$. Thus, it *is* possible to assign a consistent direction for a given realisation $P \in \mathcal{C}(\mathcal{P}, \rho_c)$ and hence break Euclidean symmetry.

However, this is not the case for the space of sprinklings $\mathcal{C}(\mathbb{M}^d, \rho_c)$ into $\mathbb{M}^d$, where the hyperboloid $\mathcal{H}^{d-1}$ now denotes the space of future directed unit vectors and is invariant under the Lorentz group $SO(n-1, 1)$ about a fixed point $p \in \mathbb{M}^{d-1}$ (see Fig. 10). To begin with, there is no "natural" direction map. Let $C \in \mathcal{C}(\mathbb{M}^d, \rho_c)$. To find an element which is closest to some fixed $e \in \Phi(C)$, one has to take the infimum over $J^+(e)$, or some suitable Lorentz invariant subset of it, which being non-compact, does not exist. Assume that *some* measurable direction map $D : \Omega_{\mathbb{M}^d} \to \mathcal{H}^{d-1}$, does exist. Then the above arguments imply that μ_D must be invariant under Lorentz boosts. The action of successive Lorentz transformations Λ can take a given measurable set

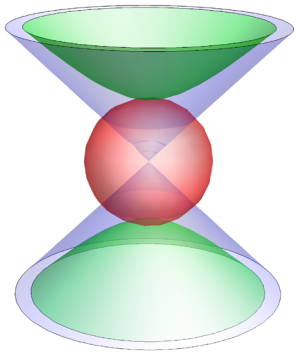

Fig. 10 The space of unit directions in $\mathbb{R}^d$ is S^{d-1}, while the space of unit timelike vectors in $\mathbb{M}^d$ is $\mathcal{H}^{d-1}$

$h \in \mathcal{H}^{d-1}$ to an infinite number of copies that are non-overlapping, and *of the same measure*. Since $\mathcal{H}^{d-1}$ is non-compact, this is not possible unless each set is of measure zero, but since this is true for any measurable set h and we require $\mu_D(\mathcal{H}^{d-1}) = 1$, this is a contradiction. This proves the following theorem (Bombelli et al. 2009):

Theorem 2 *In dimensions $n > 1$ there exists no equivariant measurable map $\mathbf{D} : \mathcal{C}(\mathbb{M}^d, \rho_c) \to \mathcal{H}$, i.e.,*

$$\mathbf{D} \circ \Lambda = \Lambda \circ \mathbf{D} \ \forall \ \Lambda \in SO(n-1, 1). \tag{13}$$

In other words, even for a given sprinkling $\omega \in \Omega_{\mathbb{M}^d}$ it is not possible to consistently pick a direction in $\mathcal{H}^{d-1}$. Consistency means that under a boost $\Lambda : \omega \to \Lambda \circ w$, and hence $D(\omega) \to \Lambda \circ D(\omega) \in \mathcal{H}^{d-1}$. Crucial to this argument is the use of the Poisson distribution.[12] Thus, an important prediction of CST is local Lorentz invariance. Tests of Lorentz invariance over the last couple of decades have produced an ever-tightening bound, which is therefore consistent with CST (Liberati and Mattingly 2016).

3.3 Forks in the road: what makes CST so "different"?

In many ways CST does not fit the standard paradigms adopted by other approaches to quantum gravity and it is worthwhile trying to understand the source of this difference. The program is minimalist but also rigidly constrained by its continuum approximation. The ensuing non-locality means that the apparatus of local physics is not readily available to CST.

Sorkin (1991) describes the route to quantum gravity and the various forks at which one has to make choices. Different routes may lead to the same destination: for example (barring interpretational questions), simple quantum systems can be described equally well by the path integral and the canonical approach. However, this need not be the case in gravity: a set of consistent choices may lead you down a unique path, unreachable from another route. Starting from broad principles, Sorkin argued that certain choices at a fork are preferable to others for a theory quantum gravity. These include the choice of Lorentzian over Euclidean, the path integral over canonical quantisation and discreteness over the continuum. This set of choices leads to a CST-like theory, while choosing the Lorentzian–Hamiltonian-continuum route leads to a canonical approach like Loop Quantum Gravity.

Starting with CST as the final destination, we can work backward to retrace our steps to see what forks had to be taken and why other routes are impossible to take. The choice at the discreteness versus continuum fork and the Lorentzian versus Euclidean fork are obvious from our earlier discussions. As we explain below, the other essential fork that *has* to be taken in CST is the histories approach to quantisation.

One of the standard routes to quantisation is via the canonical approach. Starting with the phase space of a classical system, with or without constraints, quantisation rules give rise to the familiar apparatus of Hilbert spaces and self adjoint operators. In

[12] It is interesting to ask if other choices of uniform distribution satisfy the above theorem. If so, then our criterion for a uniform distribution could not only include ones that minimise the fluctuations but also those that respect Lorentz invariance.

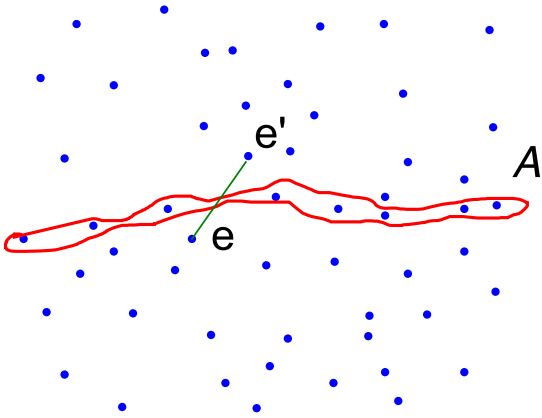

Fig. 11 A "missing link" from e to e' which "bypasses" the inextendible antichain $\mathcal{A}$

quantum gravity, apart from interpretational issues, this route has difficult technical hurdles, some of which have been partially overcome (Ashtekar and Pullin 2017). Essential to the canonical formulation is the $3+1$ split of a spacetime $M = \Sigma \times \mathbb{R}$, where Σ is a Cauchy hypersurface, on which are defined the canonical phase space variables which capture the intrinsic and extrinsic geometry of Σ.

The continuum approximation of CST however, does not allow a meaningful definition of a Cauchy hypersurface, because of the " graphical non-locality" inherent in a continuum like causal set, as we will now show. We begin by defining an *antichain* to be a set of unrelated elements in C, and an *inextendible antichain* to be an antichain $\mathcal{A} \subset C$ such that every element $e \in C \backslash \mathcal{A}$ is related to an element of $\mathcal{A}$. The natural choice for a discrete analog of a Cauchy hypersurface is therefore an inextendible antichain $\mathcal{A}$, which separates the *set* C into its future and past, so that we can express $C = \text{Fut}(\mathcal{A}) \sqcup \text{Past}(\mathcal{A}) \sqcup \mathcal{A}$, with $\sqcup$ denoting disjoint union. However, an element in $\text{Past}(\mathcal{A})$ can be related via a link to an element in $\text{Fut}(\mathcal{A})$ thus "bypassing" $\mathcal{A}$. An example of a "missing link" is depicted in Fig. 11. This means that unlike a Cauchy hypersurface, $\mathcal{A}$ is not a summary of its past, and hence a canonical decomposition using Cauchy hypersurfaces is not viable (Major et al. 2006). On the other hand, each causal set is a "history", and since the sample space of causal sets is countable, one can construct a path integral or path-sum as over causal sets. We will describe the dynamics of causal sets in more detail in Sect. 6.

Before moving on, we comment on the condition of local finiteness which, as we have pointed out, provides an intrinsic definition of spacetime discreteness, which does not need a continuum approximation. An alternative definition would be for the causal set to be countable, which along with the continuum approximation is sufficient to ensure the number to volume correspondence. This includes causal sets with order intervals of infinite cardinality, and allows us to extend causal set discretisation to more general spacetimes, like anti de Sitter spacetimes, where there exist events p, q in the spacetime for which $\text{vol}(\mathbf{A}[p, q])$ is not finite. However, what is ultimately of interest is the dynamics, and in particular, the sample space Ω of causal sets. In the growth models we will encounter in Sects. 6.1, 6.2 and 6.3 the sample space consists of past finite posets, while in the continuum-inspired dynamics of Sect. 6.4 it consists

of finite element posets. Thus, while countable posets may be relevant to a broader framework in which to study the dynamics of causal sets, it suffices for the present to focus on locally finite posets.

4 Kinematics or geometric reconstruction

In this section we discuss the program of geometric reconstruction in which topological and geometric invariants of a continuum spacetime (M, g) are "reconstructed" from the underlying ensemble of causal sets. The assumption that such a reconstruction exists for any covariant observable in (M, g) comes from the Hauptvermutung of CST discussed in Sect. 3.

In the statement of the Hauptvermutung, we used the phrase "approximately isometric", with the promise of an explanation in this section. A rigorous definition requires the notion of closeness of two Lorentzian spacetimes. In Riemannian geometry, one has the Gromov–Hausdorff distance (Petersen 2006), but there is no simple extension to Lorentzian geometry, in part because of the indefinite signature. In Bombelli and Meyer (1989) a measure of closeness of two Lorentzian manifolds was given in terms of a pseudo distance function, which however is neither symmetric nor satisfies the triangle inequality. Subsequently, in a series of papers, a true distance function was defined on the space of Lorentzian geometries, dubbed the Lorentzian Gromov–Hausdorff distance (Bombelli 2000; Noldus 2002, 2004; Bombelli and Noldus 2004; Bombelli et al. 2012). While this makes the statement of the Hauptvermutung precise, there is as yet no complete proof. Recently, a purely order theoretic criterion has been used to determine the closeness of causal sets and prove a version of the Hauptvermutung (*Sorkin and Zwane, work in progress*).

Apart from these more formal constructions, as we will describe below, a large body of evidence has accumulated in favour of the Hauptvermutung. In the program of geometric reconstruction, we look for order invariants in continuum like causal sets which correspond to manifold (either topological or geometric) invariants of the spacetime. These *manifold invariants* include dimension, spatial topology, distance functions between fixed elements in the spacetime, scalar curvature, the discrete Einstein–Hilbert action, the Gibbons–Hawking–York boundary terms, Green functions for scalar fields, and the d'Alembertian operator for scalar fields. The identification of the order invariant $\mathcal{O}$ with the manifold invariant $\mathcal{G}$ then ensures that a causal set C that faithfully embeds into (M, g) cannot faithfully embed into a spacetime with a different manifold invariant $\mathcal{G}'$.[13] Thus, in this sense two manifolds can be defined to be close with respect to their specific manifold invariants. We can then state the limited, order-invariant version of the Hauptvermutung:

$\mathcal{O}$-Hauptvermutung: *If C faithfully embeds into (M, g) and (M', g') then (M, g) and (M', g') have the same manifold invariant $\mathcal{G}$ associated with $\mathcal{O}$.*

The longer our list of correspondences between order invariants and manifold invariants, the closer we are to proving the full Hauptvermutung.

[13] This is in the sense of an ensemble, since the faithful embedding is defined statistically.

In order to correlate a manifold invariant $\mathcal{G}$ with an order invariant $\mathcal{O}$, we must recast geometry in purely order theoretic terms. Note that since locally finite posets appear in a wide range of contexts, the poset literature contains several order invariants, but these are typically not related to the manifold invariants of interest to us. The challenge is to choose the appropriate invariants that correspond to manifold invariants. Guessing and verifying this using both analytic and numerical tools is the art of geometric reconstruction.

A *labelling* of a causal set C is an injective map: $C \to \mathbb{N}$, which is the analogue of a choice of coordinate system in the continuum. By an *order invariant* in a finite causal set C we mean a function $\mathcal{O} : C \to \mathbb{R}$ such that $\mathcal{O}$ is independent of the labelling of C. For a manifold-like causal set[14] $C \in \mathcal{C}(M, \rho_c)$, associated to every order invariant $\mathcal{O}$ is the random variable $\mathbf{O}$ whose expectation value $\langle \mathbf{O} \rangle$ in the ensemble $\mathcal{C}(M, \rho_c)$ is either equal to or limits (in the large ρ_c limit) to a manifold invariant $\mathcal{G}$ of (M, g). We will typically restrict to compact regions of (M, g) in order to deal with finite values of $\mathbf{O}$.

The first candidates for geometric order invariants were defined for $\mathcal{C}(\mathbf{A}[p, q], \rho_c)$ where $\mathbf{A}[p, q]$ is an Alexandrov interval in $\mathbb{M}^d$. Some of these have been later generalised to Alexandrov intervals (or causal diamonds) in *Riemann Normal Neighbourhoods (RNN)* in curved spacetime. These manifold invariants are in this sense "local". In order to find *spatial* global invariants, the relevant spacetime region is a *Gaussian Normal Neighbourhood (GNN)* of a compact Cauchy hypersurface in a globally hyperbolic spacetime. As discussed in Sect. 3 compactness is necessary for manifold-likeness since otherwise there is a finite probability for there to be arbitrarily large voids which negates the discrete-continuum correspondence.

Before proceeding, we remind the reader that we are restricting ourselves to manifold-like causal sets in this section only because of the focus on CST kinematics and the continuum approximation. All the order invariants, however, can be calculated for *any* causal set, manifold-like or not. These order invariants give us an important class of covariant observables, essential to constructing a quantum theory of causal sets. As we will see in Sect. 6 they play an important role in the quantum dynamics.

The analytic results in this section are typically found in the continuum limit, $\rho_c \to \infty$. Strictly speaking, this limit is unphysical in CST because of the assumption of a fundamental discreteness. There are fluctuations at finite ρ_c which give important deviations from the continuum with potential phenomenological consequences. These are however not always easy to calculate analytically and hence require simulations to assess the size of fluctuations at finite ρ_c. As we will see below, CST kinematics therefore needs a combination of analytical and numerical tools.

4.1 Spacetime dimension estimators

The earliest result in CST is a dimension estimator for Minkowski spacetime due to Myrheim (1978)[15] and predates BLMS (Bombelli et al. 1987). A closely related

[14] We remind the reader that the ensemble depends on the spacetime (M, g) but we suppress the dependence on g for the sake of brevity.

[15] This remarkable preprint also contains the first expression, again without detailed proof, of the volume of a small causal diamond in an arbitrary spacetime.

dimension estimator was given by Meyer (1988), which is now collectively known as the *Myrheim–Meyer* dimension estimator.

The number of relations R in a finite n element causal set C is the number of ordered pairs $e_i, e_j \in C$ such that $e_i \prec e_j$. Since the maximum number of possible relations on n elements is $\binom{n}{2}$, the ordering fraction is defined as

$$r = \frac{2R}{n(n-1)}. \tag{14}$$

It was shown by Myrheim (1978) that r depends only on the dimension when C faithfully embeds into $\mathbb{M}^d$.

We now describe the construction of a closely related dimension estimator by Meyer (1988). Consider an Alexandrov interval $\mathbf{A}_d[p,q] \subset \mathbb{M}^d$ of volume $V >> \rho_c^{-1}$. We are interested in calculating the expectation value of the random variable $\mathbf{R}$ associated with R for the ensemble $\mathcal{C}(\mathbf{A}_d, \rho_c)$. This is the probability that a pair of elements $e_1, e_2 \in \mathbf{A}_d[p,q]$ are related. Given e_1, the probability of there being an e_2 in its future is given by the volume of the region $J^+(e_1) \cap J^-(p)$ in units of the discreteness scale, while the probability to pick e_1 is given by the volume of $\mathbf{A}_d[p,q]$. This joint probability can be calculated as follows.

Without loss of generality, choose $p = (-T/2, 0, \ldots, 0)$ and $q = (T/2, 0, \ldots, 0)$, so that the total volume

$$V = \zeta_d T^d, \quad \zeta_d \equiv \frac{\mathcal{V}_{d-2}}{2^{d-1} d(d-1)} \tag{15}$$

with $\mathcal{V}_{d-2}$ the volume of the unit $d-2$ sphere. For this choice,

$$\langle \mathbf{R} \rangle = \rho_c^2 \int_{\mathbf{A}_d} dx_1 \int_{J^+(x_1)\cap J^-(q)} dx_2 = \rho_c^2 \, \zeta_d \int_{\mathbf{A}_d} dx_1 T_1^d \,, \tag{16}$$

where T_1 is the proper time from x_1 to q, and $\mathbf{A}_d \equiv \mathbf{A}_d[p,q]$. Evaluating the integral, one finds

$$\langle \mathbf{R} \rangle = \rho_c^2 V^2 \frac{\Gamma(d+1)\Gamma(\frac{d}{2})}{4\Gamma(\frac{3d}{2})}. \tag{17}$$

Using $\langle \mathbf{n} \rangle = \rho_c V$, Meyer (1988) obtained a dimension estimator from $\langle \mathbf{R} \rangle$ by noting that the ratio

$$\frac{\langle \mathbf{R} \rangle}{\langle \mathbf{n} \rangle^2} = \frac{\Gamma(d+1)\Gamma(\frac{d}{2})}{4\Gamma(\frac{3d}{2})} \equiv f_0(d) \tag{18}$$

is a function only of d. In the large n limit this is is half of Myrheim's ordering fraction r.

However, the fluctuations in $\mathbf{R}$ are large and hence the right dimension cannot be obtained from a single realisation $C \in \mathcal{C}(\mathbf{A}_d, \rho_c)$, but rather by averaging over

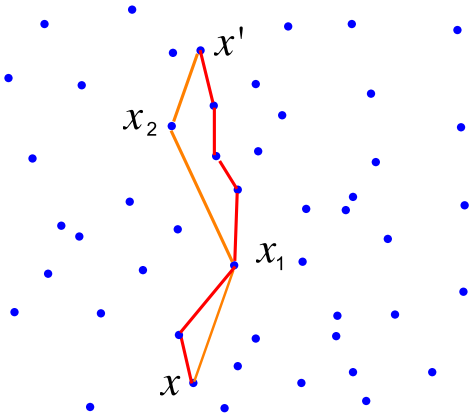

Fig. 12 Two different chains between x and x'. One is a $k = 4$ chain and the other is a $k = 7$ chain

the ensemble. For large enough ρ_c, however, the relative fluctuations should become smaller, and allow one to distinguish causal sets obtained from sprinkling into different dimensional Alexandrov intervals. Such systematic tests have been carried out numerically using sprinklings into different spacetimes by Reid (2003) and show a general convergence as ρ_c is taken to be large, or equivalently the interval size is taken to be large.

How can we use this dimension estimator in practice? Let C be a causal set of sufficiently large cardinality n. If the dimension obtained from Eq. (18) is approximately an integer d, this means that C cannot be distinguished from a causal set that belongs to $\mathcal{C}(\mathbf{A}_d, \rho_c)$ using just the dimension estimator, for $n \sim \rho_c \text{vol}(\mathbf{A}_d)$. We denote this by $C \sim_d \mathbf{A}_d$. This also means that C cannot be a typical member of $\mathcal{C}(\mathbf{A}_{d'}, 1)$ for dimension $d' \neq d$, so that $C \nsim_{d'} \mathbf{A}_{d'}$. The equivalence $C \sim_d \mathbf{A}_d$ itself does not of course imply that $C \sim \mathbf{A}_d$ or even that C is manifold-like. Rather, it is the limited statement that its dimension estimator is the same as that of a typical causal set in $\mathcal{C}(\mathbf{A}_d, \rho_c)$ for $n \sim \rho_c \text{vol}(\mathbf{A}_d)$.

This is our first example of a $\mathcal{O}$-Hauptvermutung, where the order invariant $\mathcal{O}$ is the ordering fraction r and the spacetime dimension d is the corresponding manifold invariant $\mathcal{G}$. This example provides a useful template in the search for manifold-like order invariants some of which we will describe in the next few subsections.

Using simulations Abajian and Carlip (2018) recently obtained the Myrheim–Meyer dimension as function of interval size for nested intervals in a causal set in $\mathcal{C}(\mathbf{A}_d, \rho_c)$ for $d = 3, 4, 5$. As the interval size decreases, they found that the resulting causal sets are likely to be disconnected due to the large fluctuations at small volumes. In the extreme case, there is a single point with no relations and hence the Myrheim–Meyer dimension goes to ∞ rather than 0. Using a criterion to discard such disconnected regions, it was shown that this dimension estimator gives a value of 2 at small volumes, even when $d = 3, 4, 5$, in support of the dimensional reduction conjecture in quantum gravity (Carlip 2017) which we discuss briefly in Sect. 5.

Meyer's construction is in fact more general and yields a whole family of dimension estimators. If we think of the relation $e_1 \prec e_2$ as a *chain* c_2 of two elements, then a *k-chain* c_k is the causal sequence $e_1 \prec e_2 \ldots \prec e_{k-1} \prec e_k$ (see Fig. 12), where

the *length* of c_k is defined as $k-2$. We denote the *abundance*, or number of the c_k's contained in C, by C_k. Its expectation value in $\mathcal{C}(\mathbf{A}_d[p,q],\rho_c)$ is therefore given by a sequence of k nested integrals over a sequence of nested Alexandrov intervals, $\mathbf{A}_d[p,q] \supset I(x_1,q) \supset I(x_2,q) \ldots I(x_k,q)$ which, as was shown by Meyer (1988), can be calculated inductively to give

$$\langle \mathbf{C}_k \rangle = \rho_c^k \chi_k V^k, \qquad \chi_k \equiv \frac{1}{k}\left(\frac{\Gamma(d+1)}{2}\right)^{k-1} \frac{\Gamma(\frac{d}{2})\Gamma(d)}{\Gamma(\frac{kd}{2})\Gamma(\frac{(k+1)d}{2})}. \tag{19}$$

Thus for any k, k', the ratio of $\langle \mathbf{C}_k \rangle^{1/k}$ to $\langle \mathbf{C}_{k'} \rangle^{1/k'}$ only depends on the dimension. This gives a multitude of dimension estimators.

Meyer's calculation of $\langle \mathbf{C}_k \rangle$ was generalised to a small causal diamond $\mathbf{A}_d[p,q]$ that lies in an RNN of a general spacetime, i.e., one for which $RT^2 << 1$, where T is the proper time from p to q and R denotes components of the curvature at the centre of the diamond (Roy et al. 2013). In such a region the dimension satisfies the more complicated equation

$$\begin{aligned} f_0^2(d)\Bigg(-\frac{1}{3}\frac{(d+2)}{(3d+2)} - \frac{(4d+2)}{(2d+2)}\left(\frac{\langle \mathbf{C}_3 \rangle}{\chi_3}\right)^{\frac{4}{3}} \frac{1}{\langle \mathbf{C}_1 \rangle^4} \\ + \frac{1}{3}\frac{(4d+2)(5d+2)}{(2d+2)(3d+2)}\frac{\langle \mathbf{C}_4 \rangle}{\chi_4}\frac{1}{\langle \mathbf{C}_1 \rangle^4}\Bigg) = -\frac{\langle \mathbf{C}_2 \rangle^2}{\langle \mathbf{C}_1 \rangle^4}, \end{aligned} \tag{20}$$

where $f_0(d)$ is given by Eq. (18). It is straightforward to show that the expression above reduces to the Myrheim–Meyer dimension estimator in $\mathbb{M}^d$. The calculation of Roy et al. (2013) uses a result of Khetrapal and Surya (2013), which makes explicit earlier calculations of the volume of a causal diamond in an RNN (Myrheim 1978; Gibbons and Solodukhin 2007). The C_k themselves are order invariants and hence are covariant observables for finite element causal sets.

This class of dimension estimators is just one among several that have appeared in the literature, including the mid-point scaling estimator (Bombelli 1987; Reid 2003), and more recent ones (Glaser and Surya 2013; Aghili et al. 2019). We refer the reader to the literature for more details.

4.2 Topological invariants

The next step in our reconstruction is that of topology. There are several poset topologies described in the literature (see Stanley 2011 as well as Surya 2008 for a review). However, our interest is in finding one that most closely resembles the "coarse" continuum topology. It is clear that the full manifold topology cannot be reproduced in a causal set since it requires arbitrarily small open sets. However, according to the Hauptvermutung, topological invariants like the homology groups and the fundamental groups of (M, g) should be encoded in the causal set.

A natural choice for a topology in C based on the order relation is one generated by the order intervals $\mathbf{I}[e_i, e_j] \equiv \mathrm{Fut}(e_i) \cap \mathrm{Past}(e_j)$. Indeed, in the continuum the

topology generated by their analogs, the Alexandrov intervals, can be shown to be equivalent to the manifold topology in strongly causal spacetimes (Penrose 1972). However, even for a causal set approximated by a finite region of $\mathbb{M}^d$, this order-interval topology is roughly discrete or trivial. This is because the intersection of any two intervals in the continuum can be of order the discreteness scale and hence contain just a single element of the causal set, thus trivialising the topology. A way forward is to use the causal structure to obtain a locally finite open covering of C and construct the associated "nerve simplicial complex" (see Munkres 1984).

In Major et al. (2007, 2009), a "spatial" homology of C was obtained in this manner by considering an inextendible antichain $\mathcal{A} \subset C$ (see Sect. 3.3), which is an (imperfect) analog of a Cauchy hypersurface. The natural topology on $\mathcal{A}$ is the discrete topology since there are no causal relations amongst the elements. In order to provide a topology on $\mathcal{A}$, one needs to "borrow" information from a neighbourhood of $\mathcal{A}$. The method devised was to consider elements to the future of $\mathcal{A}$ and "thicken" by a parameter v to some collar neighbourhood $\mathcal{T}_v(\mathcal{A}) \equiv \{e| \, |\mathrm{IFut}(\mathcal{A}) \cap \mathrm{IPast}(e)| \leq v\}$. Here IFut and IPast denote the *inclusive future and past* respectively, where for any $S \subset C$, $\mathrm{IFut}(S) = \mathrm{Fut}(S) \cup S$ and $\mathrm{IPast}(S) = \mathrm{Past}(S) \cup S$.

A topology can then be induced on $\mathcal{A}$ from $\mathcal{T}_v(\mathcal{A})$ by considering the open cover $\{\mathcal{O}_v \equiv \mathrm{Past}(e) \cap \mathcal{A}\}$ of $\mathcal{A}$, for $e \in \mathcal{M}_v(\mathcal{A})$, the set of future most elements of $\mathcal{T}_v(\mathcal{A})$. The "nerve" simplicial complex $\mathcal{N}_v(\mathcal{A})$ can be constructed from $\{\mathcal{O}_v\}$ for every v. For a spacetime (M, g) with compact Cauchy hypersurface Σ, and for $C \in \mathcal{C}(M, \rho_c)$ it was shown in Major et al. (2007, 2009) that there exists a range of values of v such that $\mathcal{N}_v(\mathcal{A})$ is homological to Σ (up to the discreteness scale) as long as there is a sufficient separation between the discreteness scale $\ell_c \equiv V_c^{1/d}$ and ℓ_K the scale of extrinsic curvature of Σ.

One might also imagine a similar construction on C using the nerve simplicial complex of causal intervals of a given minimal cardinality v which cover C. However, in the continuum the intersection of such intervals may not only be of order the discreteness scale, but also such that they "straddle" each other. As an example consider the equal volume intervals $\mathbf{A}[p_1, q_1]$, $\mathbf{A}[p_2, q_2]$ in $\mathbb{M}^2$ where p_1, q_1 are at $x = 0$ in a frame (x, t), with the x-coordinate of p_2 being < 0 and that of q_2 being > 0. These two intervals not only intersect, but straddle each other, i.e., the set difference $\mathbf{A}[p_1, q_1] \backslash \mathbf{A}[p_2, q_2]$ is disconnected as is $\mathbf{A}[p_2, q_2] \backslash \mathbf{A}[p_1, q_1]$. By choosing p_2, q_2 appropriately, the intersection region can be made very "thin", pushing most of the volume of $\mathbf{A}[p_2, q_2]$ out of $\mathbf{A}[p_1, q_1]$. Thus, while they intersect in $\mathbb{M}^2$ these intervals would not intersect in the corresponding causal set C. This results in a non-trivial cycle in the associated nerve simplicial complex for C, which is absent in the continuum. Such a construction can be therefore made to work only in a sufficiently localised region within C.

An example of a localised subset of C is the region sandwiched between two inextendible and non-overlapping antichains $\mathcal{A}_1$ and $\mathcal{A}_2$. The resulting homology constructed from the nerve simplicial complex of the order intervals of volume $\sim v$ is then is associated with a spacetime region rather than just space, and hence includes topology change. While preliminary investigations along these lines have been started, there is much that remains to be understood. Another possibility for characterising the

spatial homology uses chain complexes but this has only been partially investigated. A further open direction is to obtain the causal set analogues of other topological invariants.

4.3 Geodesic distance: timelike, spacelike and spatial

In Minkowski spacetime, the proper time between two events is the *longest* path between them; the shortest path between two time-like separated events is of course any zig-zag null path, which has zero length. In a causal set C, if $e_i \prec e_f$, one can construct different chains from e_i to e_f, of varying lengths. A natural choice for the discrete timelike geodesic distance between e_i and e_f is the length of the longest chain, which we denote by $l(e_i, e_j)$, as was suggested by Myrheim (1978). It was shown in Brightwell and Gregory (1991) that the expectation value of the associated random variable $\mathbf{l}$ in the ensemble $\mathcal{C}(\mathbf{A}_d, \rho_c)$ limits to a dimension dependent constant

$$\lim_{\rho_c \to \infty} \frac{\langle \mathbf{l}(x, x') \rangle}{(\rho_c V(x, x'))^{1/d}} = m_d \tag{21}$$

where

$$1.77 \leq \frac{2^{1-\frac{1}{d}}}{\Gamma(1+\frac{1}{d})} \leq m_d \leq \frac{2^{1-\frac{1}{d}} e \left(\Gamma(1+d)\right)^{\frac{1}{d}}}{d} \leq 2.62. \tag{22}$$

For a finite ρ_c, the fluctuations in $l(e_i, e_j)$ are very large (Meyer 1988; Bachmat 2007) and hence the correspondence becomes meaningful only when averaged over a large ensemble.

In Roy et al. (2013), an expression for the proper time T of a small causal diamond $\mathbf{A}_d$ in an RNN of a d dimensional spacetime was obtained to leading order correction in terms of the random variables $\mathbf{C}_k$ associated to the abundance of k-chains,

$$T^{3d} = \frac{1}{2d^2 \rho_c^3} \Big(J_1 - 2J_2 + J_3 \Big). \tag{23}$$

where

$$J_k \equiv (kd+2)((k+1)d+2) \frac{1}{\zeta_d^3} \left(\frac{\langle \mathbf{C}_k \rangle}{\chi_k} \right)^{3/k}, \tag{24}$$

with $\langle \mathbf{C}_k \rangle$ the ensemble average in $\mathcal{C}(\mathbf{A}_d, \rho_c)$ and ζ_d, χ_k defined in Eqs. (15) and (19). This definition is not intrinsic to a single causal set but requires the full ensemble. Nevertheless, it is of interest to study the intrinsic version of the expression by replacing $\langle \mathbf{C}_k \rangle$ by C_k for each causal set and then taking the ensemble average to check for convergence. Recent simulations suggest that these expressions converge fairly rapidly to their continuum values.

Spacelike distance is far less straightforward to compute from the poset, because events that are spacelike to each other have no natural relationship to each other. We

saw this already in trying to find a topology on the inextendible antichain. Thus, the relationship must be "borrowed" from the elements in the causal past and future of the spacelike events. Brightwell and Gregory (1991) defined the following, naive spatial distance function in $\mathbb{M}^d$. For a given spacelike pair $p, q \in \mathbb{M}^d$, the common future and past are defined as $J^+(p,q) \equiv J^+(p) \cap J^+(q)$ and $J^-(p,q) \equiv J^-(p) \cap J^-(q)$ respectively. For every $r \in J^+(p,q)$ and $s \in J^-(p,q)$ let $\tau(s,r)$ be the timelike distance. Then the naive distance function is given by

$$d_s(p,q) \equiv \min_{r,s} \tau(r,s). \tag{25}$$

While this is a perfectly good continuum definition of the distance in $\mathbb{M}^d$, it fails for the causal set when $d > 2$ since the number of pairs (r,s) which minimise $\tau(r,s)$ lies in the region between a co-dimension 2 hyperboloid and the light cone $\tau = 0$. In the causal set we can use the length of the maximal chain $\mathbf{l}(r,s)$ to obtain $\tau(r,s)$, but in $d > 2$ since there are an infinite number of proper time minimising pairs (r,s), there will almost surely be those for which $\mathbf{l}(r,s)$ is drastically underestimated. The minimisation in Eq. (25) will then always give 2 as the spatial distance!

Rideout and Wallden (2009) generalised the naive distance function using minimising pairs (r,s) such that either r or s is linked to both p and q. Instead of minimising over these pairs (again infinite), the *2-link* distance can be calculated by averaging over the pairs. Numerical simulations for the naive distance and the 2-link distance for sprinklings into a finite region of $\mathbb{M}^3$ show that the latter stabilises as a function of ρ_c. The former underestimates the spatial distance compared to the continuum, and the latter overestimates it. The spatial distance functions of both Brightwell and Gregory (1991) and Rideout and Wallden (2009) are however strictly "predistance" functions since they do not satisfy the triangle inequality.

Recently, a one-parameter family of discrete induced spatial distance functions was proposed for an inextendible antichain in a causal set by Eichhorn et al. (2019a). To begin with, a one parameter family of continuum induced distance functions d_ϵ was constructed for a globally hyperbolic region (M,g) of spacetime with Cauchy hypersurface Σ using only the causal structure and the volume element. In $\mathbb{M}^d$ with Σ a constant time slice in an inertial reference frame, the volume of a past causal cone from $p \succ \Sigma$ has a simple relation to the diameter D of the base of the cone $J^-(p) \cap \Sigma$

$$\mathrm{vol}(J^-(p) \cap J^+(\Sigma)) = \zeta_d \left(\frac{D}{2}\right)^d. \tag{26}$$

Since D is the distance between any two antipodal points on the $S^{d-2} \subset \Sigma$, this simple formula defines the induced distance on Σ. In a general spacetime this formula can be used to extract an approximate induced distance function in a sufficiently small region of Σ. In order to define the distance function on all of Σ, a meso-scale ϵ must be introduced, and the full distance function can then be obtained by minimising over all segmented paths, such that each segment is bounded from above by ϵ. For $\epsilon << \ell_K$, the scale of the extrinsic curvature of Σ, d_ϵ was shown in Eichhorn et al. (2019a) to converge to the induced spatial distance function d_h on (Σ, h).

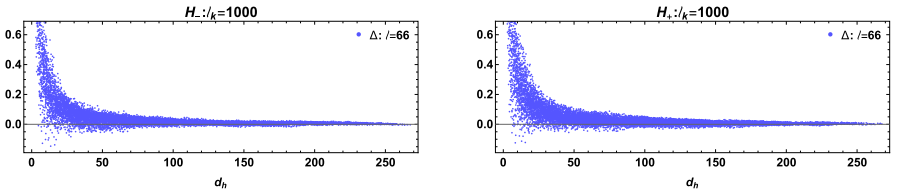

Fig. 13 The error in the discrete spatial distance is plotted as a function of the continuum induced distance on Σ for causal sets in $\mathbb{M}^2$ for Σ of both constant negative and constant positive extrinsic curvature. The discrete spatial distance always overestimates the continuum distance around the discreteness scale giving rise to "discrete asymptotic silence". For larger distances, when there is a good separation of scales, the discrete distance gives a good approximation to the continuum induced distance

Since the d_ϵ are constructed from the causal structure and volume element they are readily defined on an inextendible antichain on a causal set. For causal sets in $\mathcal{C}(M, \rho_c)$ with $\Sigma \subset M$ the discrete distance function d_ϵ was shown to significantly overestimate the continuum induced distance on Σ when the latter is close to the discreteness scale $(V_c)^{1/d}$. This discrete *"asymptotic silence"* of Eichhorn et al. (2017) mimics the narrowing of light cones in the UV, and can be traced to the large fluctuations expected around the discreteness scale. At larger distances, on the other hand, d_ϵ is a good approximation of the continuum induced distance when $(V_c)^{1/d} << \epsilon << \ell_K$. It was shown moreover that the continuum induced distance is slightly underestimated for positive curvature and slightly overestimated for negative curvature, when restricted to small regions of Σ. This was confirmed by extensive numerical simulations in $\mathbb{M}^d$ for $d = 2, 3$ (see Fig. 13). This works paves the way to recovering more spatial geometric invariants from the causal set, and is currently in progress (Eichhorn, Surya and Versteegen).

4.4 The d'Alembertian for a scalar field

One of the very first questions that comes to mind in the continuum approximation of CST is whether a tangent space can be defined naturally on a causal set. To answer this (unfortunately in the negative), we need to examine the non-local nature of a manifold-like causal set in more detail. The nearest neighbours of an element e are those that it is linked to, both in its future and its past. In a causal set approximated by Minkowski spacetime for example, and as discussed in Sect. 3, every element has an infinite number of nearest neighbours (see Fig. 8). Similarly, the "next nearest" neighbours to e are those for which the interval $|\mathbf{I}[e, e']| = 1$ or $|\mathbf{I}[e', e]| = 1$.[16] Thus, in keeping with the covariance of the causal set, we say that if $e \prec e'$ and $|\mathbf{I}[e, e']| = k$ (or $e' \prec e$ and $|\mathbf{I}[e', e]| = k$), then e' is the k-nearest neighbour of e. Examples of past k-nearest neighbours of an element in a Minkowski-like causal set are shown in Fig. 14.

It is already clear from the picture that emerges in $\mathbb{M}^d$ that, unlike a regular lattice, a simple construction of a locally defined tangent space from the set of links or next to nearest neighbours to e is not possible, since the valency of the graph is infinite.

[16] Note that this is the *exclusive* interval and hence there exists exactly one element e'' such that $e \prec e'' \prec e'$.

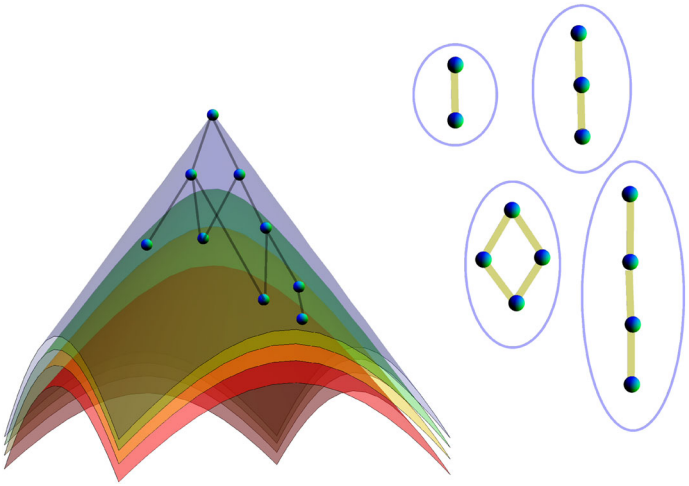

Fig. 14 The layered structure of neighbourhoods. The nearest neighbours are the links or zero intervals, the next to nearest neighbours are the 1-element intervals, the etc. Here we depict the types of 0, 1, 2 element intervals. In the figure two examples of 3 element intervals are also shown

This means in particular that derivative operators cannot also be simply defined. How then can we look for the effect of discreteness on the propagation of fields? We will discuss this in more detail in Sect. 5 but for now we notice that the best way forward is to look for scalar quantities, rather than more general tensorial ones, in making the discrete-continuum correspondence.

A scalar field d'Alembertian is a good first step. In Sorkin (2007b) and Henson (2010), a proposal was given for a discrete d'Alembertian of a free scalar field on a causal set approximated by $\mathbb{M}^2$, and extended in Benincasa and Dowker (2010), Benincasa (2013) and Dowker and Glaser (2013) to higher dimensions. For a real scalar field on a causal set $\phi : C \to \mathbb{R}$ define the $d = 4$ dimensionless discrete operator

$$B\phi(e) \equiv \frac{4}{\sqrt{6}}\left[-\phi(e) + \left(\sum_{e'\in L_0(e)} - 9\sum_{e'\in L_1(e)} + 16\sum_{e'\in L_2(e)} - 8\sum_{e'\in L_3(e)}\right)\phi(e')\right], \quad (27)$$

where $L_k(e)$ denotes the set of k-nearest neighbours to the past of $e \in C$. This is a highly non-local operator since it depends on the number of all the (possibly infinite) nearest $k = 0, 1, 2, 3$ neighbours. Notice the alternating sum whose precise coefficients turn out to be very important to the continuum limit. The expectation value of the random variable $\mathbf{B}\phi(x)$ associated with $\mathcal{C}(\mathbb{M}^4, \rho_c)$ at $x \in \mathbb{M}^4$ is

$$\frac{1}{\sqrt{\rho_c}}\langle \mathbf{B}\phi(x)\rangle = \frac{4\sqrt{\rho_c}}{\sqrt{6}}\left[-\phi(x) + \rho_c\int_{y\in J^-(x)} d^4y\,\phi(y)\,e^{-\rho_c v}\left(1 - 9\rho_c v + 8(\rho_c v)^2 - \frac{4}{3}(\rho_c v)^3\right)\right], \quad (28)$$

where $v \equiv \mathrm{vol}(\mathbf{A}(y,x))$ and we have used the probability $P_n(v)$ for v to contain n elements, Eq. (8). We have also made the expression dimensionful, in order to be able to make a direct comparison with the continuum. Let us consider the past of x in $\mathbb{M}^2$ and choose a frame $\mathcal{F}_\phi$ such that $\phi(y)$ varies slowly in the immediate past of x with respect to $\mathcal{F}_\phi$. As was shown in $\mathbb{M}^2$ by Sorkin (2007b) and in $\mathbb{M}^4$ by Benincasa and Dowker (2010) (see also Benincasa 2013), for ϕ of compact support there are miraculous cancellations that make the contributions far down the light cone negligible, thus making the operator effectively local.

In order to evaluate this integral, we first note that since ϕ is of compact support, the region of integration is compact. In $\mathcal{F}_\phi$, a small $|y-x|$ expansion of $\phi(y)$ around $\phi(x)$ can be done. Following Sorkin (2007b), Benincasa and Dowker (2010) and Benincasa (2013), the non-compact region of integration $J^-(x)$ can be split into 3 non-overlapping regions, $\mathcal{W}_1, \mathcal{W}_2, \mathcal{W}_3$ in $\mathcal{F}_\phi$: $\mathcal{W}_1$ is a neighbourhood of x, $\mathcal{W}_2$ is a neighbourhood of $\partial J^-(x)$ but bounded away from the origin and $\mathcal{W}_3$ is bounded away from $\partial J^-(x)$. The integral over $\mathcal{W}_3$ was shown in Benincasa (2013) to be bounded from above by an integral that tends to zero faster than any power of ρ_c^{-1}, while the integral over $\mathcal{W}_2$ was shown to go to zero faster than $\rho_c^{-3/2}$. The local contribution from $\mathcal{W}_1$ dominates so that

$$\lim_{\rho_c \to \infty} \frac{1}{\sqrt{\rho_c}} \langle \mathbf{B}\phi(x) \rangle = \Box\phi(x). \tag{29}$$

Thus, $B(\phi)$ is "effectively local" since its dominant contribution comes from $\mathcal{W}_1$ which is a local neighbourhood of x defined by the frame $\mathcal{F}_\phi$. In *this* frame, the contribution to $\mathbf{B}\phi(x)$ is dominated by the restrictions of L_k to $\mathbf{A}(p,q) \cap J^-(x)$. Thus, while $\mathbf{B}\phi(x)$ is not determined just by the value of ϕ at x, it depends on ϕ only in an appropriately defined compact neighbourhood of x, rather than all of $J^-(x)$. This "restoration of locality" is an important subtlety in CST kinematics.

How does a scalar field on a causal set evolve under this non-local d'Alembertian? There are indications that while the evolution in $d = 2$ is stable, it is unstable in $d = 4$ as suggested by Aslanbeigi et al. (2014). Hence it is desirable to look for generalisations of the B_k operator. An infinite family of non-local d'Alembertians has been constructed by Aslanbeigi et al. (2014) and shown to give the right continuum limit. It is still an open question whether there is a subfamily of these operators that lead to a stable evolution.

An interesting direction that has been explored by Yazdi and Kempf (2017) is to use the spectral information of the d'Alembertian operator to obtain all the information about the causal set. This was explored for $\mathbf{A}_2[p,q] \subset \mathbb{M}^2$ and it was shown that the spectrum of the d'Alembertian (or Feynman propagator) gives the link matrix (see Eq. (56) below), i.e., the matrix of all linked pairs using which the entire causal set can be reconstructed via transitivity. Extending these results to higher dimensions is an interesting open question.

4.5 The Ricci scalar and the Benincasa–Dowker action

Next we describe a very important development in CST: the construction of the discrete Einstein–Hilbert action or the *Benincasa–Dowker (BD) action* for a causal set (Benincasa and Dowker 2010; Dowker and Glaser 2013). The approach of Benincasa and Dowker (2010) was to generalise $\mathbf{B}\phi(x)$ to an RNN in curved spacetime in $d = 2, d = 4$. Again, the region of integration can be split into three parts as was done for flat spacetime. The contribution from $\mathcal{W}_3$ i.e., away from a neighbourhood of $\partial J^-(x)$ can again be shown to be bounded from above by an integral that tends to zero faster than any power of ρ_c^{-1}. In the limit, the contribution from the near region $\mathcal{W}_1$ contained in an RNN is such that

$$\lim_{\rho_c \to \infty} \frac{1}{\sqrt{\rho_c}} \langle \mathbf{B}\phi(x) \rangle |_{\mathcal{W}_1} = \Box\phi(x) - \frac{1}{2}\mathcal{R}(x)\phi(x). \tag{30}$$

where $\mathcal{R}(x)$ is the Ricci scalar (Benincasa and Dowker 2010; Benincasa 2013). However, the calculation in region $\mathcal{W}_2$ which is in the neighbourhood of $\partial J^-(x)$ but bounded away from the origin, is non-trivial, and needs a further set of assumptions to show that it does not contribute in the $\rho_c \to \infty$ limit. A painstaking calculation in Belenchia et al. (2016a) using Fermi Normal Coordinates shows that this is indeed the case in an approximately flat region of a four dimensional spacetime. Generalising this calculation to arbitrary spacetimes is highly non-trivial but is an important open question in CST.

What is of course exciting about this form for the d'Alembertian Eq. 27 is that it can be used to find the discrete Ricci curvature and hence the action. Assuming that

$$\lim_{\rho_c \to \infty} \frac{1}{\sqrt{\rho_c}} \langle \mathbf{B}\phi(x) \rangle |_{\mathcal{W}_2} = 0 \tag{31}$$

holds in all spacetimes, and putting[17] $\phi(x) = 1$

$$\lim_{\rho_c \to \infty} \frac{1}{\sqrt{\rho_c}} \langle \mathbf{B}\phi(x) \rangle = -\frac{1}{2}\mathcal{R}(x). \tag{32}$$

Thus we can write the dimensionless discrete Ricci curvature at an element $e \in C$ (Benincasa 2013) as

$$R(e) = \frac{4}{\sqrt{6}}\Big[1 - N_0(e) + 9N_1(e) - 16N_2(e) + 8N_3(e)\Big], \tag{33}$$

[17] By doing so, we violate the condition that ϕ is of compact support. However, given that the regions $\mathcal{W}_3$ and by assumption $\mathcal{W}_2$ contribute negligibly, we can always ensure this by only requiring constancy of ϕ in a neighbourhood of $\mathcal{W}_1$.

where $N_k(e) \equiv |L_k(e)|$. Summing over the n elements of a finite element causal set gives the dimensionless discrete action

$$S^{(4)}(C) = \sum_{e \in C} R(e) = \frac{4}{\sqrt{6}}\Big[n - N_0 + 9N_1 - 16N_2 + 8N_3\Big], \tag{34}$$

where N_k is the total number of k-element order intervals in C.

Benincasa and Dowker (2010) (see also Benincasa 2013) showed that (under the assumption Eq. (31)) the random variable $\mathbf{S}^{(4)}$ associated with $\mathcal{C}(M, g)$ gives the Einstein–Hilbert action in the continuum limit

$$\lim_{\rho_c \to \infty} \hbar \frac{\ell_c^2}{\ell_p^2} \langle \mathbf{S}^{(4)}(C) \rangle = \mathcal{S}_{\mathrm{EH}}(g), \tag{35}$$

up to (as yet unknown) boundary terms.

Equation (35) is exactly true in an approximately flat region of a four dimensional spacetime as shown in Belenchia et al. (2016a). Proving Eq. (31) in general is however non-trivial since there are caustics in a generic spacetime which complicate the calculation. On the other hand, numerical simulations suggest that again, up to boundary terms, the Benincasa–Dowker action S *is* the Einstein–Hilbert action (Benincasa 2013; Cunningham 2018b). We will discuss these boundary terms below.

Before doing so, we note that crucial to the validity of the causal set action are its fluctuations in a given causal set. These were shown in Sorkin (2007b) to be large for the operator B in $\mathbb{M}^2$. This can be traced to the fact that the elements in $L_k(e)$ for $k = 0, 1, 2, 3$ are very close to the discreteness scale and hence the d'Alembertian is susceptible to large Poisson fluctuations at small volumes. In order to "shield" the continuum from these fluctuations, a new *mesoscale* $\ell_\kappa > \ell_c$ and its associated density ρ_κ was introduced in Sorkin (2007b). Thus instead of a single discrete operator B, we have a one parameter family of operators:

$$B_\kappa \phi(e) \equiv \frac{4}{\sqrt{6}}\Big[-\phi(e) + \epsilon \sum_{e' \prec e} f(n(e', e), \epsilon)\phi(e')\Big], \tag{36}$$

where $\epsilon \equiv \rho_\kappa / \rho_c$ is a *non-locality parameter*,[18] $n(e, e') = |I(e, e')|$ and

$$f(n, \epsilon) = (1 - \epsilon)^n \left[1 - \frac{9\epsilon n}{1 - \epsilon} + \frac{8\epsilon^2 n(n-1)}{(1-\epsilon)^2} - \frac{4\epsilon^3 n(n-1)(n-2)}{3(1-\epsilon)^3}\right]. \tag{37}$$

This function "smears out" the contributions of the N_k into four "layers" which appear with alternating sign, as shown in Fig. 15. Each layer is thus thickened from a single value of k to a range of k values depending on ϵ. When this mesoscale matches the discreteness scale, i.e., $\epsilon = 1$, each layer collapses to a single value of k. This then gives us a one-parameter family of actions $S_\kappa(C, \epsilon)$, where ϵ can be viewed as a

[18] ϵ is a new free parameter in the theory, whose value should ultimately be decided by the fundamental dynamics.

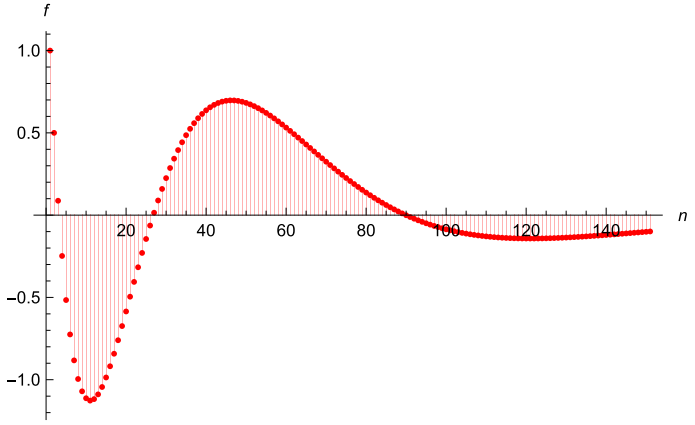

Fig. 15 The function $f(n, 0.05)$. There are 4 regions of alternating sign corresponding to 4 "smeared out" layers

tunable coupling constant. As we will see in Sect. 6, this gives rise to an interesting phase structure in 2d CST.

The result for $d = 2, 4$ are due to Benincasa and Dowker (2010) and Benincasa (2013) and were generalised to arbitrary dimensions by Dowker and Glaser (2013) and Glaser (2014), using a dimension dependent smearing function $f_d(n, \epsilon)$.

There have been other attempts to obtain the action of a causal set. In Sverdlov and Bombelli (2009), the curvature at the centre of an Alexandrov interval $\mathbf{A}_d[p, q]$ in a RNN was obtained using the leading order corrections to the volume of a small causal diamond (Gibbons and Solodukhin 2007)

$$V = V_0 \left(1 - \frac{d}{24(d+1)(d+2)} R(0) T^2 + \frac{d}{24(d+1)} R_{00} T^2\right), \tag{38}$$

where T is the proper time from p to q and V_0 is the flat spacetime volume. The expression obtained is in terms of the discrete volume and the length of the longest chain from p to q. Since R is approximately a constant in $\mathbf{A}_d[p, q]$, this also gives the approximate action. Extending it to an action on the full spacetime is however quite tricky since it is unclear how to localise the calculation.

The calculation for the abundance of k-chains C_k in an RNN in Roy et al. (2013) also gives an expression for the curvature

$$R(0) = -2(n+2)(2n+2)(3n+2) 2^{\frac{3n+2}{3n}} n^{\frac{4}{3n}-1} \frac{(K_1 - 2K_2 + K_3)}{(J_1 - 2J_2 + J_3)^{\frac{3n+2}{3n}}}. \tag{39}$$

where

$$\begin{aligned} J_k &\equiv (kn+2) K_k \\ K_k &\equiv ((k+1)n+2) Q_k, \end{aligned} \tag{40}$$

and

$$Q_k \equiv \left(\frac{\langle \mathbf{C}_k \rangle}{\rho^k \zeta_k}\right)^{3/k} = \frac{1}{\zeta_0^3}\left(\frac{\langle \mathbf{C}_k \rangle}{\rho^k \chi_k}\right)^{3/k}. \tag{41}$$

While this expression is compact, it is not defined on a single causal set, but rather, over the ensemble. Whether this can be expressed as a function on a single causal set or not is an interesting open question and under current investigation. As in the previous case, having obtained $R(0)$, however, it is non-trivial to construct the action, without a localisation requirement as was done for the BD action.

4.6 Boundary terms for the causal set action

Although the BD action gives the bulk Einstein–Hilbert action in the continuum approximation, the role of boundary terms is less clear. As shown by Benincasa et al. (2011) the expectation value for the BD action does not vanish for $\mathcal{C}(\mathbf{A}_2, \rho_c)$, where $\mathbf{A}_2[p, q] \subset \mathbb{M}^2$ as one might expect, but instead converges to a constant as $\rho_c \to \infty$ and is independent of $\mathrm{vol}(\mathbf{A}_2)$. Buck et al. (2015) showed more generally that for $\mathcal{C}(\mathbf{A}_d, \rho_c)$ with $d \geq 2$ that

$$\lim_{N\to\infty} \frac{1}{\hbar}\left\langle \mathbf{S}^d_{\mathrm{BDG}} \right\rangle = \frac{1}{l_p^{d-2}} \mathrm{vol}(\mathcal{J}^{(d-2)}), \tag{42}$$

where $\mathcal{J}^{(d-2)} \equiv \partial J^+(p) \cap \partial J^-(q)$ is the co-dimension 2 "joint" of the causal diamond $\mathbf{A}_d$, which is a round sphere S^{d-2}. In $d = 2$ this is the volume of a zero sphere S^0 which is the constant found in Benincasa et al. (2011). This in turn corresponds to the Gibbons–Hawking–York (GHY) null boundary term of Jubb et al. (2017) and Lehner et al. (2016) for a particular choice of the null affine parameter.[19] Extending this calculation to curved spacetime is challenging but would provide additional evidence that the BD action contains the null GHY term (*Dhingra, Glaser and Surya, work in progress*).

Simulations of causal sets corresponding to different regions of $\mathbb{M}^2$ moreover suggest that while the BD action contains timelike boundary terms, it does not contain spacelike boundary terms. Recent efforts by Cunningham (2018a) have been made to obtain time like boundaries in a causal set using numerical methods for $d = 2$, but it is an open question whether they admit a simple characterisation in arbitrary dimensions.

Unlike timelike boundaries, spacelike boundaries are naturally defined in a finite element causal set: a future/past spatial boundary is the future-most/past-most inextendible antichain in the causal set, which we denote as $\mathcal{F}_0, \mathcal{P}_0$ respectively. GHY terms for spacelike boundaries play an important role in the additivity of the action in the continuum path integral (though such an additivity is far from guaranteed in a causal set because of non-locality).

[19] It is an interesting question whether the choice of affine parameter along "almost" null directions can be obtained from the causal set.

The spatial causal set GHY terms were found by Buck et al. (2015), and we will describe that construction here briefly. Let (M, g) be a spacetime with initial and final spatial boundaries $(\Sigma^{\pm}, h^{\pm})$. The GHY term on $(\Sigma^{\pm}, h^{\pm})$ can be re-expressed as

$$\int_{\Sigma^{\pm}} d^{d-1}x\, \sqrt{h^{\pm}}\, K^{\pm} = \frac{\partial}{\partial n} \int_{\Sigma^{\pm}} d^{d-1}x\, \sqrt{h^{\pm}} = \frac{\partial}{\partial n} A_{\Sigma^{\pm}}, \tag{43}$$

where $\frac{\partial}{\partial n}$ is the normal derivative, and $A_{\Sigma^{\pm}}$ is the co-dimension 1 volume of $\Sigma^{\pm}$. Using the $n \sim \rho_c v$ correspondence, this suggests that $A_{\Sigma^{\pm}}$ should be given by the cardinality $F_0 \equiv |\mathcal{F}_0|$ or $P_0 \equiv |\mathcal{P}_0|$ with the normal gradient represented by the change in the cardinality. But of course this is subtle, since apart from the future most $\mathcal{F}_0$ or pastmost $\mathcal{P}_0$ antichains, one needs another "close by" antichain. Let us focus on (Σ^{+}, h^{+}) without loss of generality. There are two ways of finding this nearby antichain. To begin with if $(M, g) \subset (M', g')$ such that (Σ^{+}, h^{+}) is not a boundary in (M', g'), then we can use this embedding to define the two antichains, in any $C \in \mathcal{C}(N, \rho_c)$: one to its immediate past $\mathcal{F}_0(\Sigma^{+})$ and one to its immediate future $\mathcal{P}_0(\Sigma^{+})$. Thus the GHY term should be proportional to the difference in the cardinality of these two antichains.

However, this partitioning is not intrinsic to the causal set. Instead consider a partition $C = C^{-} \cup C^{+}$, such that $C^{+} \cap C^{-} = \emptyset$, and $\mathrm{Fut}(C^{-}) = C^{+}$, $\mathrm{Past}(C^{+}) = C^{-}$. Let $\mathcal{F}_0^{-}$ and $\mathcal{P}_0^{+}$, be the future-most and past-most antichains of C^{-} and C^{+} respectively. We can then define the dimensionless causal set "boundary term" (Buck et al. 2015)

$$S^{d}_{\mathrm{CBT}}[C, C^{-}, C^{+}] \equiv \frac{a_d}{2\Gamma\left(\frac{2}{d}\right)} \Big(F_0[C^{-}] - P_0[C^{+}] \Big), \tag{44}$$

where

$$a_d = \frac{d(d+1)}{(d+2)} \left(\frac{\mathcal{V}_{d-2}}{d(d-1)} \right)^{\frac{2}{d}}, \tag{45}$$

and $\mathcal{V}_d = (d+1)\pi^{\frac{d+1}{2}} / \Gamma\left(\frac{d+1}{2} + 1\right)$ is the volume of the unit d-sphere.

To make contact with the continuum, let (M, g) be a spacetime with compact Cauchy hypersurfaces. For a given Cauchy hypersurface (Σ, h) let $M^{\pm} = J^{\pm}(\Sigma)$ and let $C^{\pm} \in \mathcal{C}(M^{\pm}, \rho_c)$. It was shown by Buck et al. (2015) that in the limit $\rho_c \to \infty$

$$\lim_{\rho_c \to \infty} \left(\frac{\ell_c}{\ell_p} \right)^{d-2} \left\langle \mathbf{S}^{(d)}_{\mathrm{CBT}}[M, \Sigma, \rho_c] \right\rangle = \frac{1}{l_p^{d-2}} \int_{\Sigma} d^{d-1}x\, \sqrt{h}\, K = S_{\mathrm{GHY}}(\Sigma, M^{-}), \tag{46}$$

where $\mathbf{S}^{(d)}_{\mathrm{CBT}}$ is the associated random variable in (M, g). To obtain this expression, the volume of a half cone $J^{+}(p) \cap J^{-}(\Sigma)$ was calculated using a combination of RNN

coordinates and GNN coordinates[20]

$$V_{\blacktriangle}(T,\mathbf{x}) = \frac{S_{d-2}}{d(d-1)}T^d\left(1+\frac{d}{2(d+1)}K(0,\mathbf{x})T\right)+O(T^{d+2}), \tag{47}$$

for $p \in J^+(\Sigma)$ sufficiently close to Σ, where T is the proper time from p to Σ. As might be expected from dimensional considerations, the leading order correction to the flat spacetime volume of the half cone comes from the trace of the extrinsic curvature of Σ from which the GHY contribution can be obtained.

If on the other hand, (Σ, h) is a future boundary of (M, g), then we require a second antichain in Past$(\mathcal{F}_0)$ for $C \in \mathcal{C}(M, \rho_c)$. Define the antichain $\mathcal{F}_1$ in C^- to be the set of elements in C^- such that $\forall e \in \mathcal{F}_1$, $|\text{Fut}(e) \cap C^-| = 1$ (where Fut(e) excludes the element e).[21] The boundary term can then be expressed as

$$S^d_{\text{CBT}}[C, C^-, C^+] \equiv \frac{a_d}{\Gamma\left(\frac{2}{d}\right)}\Big(dF_1[C^-] - F_0[C^+]\Big), \tag{48}$$

which again yields the GHY term Eq. (46) in the limit. Indeed, a whole family of of boundary terms was obtained using the antichains $\mathcal{F}_k[C^-] = \{e \in C^- || \text{Fut}(e)| = k\}$, $\mathcal{P}_k[C^+] = \{e \in C^+ || \text{Past}(e)| = k\}$ each of which gives the GHY term in the limit Eq. (46).[22]

A by-product of the analysis of Buck et al. (2015) is that for the partitioned causal set $C = C^- \cup C^+$ described above, the quantities

$$A^d_+[C^-] \equiv \frac{b_d}{\Gamma(\frac{1}{d})}F_0[C^-], \quad A^d_-[C^+] \equiv \frac{b_d}{\Gamma(\frac{1}{d})}P_0[C^+] \tag{49}$$

for $a_d = \frac{d+1}{d(d+2)}b_d^2$ limit to the spatial volume of Σ

$$\lim_{\rho_c\to\infty}\left(\frac{\ell_p}{\ell_c}\right)^{d-2}\langle A^d_{\pm}[C^{\mp}]\rangle = \frac{1}{\ell_p^{d-1}}\int_\Sigma d^{d-1}x\sqrt{h} = A_\Sigma. \tag{50}$$

Again, as for the boundary terms, one can construct a whole family of functions $A^d[C]$ each of which limit to the spatial volume of Σ as $\rho_c \to \infty$.

4.7 Localisation in a causal set

In these calculations generalisations are made to curved spacetime using an RNN which represent a local region of a spacetime. How are we to find such local regions in a causal set using a purely order theoretic quantities? For a causal set a natural

[20] This calculation has later been extended by Jubb (2017) to higher orders to obtain more information about the spatial geometry.

[21] Note that while $\mathcal{F}_1 \cap \mathcal{F}_0 = \emptyset$, $\mathcal{F}_1$ is not necessarily an inextendible antichain.

[22] The expression in Buck et al. (2015) holds for *any* two subsets of C not just those we consider here.

definition of a local region is given by the size of an interval, but for a manifold-like causal set, this will not necessarily correspond to regions in which the curvature is small. On the other hand, many of the order invariants we have obtained so far correspond to geometric invariants only in such RNN-type regions.

A characterisation of intrinsic localisation was obtained by Glaser and Surya (2013) using the abundance N_m^d of m element order intervals for $C \in \mathcal{C}(\mathbf{A}_d, \rho_c)$. They found the following closed form expression for the associated expectation value

$$\langle \mathbf{N}_m^d(\rho, V)\rangle = \frac{(\rho V)^{m+2}}{(m+2)!}\frac{\Gamma(d)^2}{\left(\frac{d}{2}(m+1)+1\right)_{d-1}}\frac{1}{\left(\frac{d}{2}m+1\right)_{d-1}} {}_dF_d\left(\begin{matrix}1+m, \frac{2}{d}+m, \frac{4}{d}+m, \ldots, \frac{2(d-1)}{d}+m) \\ 3+m, \frac{2}{d}+m+2, \frac{4}{d}+m+2, \ldots, \frac{2(d-1)}{d}+m+2\end{matrix}\middle| -\rho V\right), \tag{51}$$

The distribution of $\langle \mathbf{N}_m^d\rangle$ with m therefore has a characteristic form which depends on dimension, and as a by-product, can be used as a dimension estimator. However, it can also be used look for intervals in a manifold-like causal set which are approximately flat by comparing the interval abundances N_m^d to the above expression for $\langle \mathbf{N}_m^d\rangle$. While one might expect the fluctuations for a given causal set C to be large, numerical simulations show that there is typically a "self averaging" which results in relatively small fluctuations even for a given realisation. This makes it an ideal diagnostic tool for checking whether a neighbourhood in a manifold-like causal set is approximately flat or not. Once such local neighbourhoods have been found, a local check of geometric estimators can be made.

In Glaser and Surya (2013), the analytic curves were compared against simulations for a range of different causal sets including those that are not manifold-like (see Fig. 16). While curvature affects the abundance of the intervals, the distribution retains its characteristic form. Hence the dependence of the abundance of intervals with size also becomes a test for manifold-likeness.

There are other ways of testing for manifold-likeness. In a similar approach, the distribution of the longest chains or linked paths of length k in a finite element causal set C has been studied in $\mathbb{M}^d$, $d = 2, 3, 4$ and shown to have a dimension-dependent peak (Aghili et al. 2019). In Bolognesi and Lamb (2016), a novel way to test for manifold-likeness was given, using the order invariant obtained from counting the number of elements with a fixed valency in a finite element causal set. In Henson (2006a), an algorithm for determining the embeddability of a causal set in $\mathbb{M}^2$ was given, which again gives an intrinsic characterisation of manifold-likeness in $d = 2$. Extending and expanding on these studies using causal sets obtained from sprinklings into different types of spacetimes would be a straightforward but useful exercise.

4.8 Kinematical entropy

Since the classical continuum geometry itself is fundamentally statistical in CST, it is interesting to ask if a kinematic entropy can be assigned even classically to the con-

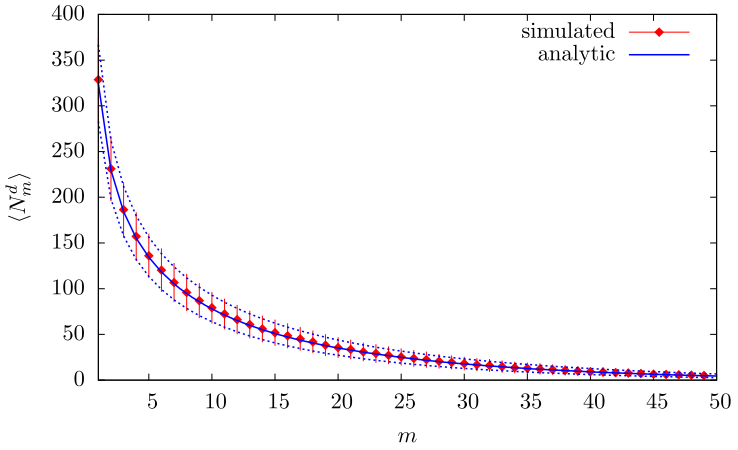

Fig. 16 The expectation value of interval abundances in a 100 element causal set $\sim \mathbb{M}^2$ as a function of interval size m. The red dots depict the average value obtained from simulations with 1000 realisations, along with error bars. The solid blue line depicts the analytic expectation value for $n = 100$ and the blue dotted lines for $n \pm \sqrt{n}$

tinuum. In Dou and Sorkin (2003), a kinematic entropy was associated with a horizon H and a spatial or null hypersurface Σ in a dimensionally reduced $d = 2$ black hole spacetime by counting links between elements in $J^-(\Sigma) \cap J^-(H)$ and those in $J^+(\Sigma) \cap J^+(H)$, with the additional requirement that the former is future-most and the latter past-most in their respective regions. A dimensionally reduced calculation showed that the number of links is proportional to the horizon area. Importantly, the calculation yields the same constant for a dimensionally reduced dynamical spacetime where a collapsing shell of null matter eventually forms a black hole. However, extending this calculation to higher dimensions proves to be tricky. In Marr (2007), an entropy formula was proposed for higher dimensions by replacing links with other sub-causal sets. While these ideas hold promise, they have not as yet been fully explored.

In analogy with Susskind's entropy bound, the maximum causal set entropy associated with a finite spherically symmetric spatial hypersurface Σ was defined by Rideout and Zohren (2006) as the number of maximal or future most elements in its future domain of dependence $D^+(\Sigma)$. It was shown that for several such examples this bound limits to the Susskind entropy bound in the continuum approximation. Again, extending this discussion to more general spacetimes is an interesting open question.

In Benincasa (2013), the mutual information between different regions of a causal set was defined using the BD action. The source of this entropy is non-locality which implies that S_{BD} is not in general additive. Dividing a causal set C into two (set-wise) disjoint regions C_1 and C_2, so that $C = C_1 \sqcup C_2$, we see that in general $S_{\mathrm{BD}}(C) \neq S_{\mathrm{BD}}(C_1) + S_{\mathrm{BD}}(C_2)$. This is because there can be order intervals between elements in C_1 and in C_2 that are not counted by either $S_{\mathrm{BD}}(C_1)$ or $S_{\mathrm{BD}}(C_2)$. The mutual information is thus defined as

$$MI[C_1, C_2] \equiv S_{\mathrm{BD}}(C_1) + S_{\mathrm{BD}}(C_2) - S_{\mathrm{BD}}(C). \tag{52}$$

In Benincasa (2013) a spacetime region with a horizon $\mathcal{H}$ and a spacelike or null hypersurface Σ was considered. Defining $X = J^+(\mathcal{H}) \cap J^-(\Sigma)$ and $Y = J^-(\mathcal{H}) \cap J^-(\Sigma)$ the mutual information between X and Y was calculated from a causal set obtained from sprinkling into $X \cup Y$. Under certain assumptions, this equal to the area of $\mathcal{H} \cap \Sigma$. These results are suggestive, but currently incomplete.

As we will see in the next section, the Sorkin spacetime entanglement entropy (SSEE) for a free scalar field provides a different avenue for exploring entropy.

4.9 Remarks

To conclude this section we note that several order invariants have been constructed on manifold-like causal sets whose expectation values limit to manifold invariants as $\rho_c \to \infty$. At finite ρ_c there are fluctuations that serve to distinguish the fundamental discreteness of causal sets from the continuum, and these have potential phenomenological consequences. Numerical simulations are often important in assessing the relative importance of these fluctuations.

For each of these invariants, one has therefore proved an *$\mathcal{O}$-Hauptvermutung*. While this collection of order invariants is not sufficient to prove the full Hauptvermutung, they lend it strong support. These order invariants are moreover important observables for the full theory. In addition to these manifold-like order invariants, there are several other order invariants that can be constructed, some of which may be important to the deep quantum regime but by themselves hold no direct continuum interpretation.

5 Matter on a continuum-like causal set

Before passing on to the dynamics of CST, we look at a phenomenologically important question, namely how quantum fields behave on a fixed manifold-like causal set. The simplest matter field is the free scalar field on a causal set in $\mathbb{M}^d$. As we noted in the previous Section, this is the only class of matter fields that we know how to study, since at present no well defined representation of non-trivial tensorial fields on causal sets is known. However, as we will see, even this very simple class of matter fields brings with it both exciting new insights and interesting conundrums.

5.1 Causal set Green functions for a free scalar field

Consider the real scalar field $\phi : \mathbb{M}^d \to \mathbb{R}$ and its CST counterpart, $\phi : C \to \mathbb{R}$ where $C \in \mathcal{C}(\mathbb{M}^d, \rho_c)$. The Klein Gordon operator of the continuum is replaced on the causal set by the B_κ operator of Sect. 4, Eq. (36). In the continuum $\Box^{-1}$ gives the Green function, and we can do the same with B_κ to obtain the discrete Green function B_κ^{-1}.

However, there are more direct ways of obtaining the Green function as was shown in Daughton (1993), Salgado (2008), Johnston (2008) and Dowker et al. (2017). The *causal matrix*

$$C_0(e, e') \equiv \begin{cases} 1 & \text{if } e' \prec e \\ 0 & \text{otherwise} \end{cases} \tag{53}$$

on a causal set C. For $C \in \mathcal{C}(\mathbb{M}^d, \rho_c)$, $C_0(e, .)$ is therefore zero everywhere except within the past light cone of e at which it is 1. In $d = 2$, this is just half the massless retarded Green's function $G_0^{(2)}(x, x') = \frac{1}{2}\theta(t - t')\theta(\tau^2(x, x'))$. Hence, we find the almost trivial relation

$$\mathbf{C}_0(x, x') = 2G_0^{(2)}(x, x'), \tag{54}$$

without having to take an expectation value, so that the dimensionless massless causal set retarded Green function is (Daughton 1993)

$$K_0^{(2)}(x, x') \equiv \frac{1}{2}C_0(x, x'). \tag{55}$$

To obtain the $d = 4$ massless causal set Green function we use the *link matrix*

$$L_0(x, x') := \begin{cases} 1 & \text{if } x' \prec x \; is \; a \; link \\ 0 & \text{otherwise} \end{cases} \tag{56}$$

For $C \in \mathcal{C}(\mathbb{M}^4, \rho_c)$ the expectation value of the associated random variable is

$$\langle \mathbf{L}_0(x, x') \rangle = \theta(x_0 - x_0')\theta(\tau^2(x, x')) \exp(-\rho_c V(x, x')), \tag{57}$$

where $V(x, x') = \text{vol}(J^-(x) \cap J^+(x')) = \frac{\pi}{24}\tau^4(x, x')$. Since the exponential in the above expression is a Gaussian which, in the $\rho_c \to \infty$ limit is proportional to $\delta(\tau^2)$, we see that it resembles the massless retarded Green function in $\mathbb{M}^4$,

$$\lim_{\rho_c \to \infty} \sqrt{\frac{\rho_c}{6}} \langle \mathbf{L}_0(x, x') \rangle = \theta(x_0 - x_0')\delta(\tau^2) = 2\pi G_0^{(4)}(x, x'). \tag{58}$$

Hence we can write the dimensionless massless causal set scalar retarded Green function as (Johnston 2008, 2010)

$$K_0^{(4)}(x, x') = \frac{1}{2\pi}\sqrt{\frac{1}{6}} L_0(x, x') \,. \tag{59}$$

In the continuum the massive Green function can be obtained from the massless Green function in $\mathbb{M}^d$ via the formal expression (Dowker et al. 2017)

$$G_m = G_0 - m^2 \, G_0 * G_0 + m^4 \, G_0 * G_0 * G_0 + \cdots = \sum_{k=0}^{\infty} (-m^2)^k \underbrace{G_0 * G_0 * \ldots G_0}_{k+1} \tag{60}$$

where

$$(A * B)(x, x') \equiv \int d^d x_1 \sqrt{-g(x_1)} A(x, x_1) B(x_1, x') \,. \tag{61}$$

Using this as a template, with the discrete convolution operation given by matrix multiplication,

$$(A * B)(e, e') \equiv \sum_{e''} A(e, e'') B(e'', e), \tag{62}$$

a candidate for the $d = 2$ dimensionless massive causal set Green function is

$$K_M^{(2)}(x, x') = \frac{1}{2} \sum_{k=0}^{\infty} (-1)^k \frac{M^{2k}}{2^k} C_k(x, x'). \tag{63}$$

Here M is dimensionless and we have used the relation $C_k(x, x') = C_0^k(x, x')$, where the product is defined by the convolution operation Eq. 61 and, $C_k(x, x')$ counts the number of k-element chains from x to x'. For $C \in \mathcal{C}(\mathbb{M}^2, \rho_c)$ it can be shown that (Johnston 2008, 2010)

$$\langle \mathbf{K}_M^{(2)}(x, x') \rangle = G_m^{(2)}(x, x') \,, \tag{64}$$

when $M^2 = \frac{m^2}{\rho_c}$. Similarly, a candidate for the $d = 4$ massive causal set Green function is

$$K_M^{(4)}(x, x') = \frac{1}{2\pi\sqrt{6}} \sum_{k=0}^{\infty} (-1)^k \left(\frac{M^2}{2\pi\sqrt{6}} \right)^k L_k(x, x'), \tag{65}$$

where we have used the fact that the number of k-element linked paths $L_k(x, x') = L_0^k(x, x')$. For $C \in \mathcal{C}(\mathbb{M}^4, \rho_c)$,

$$\lim_{\rho_c \to \infty} \sqrt{\rho_c} \langle \mathbf{K}_M^{(4)}(x, x') \rangle = G_m^{(4)}(x, x') \,, \tag{66}$$

when $M^2 = \frac{m^2}{\sqrt{\rho_c}}$.

These massive causal set Green function were first obtained by Johnston (2008, 2010) using an evocative analogy between Feynman paths and the k-chains or k-linked paths (see Fig. 17). "Amplitudes" a and b are assigned to a "hop" between two elements in the Feynman path, and to a "stop" at an intervening element, respectively. This gives a total "amplitude" $a^{k+1} b^k$ for each chain or linked path, so that the massive Green functions can be expressed as

$$K_m^{(2)}(e, e') \equiv \sum_{k=0} a_2^{k+1} b_2^k C_k(e, e'), \quad K_m^{(4)}(e, e') \equiv \sum_{k=0} a_4^{k+1} b_4^k L_k(e, e'), \tag{67}$$

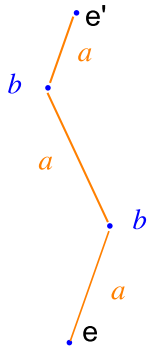

Fig. 17 The hop and stop amplitudes a and b on a 2-element chain from e to e' for a massive scalar field on a causal set

where the coefficients a_d, b_d are set by comparing with the continuum.

Finding causal set Green functions for other spacetimes is more challenging, but there have been some recent results (Dowker et al. 2017) which show that the flat spacetime form of Johnston (2008, 2010) can be used in a wider context. These include (a) a causal diamond in an RNN of a $d = 2$ spacetime with $M^2 = {\rho_c}^{-1}(m^2 + \xi R(0))$, where $R(0)$ is the Ricci scalar at the centre of the diamond and ξ is the non-minimal coupling, (b) a causal diamond in an RNN of a $d = 4$ spacetime with $R_{ab}(0) \propto g_{ab}(0)$ and $M^2 = {\rho_c}^{-1}(m^2 + \xi R(0))$ when (c) $d = 4$ de Sitter and anti de Sitter spacetimes with $M^2 = {\rho_c}^{-1}(m^2 + \xi)$.

The de Sitter causal set Green function in particular allows us to explore cosmological consequences of discreteness, one of which we will describe below. It would be useful to extend this construction to other conformally flat spacetimes of cosmological relevance like the flat FRW spacetimes. Candidates for causal set Green functions in $\mathbb{M}^3$ have also been obtained using both the volume of the causal interval and the length of the longest chain (Johnston 2010; Dowker et al. 2017), but the comparisons with the continuum need further study.

As the attentive reader would have noticed, in $d = 4$ the causal set Green function matches the continuum only for $\rho_c \to \infty$, unlike in $d = 2$. At finite ρ_c, there can be potentially observable differences with the continuum. Comparisons with observation can therefore put constraints on CST. Dowker et al. (2010a) examined a model for the propagation of a classical massless scalar field from a source to a detector on a background causal set. In $\mathbb{M}^d$, an oscillating point source with scalar charge q, frequency ω and amplitude a, and a "head-on" rectangular shaped detector was considered, so that the field produced by the source is

$$\phi(y) = \int_P G(y, x(s))q\,ds \tag{68}$$

where $\mathcal{P}$ is the world line of the source and s the proper time along this world line. If $\mathcal{D}$ represents the spacetime volume swept out by the detector during its detection time T then the output of the detector is

$$F = \int_{\mathcal{D}} \phi(y) d^4y = q \int_{\mathcal{P}} ds \int_{\mathcal{D}} d^4y G(y, x(s)) \approx \sqrt{\frac{1+\nu}{1-\nu}} \frac{q}{4\pi R} v_{\mathcal{D}} \tag{69}$$

where R is the distance between the source and detector, ν is the component of the velocity along the displacement vector between the source and detector and $v_{\mathcal{D}}$ is the spacetime volume of the detector region $\mathcal{D}$. Here, $R >> a$ and $R >> \omega^{-1}$ which in turn is much larger than the spatial and temporal extent of the detector region $\mathcal{D}$. The causal set detector output can then be defined as

$$\widetilde{F} = q \frac{1}{2\pi\sqrt{6}} \sum_{e \in \tilde{\mathcal{P}}} \sum_{e' \in \tilde{\mathcal{D}}} L_0(e', e) \tag{70}$$

where $\tilde{\mathcal{D}}$ and $\tilde{\mathcal{P}}$ correspond to the detector and source subregions in the causal set and the causal set function $L(e, e')$ is equal to some normalisation constant κ when e and e' are linked and is zero otherwise. For $C \in \mathcal{C}(\mathbb{M}^4, \rho_c)$ it was shown that, with the above constraints on R, ω, a and the dimensions of the detector, that $\langle \widetilde{\mathbf{F}} \rangle$ approximates to same continuum expression Eq. (69) when $R >> \rho_c^{-\frac{1}{4}}$. A detailed calculation gives an upper bound on the fluctuations, which, for a particular AGN model is one part in 10^{12} for $\rho_c = \rho_p$. Hence the discreteness does not seem to mess with the coherence of waves from distant sources. As we will see in Sect. 7 there are other potential signatures of the discreteness that may have phenomenological consequences (Dowker et al. 2004; Sorkin 1991, 1997; Ahmed et al. 2004).

5.2 The Sorkin–Johnston (SJ) vacuum

Having obtained the classical Green function and the d'Alembertian operator in $\mathbb{M}^2$ and $\mathbb{M}^4$, the obvious next step is to build a full quantum scalar field theory on the causal set. As we have mentioned earlier, the canonical route to quantisation is not an option for causal sets nor for fields on causal sets and hence there is a need to look at more covariant quantisation procedures.

Johnston (2009, 2010) used the the covariantly defined Peierls' bracket

$$[\widehat{\Phi}(x), \widehat{\Phi}(y)] = i\Delta(x, y) \tag{71}$$

as the starting point for quantisation, where

$$\Delta(x, y) \equiv G^R(x, x') - G^A(x, x') \tag{72}$$

is the Pauli Jordan function, and $G^{R,A}(x, x')$ are the retarded and advanced Green's functions, respectively. As we have seen, these Green functions can be defined on certain manifold-like causal sets and hence provide a natural starting point for quantisation.

However, even here, the standard route to quantisation involves the mode decomposition of the space of solutions of the Klein Gordan operator, $\ker(\Box - m^2)$. In $\mathbb{M}^d$ the space of solutions has a unique split into positive and negative frequency classes

of modes with respect to which a vacuum can be defined. In his quest for a Feynman propagator, Johnston (2009) made a bold proposal, which as we will describe below, has led to a very interesting new direction in quantum field theory even in the continuum. This is the *Sorkin–Johnston or SJ vacuum* for a free quantum scalar field theory.

Noticing that the Pauli–Jordan function on a finite causal set C is a Hermitian operator, and that $\Delta(e, e')$ itself is antisymmetric, Johnston used the fact that the eigenspectrum of $i\Delta$

$$i\widehat{\Delta} \circ v_{\mathbf{k}}(e) \equiv \sum_{e' \in C} i\Delta(e, e') v_{\mathbf{k}}(e') = \lambda_{\mathbf{k}} v_{\mathbf{k}}(e) \tag{73}$$

splits into pairs $(\lambda_{\mathbf{k}}, -\lambda_{\mathbf{k}})$, with eigenfunctions $(v_{\mathbf{k}}^{+}, v_{\mathbf{k}}^{-})$, $v_{\mathbf{k}}^{-} = v_{\mathbf{k}}^{+*}$. This provides a natural split into a positive part and a negative part, without explicit reference to $\ker(\Box - m^2)$.[23] A spectral decomposition of $i\widehat{\Delta}$ then gives

$$i\Delta(e, e') = \lambda_{\mathbf{k}} \sum_{\mathbf{k}} v_{\mathbf{k}}^{+}(e) v_{\mathbf{k}}^{+*}(e') - v_{\mathbf{k}}^{+}(e)^{*} v_{\mathbf{k}}^{+}(e'). \tag{74}$$

This decomposition is used to define the SJ Wightmann function as the positive part of $i\Delta$

$$W_{SJ}(e, e') \equiv \lambda_{\mathbf{k}} \sum_{\mathbf{k}} v_{\mathbf{k}}^{+}(e) v_{\mathbf{k}}^{+*}(e'). \tag{75}$$

Importantly, for a non-interacting theory with a Gaussian state, the Wightmann function is sufficient to describe the full theory and thus the vacuum state. Simulations in $\mathbb{M}^d$ for $d = 2, 4$ give a good agreement with the continuum (Johnston 2009, 2010).

Sorkin (2011a) noticed that the construction on the causal set, which was born out of necessity, provides a new way of thinking of the quantum field theory vacuum. A well known feature of quantum field theory in a general curved spacetime is that the vacuum obtained from mode decomposition in $\ker(\widehat{\Box} - m^2)$ is observer dependent and hence not unique. Since the SJ vacuum is intrinsically defined, at least in finite spacetime regions, one has a uniquely defined vacuum. As a result, the SJ state has generated some interest in the broader algebraic field theory community (Fewster and Verch 2012; Brum and Fredenhagen 2014; Fewster 2018). For example, while not in itself Hadamard in general, the SJ vacuum can be used to generate a new class of Hadamard states (Brum and Fredenhagen 2014).

In the continuum, the SJ vacuum was constructed for the massless scalar field in the $d = 2$ causal diamond (Afshordi et al. 2012) and recently extended to the small mass case (Mathur and Surya 2019). It has also been obtained for the trousers topology and shown to produce a divergent energy along both the future and the past light cones associated with the Morse point singularity (Buck et al. 2017). Numerical simulations

[23] The identification of $\ker(\Box - m^2)$ with $Im(i\Delta)$ is in fact well known (Wald 1994) when the latter is restricted to functions of compact support.

of the SJ vacuum on causal sets are are approximated by de Sitter spacetime suggest that the causal set SJ state differs significantly from the Mottola–Allen α vacuua (Surya et al. 2019). This has potentially far reaching observational consequences which need further investigation.

5.3 Entanglement entropy

Using the Pauli Jordan operator $i\widehat{\Delta}$ and the associated Wightman $\widehat{W}$, Sorkin (2014) defined a spacetime entanglement entropy, *Sorkins' Spacetime Entanglement Entropy (SSEE)*

$$S = \sum_i \lambda_i \ln |\lambda_i| \tag{76}$$

where λ_i are the generalised eigenvalues satisfying

$$\widehat{W} \circ v_i = i\lambda_i \widehat{\Delta} \circ v_i. \tag{77}$$

It was shown by Saravani et al. (2014) that for a causal diamond sitting at the centre of a larger one in $\mathbb{M}^2$, S has the expected behaviour in the limit that the size of the smaller diamond l is much smaller than that of the larger diamond,

$$S = b \ln \left(\frac{l}{l_{uv}} \right) + c, \tag{78}$$

where l_{uv} is the UV cut-off and b, c are constants that can be determined.

One of the promises that discretisation holds is of curing the UV divergences of quantum field theory and in particular those coming from the calculation of the entanglement entropy of Bombelli et al. (1986). As shown by Sorkin and Yazdi (2018) the causal set version of the above calculation is proportional to the volume rather than the above "area", thus differing from the continuum. This can be traced to the fact that the continuum spectrum of eigenvalues (Eq. 77) agrees with the discrete eigenvalues only up to a "knee", beyond which the effects of discreteness become important, as shown in Fig. 18. Using a double truncation of the spectrum—once in the larger diamond and once in the smaller one, Sorkin and Yazdi (2018) obtained the requisite area law. This raises very interesting and as yet unanswered puzzles about the nature of SSEE in the causal set. It is for example possible that in a fundamentally non-local theory like CST an area law is less natural than a volume law. Such a radical understanding could force us to rethink continuum inspired ideas about Black Hole entropy.

Extending the above calculation to actual black hole spacetimes is an important open problem. Ongoing simulations for causal sets obtained from sprinklings into 4d de Sitter spacetime show that this double truncation procedure gives the right de Sitter horizon entropy (*Dowker, Surya, X and Yazdi, work in progress*), but one first needs to make an ansatz for locating the knee in the causal set $i\Delta$ spectrum.

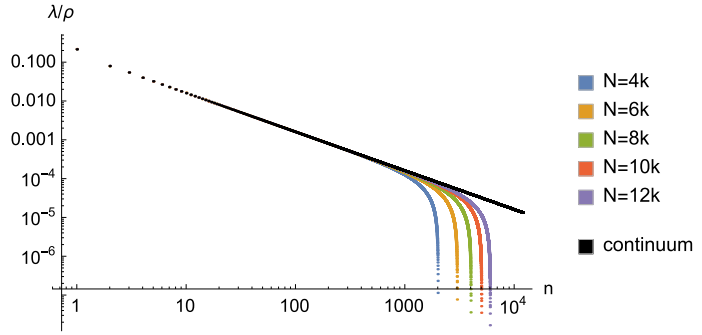

Fig. 18 A log-log plot depicting the SJ spectra for causal sets in a causal diamond in $\mathbb{M}^2$. A comparison with the continuum (the straight black line) shows that the causal set SJ spectrum matches the continuum in the IR but has a characteristic "knee" in the UV after which it deviates significantly from the continuum. As the density of the causal set increases, this knee shifts to the UV

5.4 Spectral dimensions

An interesting direction in causal set theory has been to calculate the spectral dimension of the causal set (Eichhorn and Mizera 2014; Belenchia et al. 2016c; Carlip 2017). Carlip (2017) has argued that $d = 2$ is special in the UV limit, and that several theories of quantum gravity lead to such a dimensional reduction. In light of how we have presented CST, it seems that this continuum inspired description must be limited. It is nevertheless interesting to ask if causal sets that are manifold-like might exhibit such a behaviour around the discreteness scales at which the continuum approximation is known to break down. As we have seen earlier (Sect. 4.3), one such behaviour is discrete asymptotic silence (Eichhorn et al. 2017).

Eichhorn and Mizera (2014) calculated the spectral dimension on a causal set using a random walk on a finite element causal set. It was found that in contrast, the dimension at small scales goes up rather than down. On the other hand, Belenchia et al. (2016c) showed that causal set inspired non-local d'Alembertians do give a spectral dimension of 2 in all dimensions. As we noted in Sect. 4, Abajian and Carlip (2018) showed that dimensional reduction of causal sets occurs for the Myrheim–Meyer dimension as one goes to smaller scales. Recently in Eichhorn et al. (2019b), the spectral dimension was calculated on a maximal antichain for a causal set obtained from sprinklings into $\mathbb{M}^d$, $d = 2, 3$ using the induced distance function of Eichhorn et al. (2019a). It was seen to decrease at small scales, thus bringing the results closer to that conjectured by Carlip (2017).

6 Dynamics

Until now our focus has been on manifold-like causal sets, since the aim was to find useful manifold-like covariant observables as well as to make contact with phenomenology. However, as discussed in Sect. 3, the arena for CST is a sample space Ω of locally finite posets which replaces the space of 4-geometries, and contains non-

manifold-like causal sets. A CST dynamics is given by the measure triple $(\Omega, \mathfrak{A}, \mu)$ where $\mathfrak{A}$ is an event algebra and μ is either a classical or a quantum measure. We will define these quantities later in this section.

To begin with, Ω itself can be chosen depending on the particular physical situation in mind. In the context of initial conditions for cosmology, for example, it is appropriate to restrict to the sample space of past finite countable causal sets Ω_g, while for a unimodular type dynamics using the Einstein–Hilbert action, the natural restriction is to Ω_n the sample space of causal sets of fixed cardinality n. We will see that dimensional restrictions on the sample space are also of interest and can lead to a closer comparison with other approaches to quantum gravity.

As discussed in Sects. 3 and 4, in the asymptotic $n \to \infty$ limit the sample space Ω_n is dominated by the non-manifold-like KR causal sets depicted in Fig. 9. This is the "entropy problem" of CST. These posets have approximately just three "moments" of time and hence should not play a role in the classical or continuum approximation of the theory.

For a quantum dynamics of CST we would like to start with a few basic axioms, including discrete general covariance and dynamical causality. A very important step in this direction was made by the *classical sequential growth* models (CSG)(Rideout and Sorkin 2000a), which are Markovian growth models. We will describe these in Sects. 6.1 and 6.2.

One of the main challenges in CST is to build a viable *quantum sequential growth*(QSG) dynamics. The appropriate framework for the dynamics is as a quantum measure space which is a natural quantum generalisation of classical stochastic dynamics (Sorkin 1994, 1995, 2007d). This means replacing the classical probability measure P in the measure space triple $(\Omega, \mathcal{A}, \mu_c)$ with a *quantum measure* μ. The quantum measure is defined via a decoherence functional and can also be defined as a *vector measure* in a corresponding *histories Hilbert space*. We will discuss this in Sect. 6.3.

It is also of interest to construct an effective continuum-inspired dynamics, where the discrete Einstein–Hilbert or BD action is used to give the measure for the discrete path integral or path sum. The quantum partition function can either be evaluated directly or converted into a statistical partition function over causal sets using an analytic continuation. This makes it amenable to Markov Chain Monte Carlo (MCMC) simulations as we will see below in Sect. 6.4.

6.1 Classical sequential growth models

The Rideout and Sorkin (2000a) classical sequential growth or CSG models are a class of stochastic dynamics in which causal sets are grown element by element, with the dynamics satisfying a few basic principles (Rideout and Sorkin 2000a, 2001; Martin et al. 2001; Rideout 2001; Varadarajan and Rideout 2006). The stochastic dynamics finds a natural expression in measure theory and allows for an explicit definition of covariant classical observables (Brightwell et al. 2003; Dowker and Surya 2006). This measure theoretic structure provides an important template for the quantum theory, and hence we will first flesh it out in some detail before discussing quantum dynamics.

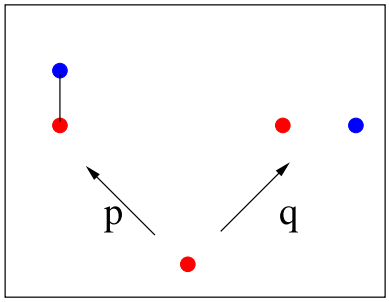

Fig. 19 The first two stages of a classical sequential growth(CSG) dynamics. The probability for a single element (red) to appear at coordinate time $n = 1$ is 1. Subsequently, the new element (blue) at $n = 2$ is added either to the future of the existing element with probability p or is unrelated to it with probability $1 - p$

Let us start with a naive picture. Imagine living on a *classical* causal set universe, with our universe represented by a single causal set. Since causal sets are locally finite, the "passage of time" occurs with the addition of a new element. If we are to respect causality, this new element cannot be added so as to disturb the past. Instead it can be added to the future of some of the existing events or it can be unrelated to all of them. Every such "atomic change" in spacetime corresponds to the causal set changing cardinality or "growing" by one. Starting with a causal set $\tilde{c}_n$ of cardinality n, the passage of time means transitioning from $\tilde{c}_n \to \tilde{c}_{n+1}$ where the new element in $\tilde{c}_{n+1}$ is to the future of some of the elements of $\tilde{c}_n$, but never in their past. In the infinite "time" limit, $n \to \infty$, the dynamics, either deterministic, probabilistic or quantum, will take you from $\tilde{c}_n$ to a countable causal set.

Working *backwards*, on the other hand, leads us to a "beginning", with $n = 0$. This gives the most natural initial condition[24] for causal sets: begin with the empty set $\emptyset$. The only way to go forward from here, is to make $n = 1$, i.e., we have a single element. For $n = 2$, the new element could either be to the future of the existing element or unrelated to it, as in Fig. 19.

Thus, one can build up the tree $\mathcal{T}$ of causal sets as $n \to \infty$ as shown in Fig. 20. As n increases, the number of possibilities grows superexponentially as expected from the KR theorem (Kleitman and Rothschild 1975), and there is no easy enumeration of this space. The growth process generates a sample space $\tilde{\Omega}_g$ of countable causal sets which are are all *past finite* and labelled by the "time" at which each element is added. A causal set $\tilde{c}$ in $\tilde{\Omega}_g$ is said to be *naturally labelled*, i.e., there exists an injective map $L : \tilde{c} \to \mathbb{N}$ (the natural numbers) which preserves the order relation in $\tilde{c}$, i.e., $e \prec e' \Rightarrow L(e) < L(e')$. In the growth process, this label is the coordinate time.

In the spirit of covariance, however, we cannot take the time label to be fundamental; the dynamics and the observables cannot depend on the order in which the elements are born. Thus, the probability to get a labelled causal set $\tilde{c}_n$ and any of its relabellings, $\tilde{c}'_n$ must be the same. Identifying relabelled causal sets as the same object in the CST tree $\mathcal{T}$ gives us a non-trivial poset of causal sets or the "postcau" $\mathcal{P}$ of Rideout and Sorkin (2000a). On $\mathcal{P}$, a covariant dynamics is thus path-independent: if there is more than one path from an unlabelled initial causal set c_{n_i} to an unlabelled final causal set c_{n_f} in $\mathcal{P}$, then in order to satisfy covariance, the measure on both paths should be the same.

[24] Of course, we could insist that there is no beginning, in which case n is never finite.

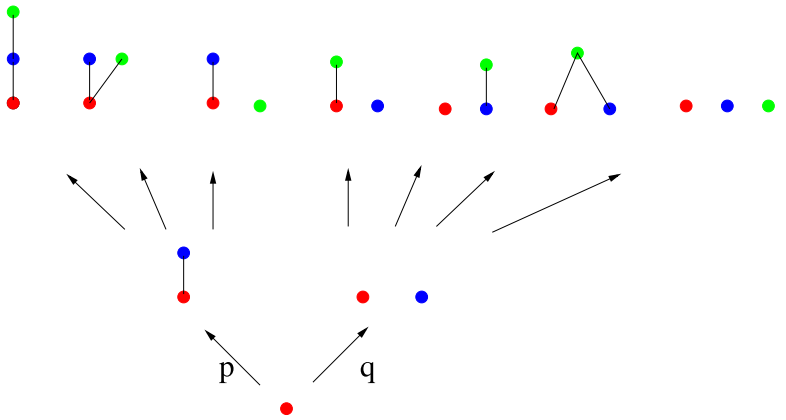

Fig. 20 The CSG tree $\mathcal{T}$. There are three ways to get the 3-element unlabelled causal set whose natural labellings are given by the 3rd, 4th and 5th 3-element labelled causal sets in the figure. One path is via the 2-element chain and the other two are from the 2-element antichain. Covariance demands that the probability along each path is the same

Apart from covariance, this dynamics also satisfies an internal causality condition, dubbed *Bell causality*. Consider the transition $\tilde{c}_n \to \tilde{c}_{n+1}$ with probability α_n where the new element e_{n+1} is added to the future of a "precursor" set $p_n \subset \tilde{c}_n$, and is unrelated to a "spectator set" $s_n \subset \tilde{c}_n$. Causality suggests that the probability for the transition should not depend on the spectator set s_n. For non-empty s_n with $|s_n| < n$, consider the causal sets $\tilde{c}_m = \tilde{c}_n \backslash s_n$ and $\tilde{c}_{m+1} = \tilde{c}_{n+1} \backslash s_n$, where $\backslash$ denotes set difference and $m + |s_n| = n$. The transition probability α_m for $\tilde{c}_m \to \tilde{c}_{m+1}$ should then be proportional to α_n. If $\tilde{c}_n \to \tilde{c}'_{n+1}$ is another transition from $\tilde{c}_n$, then defining p'_n, s'_n, α'_n, and $\tilde{c}'_{m+1} = \tilde{c}'_{n+1} \backslash s'_n$, analogously, the condition of Bell causality is

$$\frac{\alpha_n(\tilde{c}_n \to \tilde{c}_{n+1})}{\alpha'_n(\tilde{c}_n \to \tilde{c}'_{n+1})} = \frac{\alpha_m(\tilde{c}_m \to \tilde{c}_{m+1})}{\alpha'_m(\tilde{c}_m \to \tilde{c}'_{m+1})} \tag{79}$$

Though relatively easy to implement classically, a quantum version of Bell causality has been hard to find (Henson 2011).

The triple requirements of (a) covariance, (b) Bell causality and (c) Markovian evolution define the classical sequential growth dynamics of Rideout and Sorkin (2000b). Starting from the empty set, a causal set is thus grown element by element, assigning probabilities to each transition $\tilde{c}_n$ to a $\tilde{c}_{n+1}$, consistent with these requirements. Because of it being a Markovian evolution, the probability associated with any finite c_n is given by the product of the transition probabilities along a path in $\mathcal{P}$.

The dynamics was shown in Rideout and Sorkin (2000a) to be fully determined by the infinite set of coupling constants, t_n, one for each stage of the growth. If q_k denotes the transition probability from the k-element antichain to the $k+1$-element antichain, these coupling constants can be expressed as

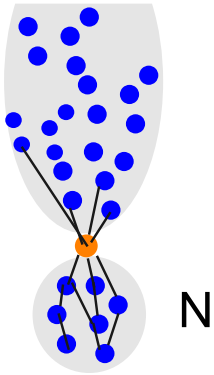

Fig. 21 A post is an analogue of a bounce in causal set cosmology

$$t_n \equiv \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} \frac{1}{q_k}. \tag{80}$$

In general, the t_n can be independent of each other. Including relations between the different t_n thus simplifies the dynamics. The simplest example is that of *transitive percolation* determined by the probability $(1-q) \geq 0$ of adding an element to the immediate future of an existing element,[25] and q of being unrelated to it. Thus, the probability of adding a new element to the immediate future of m elements of c_n and of being unrelated to m' others is $(1-q)^m q^{m'}$. In terms of the general coupling constants, $t_n = t^n \equiv \left(\frac{1-q}{q}\right)^n$.

In Varadarajan and Rideout (2006) and Dowker and Surya (2006), a generalisation of the dynamics was explored, where some of the transition probabilities were allowed to vanish, consistent with (a) (b) and (c). This requires a generalisation of the Bell causality condition. The resulting dynamics exhibits a certain "forgetfulness" when these transition probabilities vanish, but are otherwise very similar to the CSG models.

Since the generic dynamics consistent with (a), (b) and (c) does not by itself lead to constraints on the t_n, this is an embarrassment of riches. Does nature pick out one set over another? In Martin et al. (2001), an evolutionary mechanism for doing so was suggested using cosmological bounces which give rise to new epochs which "renormalise" the coupling constants towards fixed points. A cosmological bounce in a causal set is naturally described by the appearance of a *post* which is an inextendible antichain of cardinality 1 (Fig. 21).

Thus, every element in c either lies to its past or to its future. Moreover, because it is a single element maximal antichain, there are no "missing links" (see Fig. 11), and the post is indeed a summary of its past. The post is the causal set equivalent to a "bounce" but is non-singular in the causal set. We define the causal set between two posts as an "epoch", with the last epoch being the one after the last post. Let e be a post in c and let $r = |\mathrm{Past}(e)|$. Then a set of "effective" coupling constants in the epoch after e can be defined as (Martin et al. 2001)

[25] By this we mean that the new element is "linked" to an existing one, not just related to it.

$$\tilde{t}_n^{(r)} = \sum_{k=0}^{r} \binom{r}{k} t_{n+k}. \tag{81}$$

Thus, the memory of the past of the post, which is common to all the elements to the future of the post is "washed" out, but not without "dressing" up the new effective coupling constants. Denoting the set of effective couplings by $\mathcal{T}^{(i)} \equiv \{t_0^{(i)}, t_1^{(i)}, \ldots\}$ with $i = 0$ being the original set of couplings, this corresponds to applying r copies of the transform $M : \mathcal{T}^{(i)} \to \mathcal{T}^{(i+1)}$ where $t_n^{(i+1)} = t_n^{(i)} + t_{n+1}^{(i)}$, $i = 0, \ldots r-1$. In Martin et al. (2001), it was shown that the fixed points of the map M give $t_n = t^n$ (transitive percolation) for some $t \geq 0$ and moreover M does not have any other cycles. Starting from any set $\mathcal{T}^{(0)}$ for which $\lim_{n\to\infty}(t_n^{(0)})^{1/n}$ is finite, $M^r : \mathcal{T}^{(0)} \to \mathcal{T}^{(r)}$, is such that $\mathcal{T}^{(r)}$ converges pointwise to $t_n^{(r)} = t^n$ for $t = \lim_{n\to\infty}(t_n^{(0)})^{1/n}$. While this result does not guarantee that every $\mathcal{T}^{(0)}$ will converge to transitive percolation, Martin et al. (2001) examined several cases, and conjectured that the deviation from percolation-like values are "rare" and that typically, $\mathcal{T}^{(r)}$ will be nearly like transitive percolation.

Such an evolutionary renormalisation thus brings the infinite dimensional coupling constant space to a one dimensional space, which is remarkable. Assuming that this is indeed the case in general, a sufficiently late epoch will likely have a transitive percolation dynamics.

What can one say about the causal sets generated from this dynamics? A very important result from transitive percolation is that the typical causal sets obtained are not KR like posets and hence the dynamics beats their entropic dominance. The question of whether there is a continuum-like limit for transitive percolation dynamics was explored in Rideout and Sorkin (2001), using a comparison criterion. The abundance of fixed small subcausal sets was examined as a function of the coupling, by fixing the density relations. Comparisons with Poisson sprinklings in flat spacetime showed a convergence, suggestive of a continuum limit. In Ahmed and Rideout (2010), it was shown that the dynamics typically yields an exponentially expanding universe. Moreover, for $(1-q) \ll 1$ and $n \gg \frac{1}{1-q}$, after a post the universe enters a tree like phase and then a de Sitter-like phase, in which the cardinality of large causal diamonds are de Sitter like functions of the discrete proper time. In Glaser and Surya (2013), it was shown that despite this, the abundance of causal intervals is not de Sitter like, and thus, this is not strictly a manifold-like phase. In Brightwell and Georgiou (2010) and Brightwell and Luczak (2015), moreover, it was shown explicitly that in the asymptotic limit $n \to \infty$ the causal sets limit to "semi-orders" which, though temporally ordered, have no spatial structure at all, and are hence non-manifold-like. Nevertheless, the dominance of measure over entropy is important and the hope is that it will be reflected in the right quantum version of the dynamics.

Recently, Dowker and Zalel (2017) proposed a method for dealing with black hole singularities in CSG models. As in the case of cosmological bounces a new epoch is created beyond the singularity. Using "breaks" which are multi-element versions of a post, they demonstrated that a renormalisation of the coupling constants occurs in the new epoch.

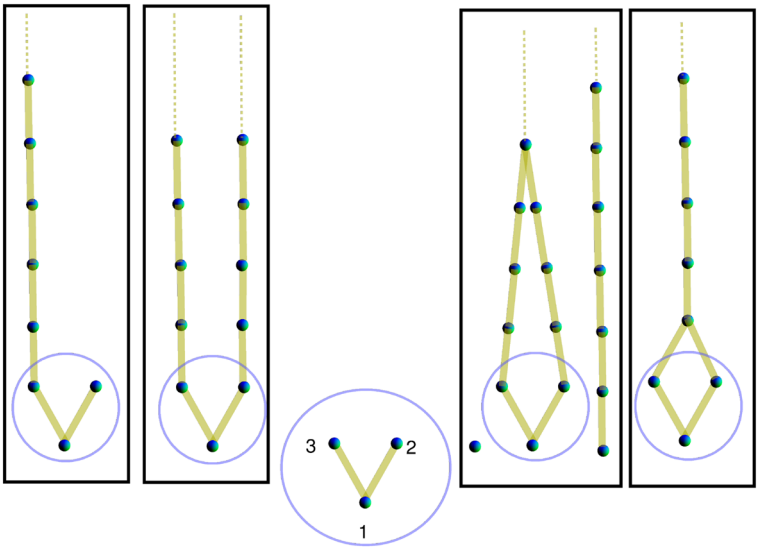

Fig. 22 The cylinder set for the "**V**" poset consists of all countable causal sets in $\tilde{\Omega}_g$ whose first three elements are the labelled "**V**" poset. Examples of causal sets that lie in cyl(**V**) are depicted in the boxes

6.2 Observables as beables

As mentioned in the introduction to this section, a dynamics for CST is given by the triple $(\Omega, \mathfrak{A}, \mu)$. In CSG this is a probability measure space, where the sample space $\tilde{\Omega}_g$ is the set of all past finite naturally labelled causal sets.

The event algebra $\mathfrak{A}$ can be constructed from the sequential growth process as follows. We define a *cylinder set* $\text{cyl}(\tilde{c}_n) \subset \tilde{\Omega}_g$ as the set of all labelled causal sets in $\tilde{\Omega}_g$ whose first n elements are the causal set $\tilde{c}_n$. Figure 22 depicts an example of a cylinder set.[26] For every finite element causal set $\tilde{c}_n$, $\text{cyl}(\tilde{c}_n) \subseteq \tilde{\Omega}_g$, and in the trivial $n = 1$ case, $\text{cyl}(\tilde{c}_1) = \tilde{\Omega}_g$. The cylinder sets in CSG satisfy a nesting property. Namely, if $n' > n$ and $\text{cyl}(\tilde{c}_{n'}) \cap \text{cyl}(\tilde{c}_n) \neq \emptyset$, then $\text{cyl}(\tilde{c}_{n'}) \subset \text{cyl}(\tilde{c}_n)$. Thus, a non-trivial intersection of two different cylinder sets is possible only if one is strictly a subset of the other.

The *event algebra* $\tilde{\mathfrak{A}}$ is generated from the cylinder sets via finite unions, intersections and set differences. It is closed under finite set operations and contains the null set $\emptyset$ as well as $\tilde{\Omega}_g$. In the growth process we assign a probability $\mu(\tilde{c}_n)$ to every finite labelled causal set $\tilde{c}_n$. By identifying $\tilde{c}_n$ with its cylinder set $\text{cyl}(\tilde{c}_n)$, we define the measure $\mu(\text{cyl}(\tilde{c}_n)) \equiv \mu(\tilde{c}_n)$ and hence on all elements of $\tilde{\mathfrak{A}}$, since μ is finitely additive. This makes $(\tilde{\Omega}_g, \tilde{\mathfrak{A}}, \tilde{\mu}')$ a "pre-measure" space.

An *event* α is an element of $\mathfrak{A}$ deemed to be *covariant* as a measurable subset $\alpha \subset \tilde{\Omega}_g$ if for every $\tilde{c} \in \alpha$, its relabelling $\tilde{c}'$ also belongs to α. Since a relabelling can

[26] A useful example to keep in mind is the 1-d random walk. Let γ^T be a finite element path in the $t - x$ plane from $t = 0$ to $t = T$. A cylinder set $\text{cyl}(\gamma^T)$ is then the set of all infinite time paths, which coincide with γ^T from $t = 0$ to $t = T$.

happen arbitrarily far into the future, no event in $\mathfrak{A}$ is covariant, since $\mathfrak{A}$ is closed only under finite set operations. Take for example the covariant *post event* which is the set of all causal sets which have a post. This is a covariant event, and is the equivalent of the return event in the random walk. In both cases, the event cannot be defined using only countable set operations, and hence the post event does not belong to $\mathfrak{A}$.

One route to obtaining covariant events is to pass to the full sigma algebra $\tilde{\mathfrak{S}}$ generated by $\tilde{\mathfrak{A}}$, which is closed under countable set operations. For classical measure spaces, the Kolmogorov–Caratheodory–Hahn extension theorem allows us to extend $\tilde{\mu}'$ to $\tilde{\mathfrak{S}}$ and hence pass with ease to a full measure space $(\tilde{\Omega}_g, \tilde{\mathfrak{S}}, \tilde{\mu})$, where $\tilde{\mu}|_{\tilde{\mathfrak{A}}} = \tilde{\mu}'$. Not every event in $\mathfrak{S}$ is covariant, but we can restrict our attention to covariant events, i.e., sets that are invariant under relabellings. If $\sim$ denotes the equivalence up to relabellings one can define the quotient algebra $\mathfrak{S} = \tilde{\mathfrak{S}}/\sim$ of covariant events. An element of $\mathfrak{S}$ is measurable covariant set, or a *covariant observable (or beable)*. Our example of the post event belongs to $\mathfrak{S}$. Another example of a covariant event is the set of *originary* causal sets, i.e., causal sets with a single initial element to the past of all other elements. Constructing more physically interesting covariant observables in $\mathfrak{S}$ is important, since it tells us what covariant questions we can ask of causal set quantum gravity.

A more covariant way to proceed is to generate the event algebra not via the cylinder sets in $\tilde{\Omega}_g$ but by using covariantly defined sets in Ω_g, the sample space of unlabelled causal sets. Because causal sets are past finite we can use the analogue of past sets $J^-(X)$ to characterise causal sets in a covariant way. A finite unlabelled sub-causal set c_n of $c \in \Omega_g$ is said to be a *partial stem* if it contains its own past. A *stem set* $\mathrm{stem}(c_n)$ is then a subset of Ω_g such that every $c \in \mathrm{stem}(c_n)$ contains the partial stem $\tilde{c_n}$. Let $\mathcal{S}$ be the sigma algebra generated by the stem sets. Although $\mathcal{S}$ is a strictly smaller subalgebra of $\mathfrak{S}$, it differs on sets of measure zero for the CSG and extended CSG models as shown by Brightwell et al. (2003) and Dowker and Surya (2006). Thus, one can characterise all the observables of CSG in terms of stem sets. This is a non-trivial result and the hope is that some version of it will carry over to the quantum case.

6.3 A route to quantisation: the quantum measure

The generalisation of CSG to QSG is, at least formally, very straightforward. One "quantises" the classical covariant probability space $(\Omega_g, \mathfrak{S}, \mu_c)$, by simply replacing the classical probability μ_c with a *quantum measure* $\mu : \mathfrak{S} \to \mathbb{R}^+$, where μ satisfies the quantum sum rule (Sorkin 1994, 1995; Salgado 2002; Sorkin 2007d)[27]

$$\mu(\alpha \cup \beta \cup \gamma) = \mu(\alpha \cup \beta) + \mu(\alpha \cup \gamma) + \mu(\beta \cup \gamma) - \mu(\alpha) - \mu(\beta) - \mu(\gamma), \quad (82)$$

for the mutually disjoint sets $\alpha, \beta, \gamma \in \mathfrak{S}$. $\mu(.)$ is not in general a probability measure since it does not satisfy additivity $\mu(\alpha \cup \beta) \neq \mu(\alpha) + \mu(\beta)$ for $\alpha \cap \beta = \emptyset$. As in the classical case, observables in this theory are simply the quantum measurable

[27] We will not discuss the very rich and interesting literature on the co-event interpretation of the quantum measure, which though incomplete, contains essential features that one would seek for a theory of quantum gravity (Sorkin 2007c).

sets in $\mathfrak{S}$. The quantum measure $\mu(.)$ can be obtained from a decoherence functional $D : \mathfrak{S} \times \mathfrak{S} \to \mathbb{C}$ of quantum theory with

$$\mu(\alpha) = D(\alpha, \alpha), \tag{83}$$

where D satisfies

- Hermiticity: $D(\alpha, \beta) = D^*(\beta, \alpha)$
- Countable biadditivity: $D(\alpha, \sqcup_i \beta_i) = \sum_i D(\alpha, \beta_i)$ and $D(\sqcup_i \alpha_i, \beta) = \sum_i D(\alpha_i, \beta)$
- Normalisation: $D(\Omega, \Omega) = 1$
- Strong positivity: $M_{ij} \equiv D(\alpha_i, \alpha_j)$ for any finite collection $\{\alpha_i\}$ is positive semi-definite

In a QSG model the transition probabilities of CSG are replaced by the decoherence functional D or quantum measure. Leaving aside Bell causality, the other principle of the growth dynamics are easy to implement. In Dowker et al. (2010c), a simple complex percolation dynamics was studied, given by a product decoherence function $\tilde{D}(\alpha, \beta) = A^*(\alpha)A(\beta)$ on $\mathfrak{A} \times \mathfrak{A}$, where $A(\alpha)$ is obtained from the transition amplitudes $q \in \mathbb{C}$, similar to transitive percolation. Thus, as in the case of CSG models, one starts with the labelled event algebra $\mathfrak{A}$ generated by the cylinder sets, and a quantum pre-measure $\tilde{D}'$. Again, in order to obtain covariant observables one has to pass to the full sigma algebra $\mathfrak{S}$ associated with $\mathfrak{A}$. However, unlike a classical measure $\tilde{D}$ need not extend to a full sigma algebra. In Dowker et al. (2010c), the quantum pre-measure was shown to be a *vector pre-measure* $\hat{\mu}'$ in the associated histories Hilbert space (Dowker et al. 2010b). Extension of $\hat{\mu}'$ to $\mathfrak{S}$ is then possible provided certain convergence conditions are satisfied.[28]

Although the vector measure is 1-dimensional in complex percolation dynamics, it was shown in Dowker et al. (2010c) not to satisfy this convergence condition and hence one cannot pass to $\mathfrak{S}$ to construct covariant observables. However a smaller algebra may be sufficient for answering physically interesting questions, which require far weaker convergence condition as suggested by Sorkin (2011b). This relaxation of conditions means that some simple measurable covariant observables can be constructed in complex percolation, including for the originary event (*Sorkin and Surya, work in progress*). Whether these results on extension are shared by all QSG models or not is of course an interesting question. Another possibility is that an extension of the measure in QSG could, for example, be a criterion for limiting the parameter space of QSG. Very recently a class of QSG dynamics that does admit an extension has been found (*Surya and Zalel, work in progress*).

The space of QSG models is largely unexplored. It is however critical to study it extensively in order to find the right CST quantum dynamics based on first principles.

[28] In general, these are given by the conditions in the Kolmogorov–Caratheodory–Hahn–Kluvanek theorem (Diestel and Uhl 1977).

6.4 A continuum-inspired dynamics

As we have seen, at a fundamental level the quantum dynamics of causal sets looks very different from that of a continuum theory of quantum gravity, even if the latter is formulated as a path integral. However, as one approaches the continuum approximation of the theory, it is possible that the effective quantum dynamics begins to resemble the continuum path integral. In CST, the quantum partition function is

$$Z_\Omega \equiv \sum_{c\in\Omega} e^{\frac{iS(c)}{\hbar}} \tag{84}$$

where $S(c)$ is an action for causal sets, and the choice of sample space Ω is determined by the problem at hand. One might also consider more generally a decoherence functional $D(c_1, c_2)$ on causal sets, inspired by the continuum, where $D(c_1, c_2) = e^{-i\frac{1}{\hbar}(S(c_1)-S(c_2))} f(c_1, c_2)$ with $f(c_1, c_2)$ a causal set analog of the delta function associated with unitarity quantum theories. This is currently an unexplored direction and we will not discuss it further in this work.

The natural choice for $S(c)$ is the d dimensional BD action $S^{(d)}_{\rm BD}(c)$ which limits to the Einstein–Hilbert action in the continuum. As discussed in Sect. 3, the sample space Ω_n of causal sets of cardinality n is dominated by KR type causal sets. An important question is whether the action $S^{(d)}_{\rm BD}(c)$ can overcome the KR entropy in the large n limit.

Indeed, there is a hierarchy of sub-dominant causal sets which are non manifold-like (Dhar 1978, 1980; Kleitman and Rothschild 1975; Promel et al. 2001), with the set of bilayer posets $\mathcal{B}$ being the next subdominant class. A recent calculation by Loomis and Carlip (2018) shows that $\mathcal{B}$ is suppressed by the BD action when the mesoscale and dimension satisfy certain conditions. The only relations in a bilayer poset are links. Given that the maximum number of relations is $\binom{n}{2}$ the causal sets in $\mathcal{B}$ can be classified by the linking fraction p given by the ratio of the total number of links N_0 to the maximal possible number of links $\binom{n}{2}$. Moreover, the action itself reduces to a simple sum over n and N_0. In the limit of large n, Loomis and Carlip (2018) consider p to be a continuous variable using which the partition function $Z_\mathcal{B}$ can be expressed as an integral over p

$$Z_\mathcal{B} = \int dp|\mathcal{B}_{p,n}|e^{iS(p)/\hbar} = e^{i\mu n}\int dp|\mathcal{B}_{p,n}|e^{\frac{1}{2}i\mu\lambda_0 pn^2+o(n^2)} \tag{85}$$

where $\mathcal{B}_{p,n}$ denotes the class of n-element causal sets in $\mathcal{B}$ with linking fraction p and μ, λ_0 are related to the mesoscale ϵ and function $f_d(n, \epsilon)$ that appears in $S^{(d)}_{\rm BD}(c)$. The challenge is then shifted to calculating $|\mathcal{B}_{p,n}|$. Using another parameter q which gives the cardinality of the upper layer as a further subclassification of $\mathcal{B}_{p,n}$, the leading order contribution to $|\mathcal{B}_{p,n}|$ was found. The resulting partition function was then shown to be strictly suppressed when $\mu\lambda_0$ satisfy the condition

$$\tan(-\mu\lambda_0/2) > \left(\frac{27}{4}e^{-\frac{1}{2}} - 1\right). \tag{86}$$

This is an important analytic calculation and paves the way for a more rigorous understanding of the CST partition function.

More than the partition function, however, it is the expectation value of observables or order invariants

$$\langle \mathcal{O} \rangle = \frac{1}{Z_c} \sum_{c \in \Omega} \mathcal{O}[c] e^{i \frac{1}{\hbar} S[c]} \tag{87}$$

that is of physical significance.[29] Evaluating this for larger values of n is a big challenge and we turn to numerical simulations to help us.

One route could be to simply "perform" the sum above. However, given that $|\Omega_n|$ grows superexponentially (to leading order it is $\sim 2^{\frac{n^2}{4}}$), this is computationally challenging even for relatively small values of n. On the other hand, Markov Chain Monte Carlo (MCMC) methods for sampling the space Ω can be used if we can convert Z_Ω into a statistical partition function.

In CST, there is no analogue of a Wick rotation: since the order relation derives from the causal structure, it cannot be "Euclideanised". On the other hand, there are other ways to analytically continue Z_Ω (see Louko and Sorkin 1997 for a continuum example). One option, first explored in Surya (2012) is to introduce a new parameter β such that

$$Z_{\Omega,\beta} \equiv \sum_{c \in \Omega} e^{i \frac{\beta}{\hbar} S(c)}. \tag{88}$$

This allows us to analytically continue $Z_{\Omega,\beta}$ from real to imaginary values of β, thus rendering the quantum partition function into a statistical partition function. We can then use standard tools in statistical physics, including MCMC methods, to find the expectation values of suitable observables (Surya 2012; Glaser and Surya 2016; Glaser et al. 2018; Glaser 2018; Cunningham and Surya 2019).

In Henson et al. (2017), MCMC methods were used to examine the sample space of naturally labelled posets $\tilde{\Omega}_n$ to determine the onset of the KR regime, using the uniform measure ($\beta = 0$). The Markov Chain was generated via a set of moves that sample Ω_n. A mixture of two moves, the *link move* and the *relation move*, was used to obtain the quickest thermalisation.

To illustrate the complexity of these moves we describe in detail the *link move*. A pair of elements e, e' are picked randomly and independently from the causal set c, and retained if $L(e) < L(e')$, where L is the natural labelling defined in Sect. 6.1. If $e \prec e'$ and moreover the relation is a link, then the move is to "unlink" them. Those relations implied by this link via transitivity also need to be removed. These are relations between elements in $\mathrm{IPast}(e)$ and those in $\mathrm{IFut}(e')$ which are "mediated" solely either by e or e'. On the other hand if e and e' are not related, then one adds in a link between e and e', *provided* that there are no existing links between elements in $\mathrm{IPast}(e)$ and $\mathrm{IFut}(e')$, after which the transitive closure is taken. In the relation move, although the existence or non-existence of a link from e to e' is also required, the move

[29] We leave out interpretational questions!

doesn't care about the sanctity of links, but is in other ways more restrictive. Thus, for both moves, picking of a pair of elements at random in c does not always lead to a possible move, let alone a probable one, and hence this MCMC model is slow to thermalise. Trying to find a more efficient move is however non-trivial precisely because of transitivity.

The simulations of Henson et al. (2017) suggest that the onset of the asymptotic KR regime occurs for n as small as $n \approx 90$. Ω_n is very large even for $n = 90$ ($\sim 2^{90^2}$!) and hence thermalisation becomes a problem very quickly. Recently, steps have been taken to incorporate the action ($\beta \neq 0$) into the measure, but again, because of thermalisation issues, the size of the posets are fairly small.

Instead of taking the full sample space, one can restrict Ω_n to causal sets that capture some gross features of a class of spacetimes. As discussed above, for large enough n, Ω_n contains causal sets that are approximated by spacetimes of arbitrary dimensions. It is thus of interest to restrict the sample space so that those causal sets that are manifold-like in the sample space are approximated only by spacetime regions of a given dimension. Such a restriction is typically hard to find, since it requires "tailoring" Ω using non-trivial order theoretic constraints determined by dimension estimators of the kind we have encountered in Sect. 4.

Somewhat fortuitously, this restriction is very natural in $d = 2$. Here, the sample space of "2-orders" $\Omega_{2\mathrm{d}}$ is one in which the continuum dimension and a particular order theoretic dimension coincide (Brightwell et al. 2008; El-Zahar and Sauer 1988; Winkler 1991). The latter is defined only for a certain class of posets, namely those obtained by the "intersection" of d totally ordered sets. For example, an n element *2-order* is the intersection of two linear orders $U = (u_1, u_2, \ldots u_n)$ and $V = (v_1, v_2, \ldots v_n)$ where each u_i and v_i are valued on a set S_n of n non-overlapping points in $\mathbb{R}$. U and V are therefore "totally ordered" by the relation $<$ in $\mathbb{R}$. Their intersection is the poset

$$U \cap V \equiv \{(u_i, v_i) \in U \times V | (u_i, v_i) \prec (u_j, v_j) \Leftrightarrow u_i < u_j \ \& \ v_i < v_j\}. \quad (89)$$

Similarly, one can define a *d-order* as the intersection of d linear orders. This is the *order theoretic dimension* referred to above.

For $d = 2$, the total orders U, V can be thought of as the set of light-cone coordinates of a causal set obtained from an embedding (not necessarily faithful) into a causal diamond in $\mathbb{M}^2$. Of special interest is the 2-order obtained from a Poisson sprinkling, an example of which is shown in Fig. 7. As shown in Brightwell et al. (2008) this is equivalent to choosing the entries of U and V from a fixed S_n at random and independently. Importantly, this *random order* dominates $\Omega_{2\mathrm{d}}$ in the large n limit as shown in El-Zahar and Sauer (1988) and Winkler (1991), and grows as $|\Omega_{2\mathrm{d}}| \sim n!/2$. Thus, unlike Ω_n, the sample space is dominated by manifold-like causal sets, though it also contains causal sets that are distinctly non-manifold-like. This makes it an ideal starting point to study the non-perturbative quantum dynamics of causal sets. Moreover, as shown in Brightwell et al. (2008), 2-orders also have trivial spatial homology in the sense of Major et al. (2007) (see Sect. 4) and hence $\Omega_{2\mathrm{d}}$ is the sample space of topologically trivial 2d causal set quantum gravity.

The continuum-inspired partition function for 2-orders or topologically trivial *2d CST* is

$$Z_{2\mathrm{d}}(\beta, n) = \sum_{c \in \Omega_{2\mathrm{d}}} \exp^{\frac{i}{\hbar} S_{2\mathrm{d}}(c,\epsilon)}, \tag{90}$$

where $S_{2\mathrm{d}}(c, \epsilon)$ is the BD action for $d = 2$ with the non-locality parameter $\epsilon = l_p^2/l_c^2 \in (0, 1]$ (see Eq. (36)). Taking $\beta \to i\beta$ allows one to obtain the expectation values of order invariants using MCMC techniques as was done by Surya (2012). The MCMC move in $\Omega_{2\mathrm{d}}$ is very straightforward, unlike that in Ω_n: a pair of elements is picked independently and at random in either U or V, and swapped. For example, if $u_i \leftrightarrow u_j$, then the elements (u_i, v_i) and (u_j, v_j) in $U \cap V$ are replaced by $(u'_i = u_j, v'_i = v_i)$ and $(u'_j = u_i, v'_j = v_j)$, hence changing the poset. *Every* move is possible, and hence one saves considerably on efficiency and thermalisation times.

Importantly, the MCMC simulations of Surya (2012) give rise to a phase transition from a continuum phase at low β to a non-manifold-like phase at high β. This is very similar to the disordered to ordered phase transition in an Ising model. The β^2 versus ϵ phase diagram moreover indicates that the continuum phase survives the analytic continuation for any value of ϵ.

It was recently demonstrated by Glaser et al. (2018) using finite size scaling arguments that that this is a first order phase transition. The analysis moreover suggests that the continuum phase corresponds to a spacetime with negative cosmological constant. This is an explicit example of a non-perturbative theory of quantum gravity in which the cosmological constant is generated via the dynamics.

This simple system also allows us to examine other physically relevant questions. Of particular interest is the Hartle–Hawking wave function using the no-boundary proposal. In 2*d* CST, this was constructed by Glaser and Surya (2016) using a natural no-boundary condition for causal sets, namely requiring the existence of an "initial" element e_0 to the past of all the other elements. $\psi_{\mathrm{HH}}(\mathcal{A}_f)$ is the wave function for a final antichain of cardinality $|\mathcal{A}_f|$, where one is summing over all causal sets that have an initial element e_0 and final boundary $\mathcal{A}_f$.

The MCMC simulations give the expectation value of the action $S_{2\mathrm{d}}$ from which the partition function can be calculated by numerical integration, up to normalisation. The normalisation itself was determined in Glaser and Surya (2016) using a combination of analytic and numerical calculations. The results of the extensive analysis was that the Hartle–Hawking wave function $\psi_{\mathrm{HH}}(\mathcal{A}_f)$ peaks at low β on antichains of small cardinality, with the peak jumping at higher β to antichains with cardinality $\sim n/2$. Interestingly, in the latter, high β (low temperature) phase, the dominant causal sets satisfy some of the rudimentary features of early universe cosmology: (a) the growth from a single element to a large antichain takes place rapidly and (b) each element in $\mathcal{A}_f$ is causally related to all the elements in its immediate past which makes $\mathcal{A}_f$ "homogeneous". However, this is a non manifold-like phase, and it is an open question how one exits this phase into a manifold-like phase. If there is a dynamical mechanism that makes β small, then this would be a promising new mechanism for generating cosmologically relevant initial conditions for the universe.

Will this analysis survive higher dimensions? One of the issues at hand is that even for 2-orders the cardinality of Ω_{2d} grows rapidly with n and hence thermalisation can become a major stumbling block. However, the finite sized scaling analysis of Glaser et al. (2018) and the techniques used therein, tell us that it suffices to be in the asymptotic regime. For 2-orders, this is already true around $n \sim 80$ and hence the results of Surya (2012) and Glaser and Surya (2016) are at least qualitatively robust. Nevertheless, to get to the asymptotic regime in $d = 4$ will require far more extensive computational power. Recently, using new sophisticated computational techniques (Cunningham 2018b), the algorithms of Surya (2012) have been updated, so that $n \sim 300$ simulations can be done in a reasonable time.

An important question, however is how to obtain a dimensionally restricted Ω_n more generally. While 2-orders are a good representation of 2d (topologically trivial) causal set quantum gravity, this is not true for higher order theoretic dimension. For $d > 2$ a d-order is an embedding into a space with "light-cubes" rather than lightcones. Though potentially interesting, this does not serve our more narrowly defined goal of obtaining a continuum-inspired dimensionally reduced sample space.

Recently, a lattice inspired method has been investigated to generate sample spaces which are both dimensionally and topologically restricted. These are obtained as embeddings (not necessarily faithful) into a fixed spacetime, and thus include manifold-like causal sets. In $d = 2$, the simplest example comes from causal sets obtained from sprinkling into the flat cylinder spacetime $ds^2 = -dt^2 + d\theta^2$, $\theta \in [0, 2\pi]$. Recently it has been demonstrated that topologically non-trivial lattice inspired models in $d = 2$ and $d = 3$ also exhibit phase transitions similar to 2-orders (Cunningham and Surya 2019). The next step is to include a wider class of embeddings as well as topology change into the model, and hence bring it closer to a full 2d theory of quantum gravity.

Of course, 2d causal set quantum gravity without matter does not have a continuum counterpart, since 2d continuum quantum gravity is coupled to a scalar field (for example, Liouville gravity). Studying 2d CST with matter is therefore an open interesting question. In Glaser (2018), Ising spins were coupled to the causal set by placing a spin $s_i = \pm 1$ at every element e_i and coupling spins along the links, i.e.,

$$S_I(j) \equiv j \sum_{ik} s_i s_k L_{ik} \,, \tag{91}$$

where L_{ik} is the link matrix and j the spin coupling constant. The phase structure of this model coupled to the BD action is substantially richer. In particular, the hope is that some of the resulting phase transitions are of higher order and hence comparisons with conformal field theories might be possible. Further analysis of this model would definitely be useful and interesting.

In the MCMC simulations discussed above, labelled posets are used for practical reasons, since this is how they are stored on the computer. A single unlabelled poset admits many relabellings or "automorphisms", but the number of relabellings varies from poset to poset even for the same cardinality. For example, in the list of coloured or labelled 3-element causal sets in Fig. 20, we see that there is only one 3-element causal set with three distinct natural labellings, while all the others admit only one

natural labelling. Enumerating the number of automorphisms for a given causal set quickly becomes very difficult as n increases.

In the continuum path integral, the "correct" measure in a gauge theory involves the volume of the gauge orbits. In this discrete setting, as we have discussed above, the analogous gauge orbits corresponding to to the automorphisms, are not of the same cardinality for each $c \in \tilde{\Omega}_n$.

Indeed, the choice of measure is not obvious in CST since it is not merely a discretisation of the continuum theory, with the path sum Eq. (84) including causal sets that are non-manifold-like. There is no underlying order theoretic reason to pick the specific BD action; we have done so, "inspired" by the continuum. For continuum like causal sets of a fixed dimension the number of relabellings is approximately the same, so that they appear roughly with the same weight in the path integral. However, it is the relative weight compared the non-continuum-like causal sets that depends on the relabellings. In the classical sequential growth model described above, the labelling is related to temporality and hence the choice of a uniform measure on the set of labelled causal sets $\tilde{\Omega}_g$ is a natural one. In the MCMC simulations, therefore we pick a measure that is uniform on $\tilde{\Omega}_n$, rather than on the unlabelled sample space Ω_n. Causal sets that admit more relabellings come with a higher natural weight than those that admit fewer relabellings. However, discrete covariance or label invariance is not compromised since the observables themselves are label independent.

While these numerical simulations have uncovered a wealth of information about the statistical thermodynamics of causal sets, one must pause to ask how this is related to the quantum dynamics, as $\beta \to -i\beta$. There is for example no analogue of the Osterwalder–Schrader theorems to protect the results we have obtained in the MCMC simulations. Pursuing these questions further is important, though finding definitive and rigorous answers is perhaps beyond the scope of our present understanding of CST.

7 Phenomenology

While the deep realm of quantum gravity is extremely well shielded from experimental probes in the foreseeable future, it is possible that certain properties of quantum gravity can "leak" into observationally accessible regimes. This is the reason for the push, in the last couple of decades, for exploring quantum gravity phenomenology. Without a full theory of quantum gravity, of course there is little hope that any phenomenology is entirely believable, since it requires assumptions about an incomplete theory. Nevertheless, quantum gravity phenomenology can be useful in setting realistic bounds on these leaked out properties, and hence constrain theories of quantum gravity, albeit weakly. Models of quantum gravity phenomenlogy moreover use distilled properties of the underlying theory to build reasonable models that can be tested. Some of these properties are unique to a given approach.

In CST spacetime discreteness takes a special form and brings with it a special type of non-locality that can affect observable physics. We have already encountered the possibility of voids in Sect. 3 as well as the propagation of scalar fields from distance sources in Sect. 5. The continuum approximation of CST is Lorentz invariant and consistent with stringent observational bounds as summarised in Liberati and Mattingly

(2016). In addition, as suggested by Dowker et al. (2004), there is the possibility of generating very high energies particles through long time diffusion in momentum space. This arises from the randomness of CST discreteness, which cause particles to "swerve", or suddenly change their momentum, as they traverse the causal set underlying our universe (Philpott et al. 2009; Contaldi et al. 2010). This spacetime Brownian motion was calculated in $\mathbb{M}^d$ and can be constrained by observations (Kaloper and Mattingly 2006), but an open question is how to extend the calculation to our FRW universe.

There have been some very interesting recent ideas by Belenchia et al. (2016b) for testing CST type non-locality via its effect on propagation in the continuum using the d'Alembertian operator. Belenchia et al. (2015) have looked at the associated quantum field theory which contain critical instabilities. These can be removed by modifying the d'Alembertian, but the relationship to CST is unclear. Saravani and Afshordi (2017) have proposed a candidate for dark matter as off-shell modes of the non-local CST d'Alembertian. This is an exciting proposal and should be investigated in more detail.

We will not review these very interesting ideas on CST phenomenology here, except one, namely the prediction of Λ.

7.1 The 1987 prediction for Λ

One of the most outstanding questions in theoretical physics is understanding the origin of "dark energy" which observationally has been seen to make up $\sim 70\%$ of the total energy of the universe. The current observational value is $\sim 2.888 \times 10^{-122}$ in Planck units. Quantum field theory predictions for dark energy interpreted as the energy of vacuum fluctuations of quantum fields on the other hand gives a huge value, perhaps as large as ~ 1 in Planck units. The gross conflict with observation obviously implies that this cannot be the source of Λ.[30]

In light of this conundrum, the CST prediction for Λ due to Sorkin (1991) is startling in its simplicity and accuracy, especially since it was made several years before the 1998 observation. One begins with the framework of unimodular gravity (Sorkin 1997; Unruh and Wald 1989) in which the spacetime volume element is fixed. Λ then appears as a Lagrange multiplier in the action, with $\Lambda \int dV = \Lambda V = \text{constant}$, for any finite spacetime region of volume V. In a canonical formulation of the theory, therefore Λ and V are conjugate to each other, so that on quantisation there is an uncertainty relation

$$\Delta V \Delta \Lambda \sim 1. \tag{92}$$

Using the fact that ΔV is generated from Poisson fluctuations of the underlying causal set ensemble

$$\Delta V \sim \sqrt{V}. \tag{93}$$

[30] On the other hand, it would be interesting to understand *why* the back of the envelope quantum field theory calculation is not observationally relevant. Interesting insights into this question could come from a better understanding of the SJ vacuum in de Sitter spacetime.

Assuming $\langle \Lambda \rangle = 0$, moreover, we see that

$$\Lambda \sim \frac{1}{\sqrt{V}} \sim H^2 = \frac{1}{3}\rho_{\text{critical}} \tag{94}$$

where H is the Hubble constant. If V is taken to be the volume of the visible universe,

$$\Lambda = \Delta\Lambda \sim 10^{-120}, \tag{95}$$

in Planck units. This is very close to the subsequently observed value of Λ! Importantly, the prediction also states that that under these assumptions, Λ always tracks the critical density and is hence "everpresent".

This argument is general and requires three important ingredients: (i) the assumption of unimodularity and hence the conjugacy between Λ and V, (ii) the number to volume correspondence $V \sim n$ and (iii) that there *are* fluctuations in V which are Poisson, with $\delta V = \sqrt{V} \sim \sqrt{n}$. While (i) can be motivated by a wide range of theories of quantum gravity, (ii) and (iii) are both *distinctive* to causal set theory. No other discrete approach to quantum gravity makes the $n \sim V$ correspondence at a fundamental level and also incorporates Poisson fluctuations kinematically in the continuum approximation. Quoting from Sorkin (1991), "Fluctuations in Λ arise as residual nonlocal quantum effects of spacetime discreteness". Interestingly, as shown by Sorkin (2005a), if spacetime admits large extra directions, then the contribution to V is very different and gives the wrong answer for $\Delta\Lambda$.

Of course, an important question that arises in this quick calculation is why we should assume that $\langle \Lambda \rangle = 0$.[31] The answer to this may well lie in the full and as yet unknown quantum dynamics. Nevertheless, phenomenologically this assumption leads to further predictions that can already be tested. The first conclusion is that a fluctuating Λ must violate conservation of the stress energy tensor, and hence the Einstein field equations.

In Ahmed et al. (2004), a dynamical model for generating fluctuations of Λ was constructed, starting with the flat $k = 0$ FRW spacetime. In order to accommodate a fluctuating Λ, one of the two Friedmann equations must be dropped. In Ahmed et al. (2004), the Friedmann equation

$$3\left(\frac{\dot{a}}{a}\right)^2 = \rho + \rho_\Lambda \tag{96}$$

was retained,[32] with

$$\rho_\Lambda = \Lambda, \qquad p_\Lambda = -\Lambda - \dot{\Lambda}/3H, \tag{97}$$

[31] In Samuel and Sinha (2006), a very striking analogy was made between a fluctuating Λ and the surface tension T of a fluid membrane. In addition, using the atomicity of the model, the mean value of T was shown to be zero, with a suggestion of how this might extend to CST.

[32] Subsequently, more general "mixed equation" models were examined in Ahmed and Sorkin (2013), which indicate that the results of Ahmed et al. (2004) are robust to these modifications.

and Λ modelled as a stochastic function of V, such that

$$\Delta\Lambda \sim \frac{1}{\sqrt{V}}. \tag{98}$$

More generally, Λ can be thought of as the action S per unit volume, which for causal sets means that $\Lambda \sim S/V$. A very simple stochastic dynamics is then generated by assuming that every element contributes $\pm\hbar$ to S, so that

$$S = \sum_{\text{elements}} \pm\hbar \Rightarrow S/\hbar \sim \pm\sqrt{N} \sim \pm\sqrt{V/l_p^4} \Rightarrow \Lambda \sim \pm\frac{\hbar/l_p^2}{\sqrt{V}}, \tag{99}$$

where we have equated the discreteness scale l_c with the Planck length l_p. One then gets the integro-differential equations

$$\frac{da}{a} = \sqrt{\frac{\rho+\Lambda}{3}}d\tau$$
$$Vd\Lambda = Vd(S/V) = dS - \Lambda\dot{V}d\tau,$$

where

$$V(\tau) = \frac{4\pi}{3}\int_0^t dt' a(t')^3\left(\int_0^{t'} dt''\frac{1}{a(t'')}\right)^3 \tag{100}$$

is the volume of the entire causal past of an event in the FRW spacetime. The stochastic equation is then generated as follows. At the i^{th} step one has the variables a_i (scale factor), N_i, V_i, S_i and Λ_i. The scale factor is updated using the discrete Friedmann equation $a_{i+1} = a_i + a_i\sqrt{\frac{\rho+\Lambda}{3}}(\tau_{i+1}-\tau_i)$, from which $V_i = V(\tau)$ can be calculated and thence $N_{i+1} = V_{i+1}/\ell^4$. The action is then updated via $S_{i+1} = S_i + \alpha\ \xi\sqrt{N_{i+1}-N_i}$, where ξ is a Gaussian random variable, with $\Delta\xi = 1$, and α is a tunable free parameter which controls the magnitude of the fluctuations. Finally, $\Lambda_{i+1} = S_{i+1}/V_{i+1}$, with $S_0 = 0$. It was shown in Ahmed et al. (2004) that in order to be consistent with astrophysical observations, $0.01 < \alpha < 0.02$. The results of simulations moreover suggest that Λ is *"everpresent"* and tracks the energy density of the universe.

This model assumes spatial homogeneity and it is important to check how inhomogeneities affect these results. In Barrow (2007) and Zuntz (2008), inhomogeneities were modelled by taking $\Lambda(x^\mu)$, such that $\Delta\Lambda(x)$ is dependent only on $\Lambda(y)$ for $y \in J^-(x)$. This would mean that well separated patches in the CMB sky would contain uncorrelated fluctuations in Ω_Λ, which in turn are strongly constrained to $< 10^{-6}$ by observations and hence insufficient to account for Λ. In Ahmed et al. (2004) and Zwane et al. (2018), it was suggested that quantum Bell correlations may be a possible way to induce correlations in the CMB sky. However, incorporating inhomogeneities into the dynamics in a systematic way remains an important open question.

In Zwane et al. (2018), a phenomenological model was adopted which uses the homogeneous temporal fluctuations in Λ to model a quintessence type spatially inhomogeneous scalar field with a potential term that varies from realisation to realisation. Using MCMC methods to sample the cosmological parameter space, and generate different stochastic realisations, it was shown that these CST inspired models agrees with the observations as well as ΛCDM models and in fact does better for the Baryonic Acoustic Oscillations (BAO) measurements. The very extensive and detailed analysis of Zwane et al. (2018) sets the stage for direct comparisons with future observations and heralds an exciting phase of quantum gravity phenomenology.

8 Outlook

CST has come a long way in the last three decades, despite the fact that there are only a few practitioners who have been able to dedicate their time to it. Over the last decade, in particular, there has been a growth of interest with inputs from the wider quantum gravity community. This is heartening, since an extensive exploration of the theory is required in order to make significant progress. It is our hope that this review will spark the interest of the larger quantum gravity community, and continue what has been a productive dialogue.

We have in this review touched upon several open questions, many of which are challenging but some of which are straightforward to carry out. We will not summarise these but just pick two that are of utmost importance. One is the the pursuit of CST-inspired inhomogeneous models of fluctuating Λ which can be tested against the most recent observations. The second, on the other side of the quantum gravity spectrum, is the construction from first principles of a viable quantum dynamics for causal sets. Between these two ends lie myriad interesting questions. We invite you to join us.

Acknowledgements I am indebted to Rafael Sorkin for his deep insights and vast knowledge, that have directly and indirectly shaped this review. I am also deeply indebted to Fay Dowker for our interactions and collaborations over the past 25 years, which have helped enrich my understanding of quantum gravity. I am grateful to my other collaborators, including David Rideout, Joe Henson, Graham Brightwell, Petros Wallden, Lisa Glaser, Denjoe O'Connor, Ian Jubb, Yasaman Yazdi and my students Nomaan X and Abhishek Mathur, for their active and continuous engagement with the questions in CST, which have led to fruitful discussions, arguments, disagreements and debates over the years. Finally, I would like to thank Yasaman Yazdi and Stav Zalel for a careful reading through the first draft of the manuscript and giving me useful feedback. This research was supported in part by the Emmy Noether Fellowship (2017–2018) and also by a Visiting Fellowship (2019–2022) at the Perimeter Institute of Theoretical Physics.

A Notation and terminology

We list some of the more widely used definitions as well as the abbreviations used in the paper.

Definitions

Relation: $e, e' \in C$ are said to be *related* if $e \prec e'$ or $e \succ e'$.

Link: $e \prec e' \in C$ is said to be a *link* if $\nexists\, e'' \in C$ such that $e'' \neq e, e'$ and $e \prec e'' \prec e'$.

Hasse diagram: In a *Hasse diagram*, only the nearest neighbour relations or links are depicted with the remaining relations following from transitivity (see Fig. 4).

Valency: The *valency* $v(e)$ of an element e in a causal set C is the set of elements in C that are linked to e.

Order Interval: The *order interval* between the pair $e_i, e_j \in C$ is the set $\mathbf{I}[e_i, e_j] \equiv \text{Fut}(e_i) \cap \text{Past}(e_j)$ where $\text{Fut}(x), \text{Past}(x)$ are the exclusive future and past of x.

Labelling: A *labelling* of the causal set C of cardinality n is an injective map $L : C \to \mathbb{N}$, where $\mathbb{N}$ is the set of natural numbers.

Natural Labelling: A labelling $L : C \to \mathbb{N}$ is called *natural* if $e_i \prec e_j \Rightarrow L(e_i) < L(e_j)$.

Total Order: A poset C is *totally ordered* if for each pair $e_i, e_j \in C$, either $e_i \prec e_j$ or $e_j \prec e_i$.

Chain: A k-element set C is called a *chain* (or k-*chain*) if it is a totally ordered set, i.e., for every $e_i, e_j \in C$ either $e_i \prec e_j$ or $e_j \prec e_i$.

Length of a chain: The *length of a* k-*chain* is $k - 2$.

Antichain: A causal set C is an *antichain* if no two elements are related to each other.

Inextendible Antichain: A subset $\mathcal{A} \subseteq C$ is an *inextendible antichain* in C if it is an antichain and for every element $e \in C \backslash \mathcal{A}$ (where $\backslash$ is set difference) either $e \in \text{Past}(\mathcal{A})$ or $e \in \text{Fut}(\mathcal{A})$ (see Eq. (3)).

Order Invariant: $\mathcal{O} :\to \mathbb{R}$ is an *order invariant* if it is independent of the labelling of the causal set C. It is possible to replace $\mathbb{R}$ by a more general field, but since this has not been explicitly used here, the above definition is sufficient.

Manifold-like: A causal set C is said to be manifold-like if C has a continuum approximation.

Alexandrov interval: This is the generalised causal diamond in (M, g), $\mathbf{A}[p, q] \equiv I^+(p) \cap I^-(q)$, $p, q \in M$.

Sample Space Ω: This is a collection or space of causal sets.

Non-locality parameter: $\epsilon \equiv \rho_\kappa / \rho_c$ appears in the BD action.

Abbreviations in alphabetical order

BD action: Benincasa–Dowker action (see Sect. 4.5).

BLMS: Bombelli, Lee, Meyer and Sorkin's CST proposal (Bombelli et al. 1987).

CSG: Classical sequential growth dynamics (see Sect. 6.1).

CST: Causal set theory.

GHY: Gibbons–Hawking–York (see Sect. 4.6).

GNN: Gaussian normal neighbourhood.
HKMM theorem: Hawking–King–McCarthy–Malament theorem (see Sect. 2).
KR posets: Kleitman–Rothschild posets (see Sect. 3.1).
MCMC: Markov Chain Monte Carlo (see Sect. 6.4).
QSG: Quantum sequential growth dynamics (see Sect. 6.3).
RNN: Riemann normal neighbourhood.
SJ vacuum: Sorkin–Johnston vacuum (see Sect. 5.2).
SSEE: Sorkin spacetime entanglement entropy (see Sect. 5.3).

References

Abajian J, Carlip S (2018) Dimensional reduction in manifoldlike causal sets. Phys Rev D 97:066007. https://doi.org/10.1103/PhysRevD.97.066007. arXiv:1710.00938

Afshordi N, Buck M, Dowker F, Rideout D, Sorkin RD, Yazdi YK (2012) A ground state for the causal diamond in 2 dimensions. JHEP 10:088. https://doi.org/10.1007/JHEP10(2012)088. arXiv:1207.7101

Aghili M, Bombelli L, Pilgrim BB (2019) Discrete spacetime: a web of chains. Class Quantum Grav 36:185015. https://doi.org/10.1088/1361-6382/ab30e3. arXiv:1807.08701

Ahmed M, Rideout D (2010) Indications of de Sitter spacetime from classical sequential growth dynamics of causal sets. Phys Rev D 81:083528. https://doi.org/10.1103/PhysRevD.81.083528. arXiv:0909.4771

Ahmed M, Sorkin R (2013) Everpresent Λ. II. Structural stability. Phys Rev D 87:063515. https://doi.org/10.1103/PhysRevD.87.063515. arXiv:1210.2589

Ahmed M, Dodelson S, Greene PB, Sorkin R (2004) Everpresent Λ. Phys Rev D 69:103523. https://doi.org/10.1103/PhysRevD.69.103523

Ashtekar A, Pullin J (2017) Applications. In: Ashtekar A, Pullin J (eds) Loop quantum gravity: the first 30 years. World Scientific, Singapore, p 181. https://doi.org/10.1142/9789813220003_others03

Aslanbeigi S, Saravani M, Sorkin RD (2014) Generalized causal set d'Alembertians. JHEP 06:024. https://doi.org/10.1007/JHEP06(2014)024. arXiv:1403.1622

Bachmat E (2007) Discrete spacetime and its applications. arXiv e-prints arXiv:gr-qc/0702140

Barrow JD (2007) A strong constraint on ever-present lambda. Phys Rev D 75:067301. https://doi.org/10.1103/PhysRevD.75.067301. arXiv:gr-qc/0612128

Beem J, Ehrlich P, Easley K (1996) Global Lorentzian geometry. Marcel Dekker, New York

Belenchia A, Benincasa DMT, Liberati S (2015) Nonlocal scalar quantum field theory from causal sets. JHEP 03:036. https://doi.org/10.1007/JHEP03(2015)036. arXiv:1411.6513

Belenchia A, Benincasa DMT, Dowker F (2016a) The continuum limit of a 4-dimensional causal set scalar d'Alembertian. Class Quantum Grav 33:245018. https://doi.org/10.1088/0264-9381/33/24/245018. arXiv:1510.04656

Belenchia A, Benincasa DMT, Liberati S, Marin F, Marino F, Ortolan A (2016b) Testing quantum gravity induced nonlocality via optomechanical quantum oscillators. Phys Rev Lett 116:161303. https://doi.org/10.1103/PhysRevLett.116.161303. arXiv:1512.02083

Belenchia A, Benincasa DMT, Marciano A, Modesto L (2016c) Spectral dimension from nonlocal dynamics on causal sets. Phys Rev D 93:044017. https://doi.org/10.1103/PhysRevD.93.044017. arXiv:1507.00330

Bell JL, Korté H (2016) Hermann Weyl. In: Zalta EN (ed) The Stanford encyclopedia of philosophy, winter 2016 edn. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2016/entries/weyl/

Benincasa DM, Dowker F (2010) The scalar curvature of a causal set. Phys Rev Lett 104:181301. https://doi.org/10.1103/PhysRevLett.104.181301

Benincasa DM, Dowker F, Schmitzer B (2011) The random discrete action for 2-dimensional spacetime. Class Quantum Grav 28:105018. https://doi.org/10.1088/0264-9381/28/10/105018

Benincasa DMT (2013) The action of a casual set. PhD thesis, Imperial College, London http://hdl.handle.net/10044/1/14170
Bolognesi T, Lamb A (2016) Simple indicators for Lorentzian causets. Class Quantum Grav 33:185004. https://doi.org/10.1088/0264-9381/33/18/185004. arXiv:1407.1649
Bombelli L (1987) Space-time as a causal set. PhD thesis, Syracuse University
Bombelli L (2000) Statistical Lorentzian geometry and the closeness of Lorentzian manifolds. J Math Phys 41:6944–6958. https://doi.org/10.1063/1.1288494. arXiv:gr-qc/0002053
Bombelli L, Meyer DA (1989) The origin of Lorentzian geometry. Phys Lett A 141:226–228. https://doi.org/10.1016/0375-9601(89)90474-X
Bombelli L, Noldus J (2004) The moduli space of isometry classes of globally hyperbolic space-times. Class Quantum Grav 21:4429–4454. https://doi.org/10.1088/0264-9381/21/18/010. arXiv:gr-qc/0402049
Bombelli L, Koul RK, Lee J, Sorkin RD (1986) A quantum source of entropy for black holes. Phys Rev D 34:373–383. https://doi.org/10.1103/PhysRevD.34.373
Bombelli L, Lee J, Meyer D, Sorkin R (1987) Space-time as a causal set. Phys Rev Lett 59:521–524. https://doi.org/10.1103/PhysRevLett.59.521
Bombelli L, Henson J, Sorkin RD (2009) Discreteness without symmetry breaking: a theorem. Mod Phys Lett A24:2579–2587. https://doi.org/10.1142/S0217732309031958
Bombelli L, Noldus J, Tafoya J (2012) Lorentzian manifolds and causal sets as partially ordered measure spaces. arXiv e-prints arXiv:1212.0601
Brightwell G, Georgiou N (2010) Continuum limits for classical sequential growth models. Rand Struct Alg 36:218–250. https://doi.org/10.1002/rsa.20278
Brightwell G, Gregory R (1991) The structure of random discrete space-time. Phys Rev Lett 66:260–263. https://doi.org/10.1103/PhysRevLett.66.260
Brightwell G, Luczak M (2011) Order-invariant measures on causal sets. Ann Appl Probab 21:1493–1536. https://doi.org/10.1214/10-AAP736. arXiv:0901.0240
Brightwell G, Luczak M (2012) Order-invariant measures on fixed causal sets. Comb Prob Comput 21:330–357. https://doi.org/10.1017/S0963548311000721. arXiv:0901.0242
Brightwell G, Luczak M (2015) The mathematics of causal sets. arXiv e-prints arXiv:1510.05612
Brightwell G, Dowker HF, Garcia RS, Henson J, Sorkin RD (2003) Observables in causal set cosmology. Phys Rev D 67:084031. https://doi.org/10.1103/PhysRevD.67.084031
Brightwell G, Henson J, Surya S (2008) A 2D model of causal set quantum gravity: the emergence of the continuum. Class Quantum Grav 25:105025. https://doi.org/10.1088/0264-9381/25/10/105025
Brum M, Fredenhagen K (2014) 'Vacuum-like' Hadamard states for quantum fields on curved spacetimes. Class Quantum Grav 31:025024. https://doi.org/10.1088/0264-9381/31/2/025024. arXiv:1307.0482
Buck M, Dowker F, Jubb I, Surya S (2015) Boundary terms for causal sets. Class Quantum Grav 32:205004. https://doi.org/10.1088/0264-9381/32/20/205004. arXiv:1502.05388
Buck M, Dowker F, Jubb I, Sorkin R (2017) The Sorkin–Johnston state in a patch of the trousers spacetime. Class Quantum Grav 34:055002. https://doi.org/10.1088/1361-6382/aa589c. arXiv:1609.03573
Carlip S (2017) Dimension and dimensional reduction in quantum gravity. Class Quantum Grav 34:193001. https://doi.org/10.1088/1361-6382/aa8535. arXiv:1705.05417
Christ NH, Friedberg R, Lee TD (1982) Random lattice field theory: general formulation. Nucl Phys B 202:89. https://doi.org/10.1016/0550-3213(82)90222-X
Contaldi CR, Dowker F, Philpott L (2010) Polarization diffusion from spacetime uncertainty. Class Quantum Grav 27:172001. https://doi.org/10.1088/0264-9381/27/17/172001. arXiv:1001.4545
Cortês M, Smolin L (2014) Quantum energetic causal sets. Phys Rev D 90:044035. https://doi.org/10.1103/PhysRevD.90.044035. arXiv:1308.2206
Cunningham W (2018a) Inference of boundaries in causal sets. Class Quantum Grav 35:094002. https://doi.org/10.1088/1361-6382/aaadc4. arXiv:1710.09705
Cunningham WJ (2018b) High performance algorithms for quantum gravity and cosmology. PhD thesis, Northeastern University arXiv:1805.04463
Cunningham, WJ, Surya S (2019) Dimensionally restricted causal set quantum gravity: examples in two and three dimensions. arXiv:1908.11647
Daughton AR (1993) The recovery of locality for causal sets and related topics. PhD thesis, Syracuse University
Dhar D (1978) Entropy and phase transitions in partially ordered sets. J Math Phys 19:1711. https://doi.org/10.1063/1.523869
Dhar D (1980) Asymptotic enumeration of partially ordered sets. Pacific J Math 90:299–305

Diestel J, Uhl JJ Jr (1977) Vector measures, Mathematical Surveys and Monographs, vol 15. American Mathematical Society, Providence, RI. https://doi.org/10.1090/surv/015
Dou D, Sorkin RD (2003) Black hole entropy as causal links. Found Phys 33:279–296. https://doi.org/10.1023/A:1023781022519. arXiv:gr-qc/0302009
Dowker F (2005) Causal sets and the deep structure of spacetime. In: Ashtekar A (ed) 100 Years of relativity: space-time structure: Einstein and beyond. World Scientific, Singapore, pp 445–464. https://doi.org/10.1142/9789812700988_0016. arXiv:gr-qc/0508109
Dowker F, Ghazi-Tabatabai Y (2008) The Kochen–Specker theorem revisited in quantum measure theory. J Phys A 41:105301. https://doi.org/10.1088/1751-8113/41/10/105301. arXiv:0711.0894
Dowker F, Glaser L (2013) Causal set d'Alembertians for various dimensions. Class Quantum Grav 30:195016
Dowker F, Surya S (2006) Observables in extend percolation models of causal set cosmology. Class Quantum Grav 23:1381–1390. https://doi.org/10.1088/0264-9381/23/4/018. arXiv:gr-qc/0504069
Dowker F, Zalel S (2017) Evolution of universes in causal set cosmology. C R Physique 18:246–253. https://doi.org/10.1016/j.crhy.2017.03.002. arXiv:1703.07556
Dowker F, Henson J, Sorkin RD (2004) Quantum gravity phenomenology, Lorentz invariance and discreteness. Mod Phys Lett A 19:1829–1840. https://doi.org/10.1142/S0217732304015026. arXiv:gr-qc/0311055
Dowker F, Henson J, Sorkin R (2010a) Discreteness and the transmission of light from distant sources. Phys Rev D 82:104048. https://doi.org/10.1103/PhysRevD.82.104048. arXiv:1009.3058
Dowker F, Johnston S, Sorkin RD (2010b) Hilbert spaces from path integrals. J Phys A 43:275302. https://doi.org/10.1088/1751-8113/43/27/275302. arXiv:1002.0589
Dowker F, Johnston S, Surya S (2010c) On extending the quantum measure. J Phys A 43:505305. https://doi.org/10.1088/1751-8113/43/50/505305. arXiv:1007.2725
Dowker F, Surya S, Nomaan X (2017) Scalar field Green functions on causal sets. Class Quantum Grav 34:124002. https://doi.org/10.1088/1361-6382/aa6bc7. arXiv:1701.07212
Eichhorn A (2018) Towards coarse graining of discrete Lorentzian quantum gravity. Class Quantum Grav 35:044001. https://doi.org/10.1088/1361-6382/aaa0a3. arXiv:1709.10419
Eichhorn A, Mizera S (2014) Spectral dimension in causal set quantum gravity. Class Quantum Grav 31:125007. https://doi.org/10.1088/0264-9381/31/12/125007. arXiv:1311.2530
Eichhorn A, Mizera S, Surya S (2017) Echoes of asymptotic silence in causal set quantum gravity. Class Quantum Grav 34(16):16LT01. https://doi.org/10.1088/1361-6382/aa7d1b. arXiv:1703.08454
Eichhorn A, Surya S, Versteegen F (2019a) Induced spatial geometry from causal structure. Class Quantum Grav 36:105005. https://doi.org/10.1088/1361-6382/ab114b. arXiv:1809.06192
Eichhorn A, Surya S, Versteegen F (2019b) Spectral dimension on spatial hypersurfaces in causal set quantum gravity. arXiv e-prints arXiv:1905.13498
El-Zahar MH, Sauer NW (1988) Asymptotic enumeration of two-dimensional posets. Order 5:239. https://doi.org/10.1007/BF00354891
Fewster CJ (2018) The art of the state. Int J Mod Phys D 27:1843007. https://doi.org/10.1142/S0218271818430071. arXiv:1803.06836
Fewster CJ, Verch R (2012) On a recent construction of 'vacuum-like' quantum field states in curved spacetime. Class Quantum Grav 29:205017. https://doi.org/10.1088/0264-9381/29/20/205017. arXiv:1206.1562
Feynman RP (1944) The character of physical law. Modern Library, New York
Finkelstein D (1969) Space-time code. Phys Rev 184:1261–1271. https://doi.org/10.1103/PhysRev.184.1261
Gibbons GW, Solodukhin SN (2007) The geometry of small causal diamonds. Phys Lett B 649:317–324. https://doi.org/10.1016/j.physletb.2007.03.068. arXiv:hep-th/0703098
Glaser L (2014) A closed form expression for the causal set d'Alembertian. Class Quantum Grav 31:095007
Glaser L (2018) The Ising model coupled to 2d orders. Class Quantum Grav 35:084001. https://doi.org/10.1088/1361-6382/aab139. arXiv:1802.02519
Glaser L, Surya S (2013) Towards a definition of locality in a manifoldlike causal set. Phys Rev D 88:124026. https://doi.org/10.1103/PhysRevD.88.124026. arXiv:1309.3403
Glaser L, Surya S (2016) The Hartle–Hawking wave function in 2D causal set quantum gravity. Class Quantum Grav 33:065003. https://doi.org/10.1088/0264-9381/33/6/065003. arXiv:1410.8775
Glaser L, O'Connor D, Surya S (2018) Finite size scaling in 2d causal set quantum gravity. Class Quantum Grav 35:045006. https://doi.org/10.1088/1361-6382/aa9540. arXiv:1706.06432

Greene BR, Plesser MR (1991) Mirror manifolds: a brief review and progress report. In: 2nd International symposium on particles, strings and cosmology (PASCOS 1991) Boston, MA, 25–30 March 1991, pp 648–666 arXiv:hep-th/9110014

Hawking S, Ellis G (1973) Large scale structure of space-time. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511524646

Hawking S, King A, McCarthy P (1976) A new topology for curved space-time which incorporates the causal, differential, and conformal structures. J Math Phys 17:174–181. https://doi.org/10.1063/1.522874

Hemion G (1988) A quantum theory of space and time. Int J Theor Phys 27:1145

Henson J (2005) Comparing causality principles. Stud Hist Phil Sci B 36:519–543. https://doi.org/10.1016/j.shpsb.2005.04.003. arXiv:quant-ph/0410051

Henson J (2006a) Constructing an interval of Minkowski space from a causal set. Class Quantum Grav 23:L29–L35. https://doi.org/10.1088/0264-9381/23/4/L02. arXiv:gr-qc/0601069

Henson J (2006b) The causal set approach to quantum gravity. In: Oriti D (ed) Approaches to quantum gravity. Cambridge University Press, Cambridge, pp 393–413 arXiv:gr-qc/0601121

Henson J (2010) Discovering the discrete universe. In: Proceedings, foundations of space and time: reflections on quantum gravity: Cape Town, South Africa arXiv:1003.5890

Henson J (2011) Causality, Bell's theorem, and ontic definiteness. arXiv e-prints arXiv:1102.2855

Henson J, Rideout D, Sorkin RD, Surya S (2017) Onset of the asymptotic regime for (uniformly random) finite orders. Exper Math 26:253–266. https://doi.org/10.1080/10586458.2016.1158134

Johnston S (2008) Particle propagators on discrete spacetime. Class Quantum Grav 25:202001. https://doi.org/10.1088/0264-9381/25/20/202001. arXiv:0806.3083

Johnston S (2009) Feynman propagator for a free scalar field on a causal set. Phys Rev Lett 103:180401. https://doi.org/10.1103/PhysRevLett.103.180401. arXiv:0909.0944

Johnston SP (2010) Quantum fields on causal sets. PhD thesis, Imperial College London arXiv:1010.5514

Jubb I (2017) The geometry of small causal cones. Class Quantum Grav 34:094005. https://doi.org/10.1088/1361-6382/aa68b7. arXiv:1611.00785

Jubb I, Samuel J, Sorkin R, Surya S (2017) Boundary and corner terms in the action for general relativity. Class Quantum Grav 34:065006. https://doi.org/10.1088/1361-6382/aa6014. arXiv:1612.00149

Kaloper N, Mattingly D (2006) Low energy bounds on Poincaré violation in causal set theory. Phys Rev D 74:106001. https://doi.org/10.1103/PhysRevD.74.106001. arXiv:astro-ph/0607485

Khetrapal S, Surya S (2013) Boundary term contribution to the volume of a small causal diamond. Class Quantum Grav 30:065005. https://doi.org/10.1088/0264-9381/30/6/065005. arXiv:1212.0629

Kleitman DJ, Rothschild BL (1975) Asymptotic enumeration of partial orders on a finite set. Trans Amer Math Soc 205:205–220

Kronheimer E, Penrose R (1967) On the structure of causal spaces. Proc Camb Phil Soc 63:481

Lehner L, Myers RC, Poisson E, Sorkin RD (2016) Gravitational action with null boundaries. Phys Rev D 94:084046. https://doi.org/10.1103/PhysRevD.94.084046. arXiv:1609.00207

Levichev AV (1987) Prescribing the conformal geometry of a lorentz manifold by means of its causal structure. Sov Math Dokl 35:452–455

Liberati S, Mattingly D (2016) Lorentz breaking effective field theory models for matter and gravity: theory and observational constraints. In: Peron R, Colpi M, Gorini V, Moschella U (eds) Gravity: where do we stand?. Springer, Cham, pp 367–417. https://doi.org/10.1007/978-3-319-20224-2_11. arXiv:1208.1071

Loomis SP, Carlip S (2018) Suppression of non-manifold-like sets in the causal set path integral. Class Quantum Grav 35:024002. https://doi.org/10.1088/1361-6382/aa980b. arXiv:1709.00064

Louko J, Sorkin RD (1997) Complex actions in two-dimensional topology change. Class Quantum Grav 14:179–204. https://doi.org/10.1088/0264-9381/14/1/018

Major S, Rideout D, Surya S (2007) On recovering continuum topology from a causal set. J Math Phys 48:032501. https://doi.org/10.1063/1.2435599. arXiv:gr-qc/0604124

Major S, Rideout D, Surya S (2009) Stable homology as an indicator of manifoldlikeness in causal set theory. Class Quantum Grav 26:175008. https://doi.org/10.1088/0264-9381/26/17/175008. arXiv:0902.0434

Major SA, Rideout D, Surya S (2006) Spatial hypersurfaces in causal set cosmology. Class Quantum Grav 23:4743. https://doi.org/10.1088/0264-9381/23/14/011

Malament DB (1977) The class of continuous timelike curves determines the topology of spacetime. J Math Phys 18:1399–1404. https://doi.org/10.1063/1.523436

Marr S (2007) Black hole entropy from causal sets. PhD thesis, Imperial College http://hdl.handle.net/10044/1/11818
Martin X, O'Connor D, Rideout DP, Sorkin RD (2001) On the 'renormalization' transformations induced by cycles of expansion and contraction in causal set cosmology. Phys Rev D 63:084026. https://doi.org/10.1103/PhysRevD.63.084026. arXiv:gr-qc/0009063
Mathur A, Surya S (2019) Sorkin–Johnston vacuum for a massive scalar field in the 2D causal diamond. Phys Rev D 100:045007. https://doi.org/10.1103/PhysRevD.100.045007. arXiv:1906.07952
Meyer D (1988) The dimension of causal sets. PhD thesis, M.I.T
Munkres JR (1984) Elements of algebraic topology. Addison-Wesley, Reading
Myrheim J (1978) Statistical geometry. Technical Report CERN-TH-2538, CERN
Noldus J (2002) A new topology on the space of Lorentzian metrics on a fixed manifold. Class Quantum Grav 19:6075–6107. https://doi.org/10.1088/0264-9381/19/23/313. arXiv:1104.1811
Noldus J (2004) A Lorentzian Gromov–Hausdorff notion of distance. Class Quantum Grav 21:839–850. https://doi.org/10.1088/0264-9381/21/4/007. arXiv:gr-qc/0308074
Parrikar O, Surya S (2011) Causal topology in future and past distinguishing spacetimes. Class Quantum Grav 28:155020. https://doi.org/10.1088/0264-9381/28/15/155020. arXiv:1102.0936
Penrose R (1972) Techniques of differential topology in relativity. SIAM, Philadelphia
Petersen P (2006) Riemannian geometry, 2nd edn. Springer, New York. https://doi.org/10.1007/978-0-387-29403-2
Philpott L, Dowker F, Sorkin RD (2009) Energy-momentum diffusion from spacetime discreteness. Phys Rev D 79:124047. https://doi.org/10.1103/PhysRevD.79.124047. arXiv:0810.5591
Promel H, Steger A, Taraz A (2001) Phase transitions in the evolution of partial orders. J Combin Theory Ser A 94:230
Reid DD (2003) Manifold dimension of a causal set: tests in conformally flat spacetimes. Phys Rev D 67:024034. https://doi.org/10.1103/PhysRevD.67.024034. arXiv:gr-qc/0207103
Rideout D, Wallden P (2009) Spacelike distance from discrete causal order. Class Quantum Grav 26:155013. https://doi.org/10.1088/0264-9381/26/15/155013. arXiv:0810.1768
Rideout D, Zohren S (2006) Evidence for an entropy bound from fundamentally discrete gravity. Class Quantum Grav 23:6195–6213. https://doi.org/10.1088/0264-9381/23/22/008. arXiv:gr-qc/0606065
Rideout DP (2001) Dynamics of causal sets. PhD thesis, Syracuse University arXiv:gr-qc/0212064
Rideout DP, Sorkin RD (2000a) Classical sequential growth dynamics for causal sets. Phys Rev D 61:024002. https://doi.org/10.1103/PhysRevD.61.024002
Rideout DP, Sorkin RD (2000b) Classical sequential growth dynamics for causal sets. Phys Rev D 61:024002. https://doi.org/10.1103/PhysRevD.61.024002. arXiv:gr-qc/9904062
Rideout DP, Sorkin RD (2001) Evidence for a continuum limit in causal set dynamics. Phys Rev D 63:104011. https://doi.org/10.1103/PhysRevD.63.104011. arXiv:gr-qc/0003117
Riemann B (1873) On the hypotheses which lie at the bases of geometry. Nature VIII(183, 184):14–17, 36, 37. https://doi.org/10.1038/008036a0, translated by W. K. Clifford from Vol. VIII of the Göttingen Abhandlungen
Robb A (1914) A theory of time and space. Cambridge University Press, Cambridge
Robb A (1936) Geometry of time and space. Cambridge University Press, Cambridge
Roy M, Sinha D, Surya S (2013) Discrete geometry of a small causal diamond. Phys Rev D 87:044046. https://doi.org/10.1103/PhysRevD.87.044046. arXiv:1212.0631
Salgado RB (2002) Some identities for the quantum measure and its generalizations. Mod Phys Lett A 17:711–728. https://doi.org/10.1142/S0217732302007041. arXiv:gr-qc/9903015
Salgado RB (2008) Toward a quantum dynamics for causal sets. PhD thesis, Syracuse University
Samuel J, Sinha S (2006) Surface tension and the cosmological constant. Phys Rev Lett 97:161302. https://doi.org/10.1103/PhysRevLett.97.161302. arXiv:cond-mat/0603804
Saravani M, Afshordi N (2017) Off-shell dark matter: a cosmological relic of quantum gravity. Phys Rev D 95:043514. https://doi.org/10.1103/PhysRevD.95.043514
Saravani M, Aslanbeigi S (2014) On the causal set-continuum correspondence. Class Quantum Grav 31:205013. https://doi.org/10.1088/0264-9381/31/20/205013. arXiv:1403.6429
Saravani M, Sorkin RD, Yazdi YK (2014) Spacetime entanglement entropy in 1 + 1 dimensions. Class Quantum Grav 31:214006. https://doi.org/10.1088/0264-9381/31/21/214006. arXiv:1311.7146
Sorkin RD (1991) Spacetime and causal sets. In: D'Olivo JC (ed) Relativity and gravitation: classical and quantum, proceedings of the SILARG VII conference, Cocoyocan, Mexico. World Scientific, Singapore, pp 150–173

Sorkin RD (1994) Quantum mechanics as quantum measure theory. Mod Phys Lett A 9:3119–3128. https://doi.org/10.1142/S021773239400294X. arXiv:gr-qc/9401003
Sorkin RD (1995) Quantum measure theory and its interpretation. In: Feng DH, Hu B-L (eds) Quantum classical correspondence: proceedings of the 4th Drexel symposium on quantum nonintegrability, Philadelphia, 8–11 Sept 1994, pp 229–251. International Press, Cambridge arXiv:gr-qc/9507057
Sorkin RD (1997) Forks in the road, on the way to quantum gravity. Int J Theor Phys 36:2759–2781. https://doi.org/10.1007/BF02435709. arXiv:gr-qc/9706002
Sorkin RD (2005a) Big extra dimensions make Λ too small. Braz J Phys 35:280–283. https://doi.org/10.1590/S0103-97332005000200012. arXiv:gr-qc/0503057
Sorkin RD (2005b) Causal sets: discrete gravity. In: Gomberoff A, Marolf D (eds) Lectures on quantum gravity. Series of the Centro de Estudios Científicos. Springer, New York, pp 305–327. https://doi.org/10.1007/0-387-24992-3_7. arXiv:gr-qc/0309009
Sorkin RD (2007a) An exercise in 'anhomomorphic logic'. J Phys: Conf Ser 67:012018. https://doi.org/10.1088/1742-6596/67/1/012018. arXiv:quant-ph/0703276
Sorkin RD (2007b) Does locality fail at intermediate length-scales. In: Oriti D (ed) Approaches to quantum gravity. Cambridge University Press, Cambridge, pp 26–43 arXiv:gr-qc/0703099
Sorkin RD (2007c) An exercise in "anhomomorphic logic". J Phys: Conf Ser 67:012018. https://doi.org/10.1088/1742-6596/67/1/012018
Sorkin RD (2007d) Quantum dynamics without the wave function. J Phys A 40:3207–3222. https://doi.org/10.1088/1751-8113/40/12/S20. arXiv:quant-ph/0610204
Sorkin RD (2009) Light, links and causal sets. J Phys: Conf Ser 174:012018. https://doi.org/10.1088/1742-6596/174/1/012018. arXiv:0910.0673
Sorkin RD (2011a) Scalar field theory on a causal set in histories form. J Phys: Conf Ser 306:012017. https://doi.org/10.1088/1742-6596/306/1/012017. arXiv:1107.0698
Sorkin RD (2011b) Toward a 'fundamental theorem of quantal measure theory'. arXiv e-prints arXiv:1104.0997
Sorkin RD (2014) Expressing entropy globally in terms of (4D) field-correlations. J Phys: Conf Ser 484:012004. https://doi.org/10.1088/1742-6596/484/1/012004. arXiv:1205.2953
Sorkin RD, Yazdi YK (2018) Entanglement entropy in causal set theory. Class Quantum Grav 35:074004. https://doi.org/10.1088/1361-6382/aab06f. arXiv:1611.10281
Stachel J (1986) Einstein and the quantum: fifty years of struggle. In: Colodny R (ed) From quarks to quasars, philosophical problems of modern physics. University of Pittsburgh Press, Pittsburgh, p 379
Stanley RP (2011) Enumerative combinatorics, vol I, 2nd edn. Cambridge University Press, Cambridge
Stoyan D, Kendall W, Mecke J (1995) Stochastic geometry and its applications. Wiley, Chichester
Surya S (2008) Causal set topology. Theor Comput Sci 405:188–197. https://doi.org/10.1016/j.tcs.2008.06.033. arXiv:0712.1648
Surya S (2012) Evidence for the continuum in 2D causal set quantum gravity. Class Quantum Grav 29:132001. https://doi.org/10.1088/0264-9381/29/13/132001
Surya S, Nomann X, Yazdi YK (2019) Studies on the SJ vacuum in de Sitter spacetime. JHEP 07:009. https://doi.org/10.1007/JHEP07(2019)009. arXiv:1812.10228
Sverdlov R, Bombelli L (2009) Gravity and matter in causal set theory. Class Quantum Grav 26:075011. https://doi.org/10.1088/0264-9381/26/7/075011. arXiv:0801.0240
Unruh WG, Wald RM (1989) Time and the interpretation of canonical quantum gravity. Phys Rev D 40:2598. https://doi.org/10.1103/PhysRevD.40.2598
Varadarajan M, Rideout D (2006) A general solution for classical sequential growth dynamics of causal sets. Phys Rev D 73:104021. https://doi.org/10.1103/PhysRevD.73.104021. arXiv:gr-qc/0504066
Wald R (1984) General relativity. University of Chicago Press, Chicago
Wald RM (1994) Quantum field theory in curved spacetime and black hole thermodynamics. University of Chicago Press, Chicago
Wallden P (2013) Causal sets dynamics: review & outlook. J Phys: Conf Ser 453:012023. https://doi.org/10.1088/1742-6596/453/1/012023
Winkler P (1991) Random orders of dimension 2. Order 7:329. https://doi.org/10.1007/BF00383197
Yazdi YK, Kempf A (2017) Towards spectral geometry for causal sets. Class Quantum Grav 34:094001. https://doi.org/10.1088/1361-6382/aa663f. arXiv:1611.09947
Zeeman EC (1964) Causality implies the Lorentz group. J Math Phys 5(4):490–493. https://doi.org/10.1063/1.1704140

Zuntz JA (2008) The cosmic microwave background in a causal set universe. Phys Rev D 77:043002. https://doi.org/10.1103/PhysRevD.77.043002. arXiv:0711.2904

Zwane N, Afshordi N, Sorkin RD (2018) Cosmological tests of everpresent Λ. Class Quantum Grav 35:194002. https://doi.org/10.1088/1361-6382/aadc36. arXiv:1703.06265