

Local d'Alembertian for causal sets

Marián Boguñá^{1,2,*} and Dmitri Krioukov^{3,†}

¹*Departament de Física de la Matèria Condensada, Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona, Spain*

²*Universitat de Barcelona Institute of Complex Systems (UBICS), Universitat de Barcelona, Barcelona, Spain*

³*Department of Physics, Network Science Institute, Department of Mathematics, and Department of Electrical and Computer Engineering, Northeastern University, Boston, Massachusetts 02115, USA*



(Received 31 July 2025; accepted 24 December 2025; published 22 January 2026)

Causal set theory is an intrinsically nonlocal approach to quantum gravity, inheriting its nonlocality from Lorentzian nonlocality. This nonlocality causes problems in defining differential operators—such as the d'Alembert operator, a cornerstone of any relativistic field theory—within the causal set framework. It has been proposed that d'Alembertians in causal sets must themselves be nonlocal to respect causal set nonlocality. However, we show that such nonlocal d'Alembertians do not converge to the standard continuum d'Alembertian for some basic fields. To address this problem, we introduce a local d'Alembert operator for causal sets and demonstrate its convergence to its continuum counterpart for arbitrary fields in Minkowski spacetimes. Our construction leverages recent developments in measuring both timelike and spacelike distances in causal sets using only their intrinsic structure. This approach thus reconciles locality with Lorentz invariance, and paves a way toward defining converging discrete approximations to locality-based differential operators even in theories that are inherently nonlocal.

DOI: [10.1103/mp89-qzjp](https://doi.org/10.1103/mp89-qzjp)

I. INTRODUCTION

Causal set theory (CST) [1] is a peculiar actor in the never-ending dramatic quest for a theory of quantum gravity. CST is one of the most minimalistic approaches there, in terms of the assumptions it makes. It simply posits that spacetime, at the Planck scale, is not a continuous manifold but rather a discrete object made up of elementary spacetime atoms, with no internal structure, connected by causal relations [2–9]. The approach finds its origins in the work of Hawking *et al.* [10], as well as Malament [11], who demonstrated that spacetimes that satisfy some natural niceness assumptions, and that have the same causal structure, are equivalent modulo conformal rescaling.

CST suggests that the smooth continuum of spacetime that we observe at macroscopic scales is an illusion, a coarse-grained smearing of an underlying fine-grained discrete structure at the Planck scale. The process by which this discrete structure transitions to the smooth spacetime continuum is known as the continuum limit of causal sets [12–17]. Achieving this limit requires an ability to measure distances using only the causal set structure, which is relatively straightforward for timelike separated events,

but considerably more challenging for spacelike separated events [18–21].

CST has traditionally been viewed as a nonlocal theory [22], a perspective that arises from the observation that in causal sets obtained from random sprinklings on Lorentzian manifolds, the number of nearest neighbors of a given event is infinite and distributed on a hyperbolic surface of constant proper time of the order of the Planck time. These nearest neighbors, although separated by one unit of proper time, can be arbitrarily distant in space. Yet thanks to Lorentz invariance, there is seemingly nothing that can differentiate one such neighbor from another, so all of them must be considered on the same footing in any field theory defined on the causal set.

This line of reasoning inspired the development of nonlocal d'Alembert operators for causal sets [22–27]. These operators are nonlocal in that they aim to account equally for all events in the aforementioned infinite set of the nearest neighbors of a given event in a causal set. Yet at the same time they attempt to resurrect locality in the continuum limit attained at the infinite event density. While these two goals appear mutually contradictory, nonlocal discrete d'Alembertians have nevertheless succeeded in converging to the continuum limit when acting upon many types of scalar fields. Many but not all. Unlike their continuous parent, the standard d'Alembertian, these nonlocal operators lead to

* Contact author: marian.boguna@ub.edu

† Contact author: dima@northeastern.edu

divergences when acting on certain scalar fields, including the simplest one, a constant field, as we show in Sec. III. These divergences can be suppressed by requiring that the scalar field upon which the operator acts is zero everywhere except a bounded chunk of spacetime. However, this requirement is equivalent to predetermining the set of events used in the evaluation of the operator, which amounts to defining a specific reference frame. More importantly, it is reasonable to require a discrete d'Alembertian to converge to the continuous one on any good-looking field, not only on some specific fields.

The key idea behind our work presented here is very simple. While all neighbors of a given event in a causal set are indeed equivalent with respect to this event, they are not equivalent among themselves. In [21], we have shown that using only the causal set structure, it is possible to measure spatial distances between causal set events all the way down to the smallest possible distances of the order of the Planck length. This means that we can take one of those neighboring events, and then sort the rest of them in the order of increasing proper length from the selected neighbor. This implies that local neighborhoods can be defined and, consequently, local differential operators can be constructed without breaking Lorentz invariance.

Our main contributions in this paper are a definition of a local discrete d'Alembert operator based on the idea above, and analytic and numerical evidence that this operator converges to the continuous d'Alembertian on any analytic scalar field in the $(d + 1)$ -dimensional Minkowski spacetime.

We begin with an illustration of key components of our construction in much simpler settings in Sec. II, where we define a discrete field Laplacian for random geometric graphs in the Euclidean plane, and show that this Laplacian converges to the standard continuous Laplacian acting on the same field. We then proceed to Sec. III where we discuss key complications arising from Lorentzian non-locality, including the divergences of nonlocal discrete d'Alembertians mentioned above. Finally, by far the longest Sec. IV, with its own internal structure, contains a detailed account of our discrete d'Alembertian construction, and the analytical and numerical demonstrations of its convergence to the continuous d'Alembert operator in the continuum limit. Some concluding remarks are in Sec. V.

II. LAPLACIAN IN RANDOM GEOMETRIC GRAPHS

To illustrate the key ideas behind our discrete d'Alembertian and its convergence to the continuum one, we first consider the Riemannian case, which is simpler than the Lorentzian case, yet all the main ingredients are present, except one—nonlocal connections. In Riemannian manifolds, the d'Alembertian is the Laplacian, while the role of causal sets is played by random geometric graphs. In these graphs, only those node pairs that are close to each other in the space are connected. In this section, we

illustrate how the discrete structure of these only-local connections can be used to evaluate the action of the Laplacian on a scalar field. This example provides a clear introduction to the techniques that we will later apply to the more complex Lorentzian settings that involve nonlocal connections. To the best of our knowledge, the results in [28], dealing with the convergence of a discrete Laplacian defined for point clouds in Riemannian manifolds, are closest in spirit to the results presented in this section; there are also many results concerning the convergence of the standard graph Laplacian of random geometric graphs [29,30], as opposed to the field Laplacian we are considering here.

Consider a random geometric graph in the Euclidean plane, generated by a Poisson point process with density ρ , and a connection radius r_c . In this graph, two nodes are connected if and only if the Euclidean distance between them is less than r_c . The degrees of nodes are Poisson-distributed random variables with mean $\langle k \rangle = \rho \pi r_c^2$ and variance $\sigma_k^2 = \langle k \rangle$. Therefore, in the limit $\langle k \rangle \gg 1$, the relative fluctuations $\sigma_k / \langle k \rangle$ vanish, so we can assume that all nodes have approximately the same number of neighbors $\langle k \rangle$. Figure 1 shows a visualization of a random geometric graph, highlighting a central node and its neighbors.

Since the graph can be thought of as a random discretization of the Euclidean plane and distances in it, we are tempted to ask to what extent this discrete sampling of a continuous manifold can recover its continuous properties.

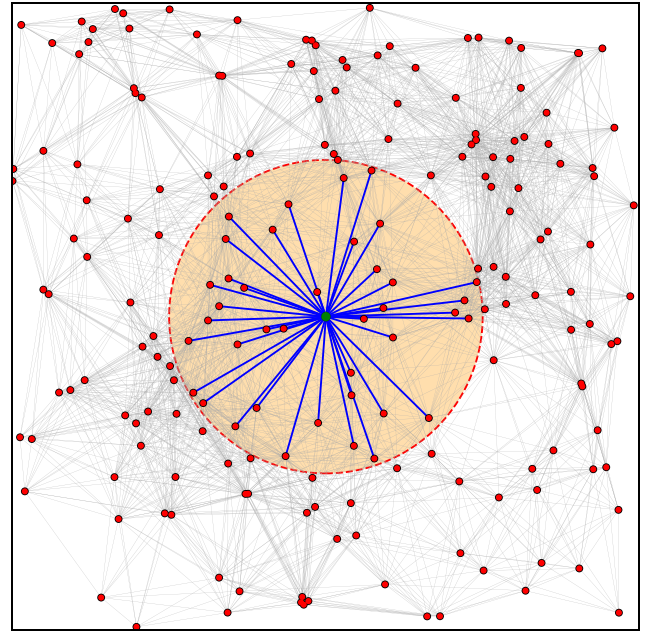


FIG. 1. Random geometric graphs. The figure shows a random geometric graph generated by the Poisson point process on the Euclidean unit square with point density $\rho = 200$ and connection radius $r_c = 0.26$. The blue links within the shaded circle represent the neighbors of the central node in green. The discrete Laplacian in Eq. (5), evaluated at the central node, is computed using the values of the scalar field at these neighboring nodes.

Specifically, we are interested in evaluating the Laplace operator acting on a scalar field $\phi(x, y)$ defined at every point on the plane, and, consequently, at any node of the graph. If we evaluate the Laplacian using only the graph structure and the value of the ϕ -field at its nodes, will it converge to the Laplacian evaluated on the plane in a certain limit?

The answer to this question is positive, and to illustrate this convergence, let us evaluate the graph Laplacian at a particular node, which we assume, without loss of generality, is located at the origin, so we call this node 0. In the limit $r_c \rightarrow 0$, the scalar field at 0's neighbor i at coordinates $\mathbf{r}_i = (x_i, y_i)$ can be approximated using the Taylor series as

$$\begin{aligned} \phi(x_i, y_i) \equiv \phi_i &= \phi_0 + \nabla\phi_0 \cdot \mathbf{r}_i \\ &+ \frac{1}{2} \left(\frac{\partial^2 \phi}{\partial x^2} \Big|_0 x_i^2 + \frac{\partial^2 \phi}{\partial y^2} \Big|_0 y_i^2 + 2 \frac{\partial^2 \phi}{\partial x \partial y} \Big|_0 x_i y_i \right) \\ &+ \mathcal{O}(r_c^3), \quad x_i, y_i < r_c, \end{aligned} \quad (1)$$

where the subscript 0 indicates that the field and its derivatives are evaluated at the node 0, located at the origin $x = y = 0$. We now want to take the average of Eq. (1) over the set of k_0 neighbors of the 0 node. For that end, we define their average x -coordinates (squared) as

$$\bar{x} \equiv \frac{1}{k_0} \sum_{i=1}^{k_0} x_i, \quad \bar{x}^2 \equiv \frac{1}{k_0} \sum_{i=1}^{k_0} x_i^2, \quad (2)$$

and similarly for the y -coordinates. The quantities $\bar{x}$ and $\bar{x}^2$ are still random variables since they depend on a particular realization of the Poisson point process. Due to the Euclidean symmetries, the mean of $\bar{x}$ is zero, $\langle \bar{x} \rangle = 0$, while its variance is $\langle x^2 \rangle / \langle k \rangle$, where $\langle x^2 \rangle \equiv \langle \bar{x}^2 \rangle$ is the mean of $\bar{x}^2$. One can check that $\langle x^2 \rangle = r_c^2/4$, substituting which into the average of (1), we observe that the average $\bar{\phi}$ of the values of the scalar field ϕ over the neighbors of node 0 can be written as

$$\bar{\phi} = \phi_0 + \frac{1}{2} \nabla^2 \phi_0 \langle x^2 \rangle + \mathcal{O} \left(\sqrt{\frac{\langle x^2 \rangle}{\langle k \rangle}} \right) + \mathcal{O}(r_c^4), \quad (3)$$

where the first error term comes from $\bar{x}$ and the second from the truncated Taylor expansion up to fourth order. Using $\langle x^2 \rangle = r_c^2/4$ and $\langle k \rangle = \rho \pi r_c^2$, we can then write the Laplacian as

$$\nabla^2 \phi_0 = \frac{8}{r_c^2} (\bar{\phi} - \phi_0) + \mathcal{O} \left(\frac{1}{r_c^2 \sqrt{\rho}} \right) + \mathcal{O}(r_c^2). \quad (4)$$

The last equation says that the continuous Laplacian $\nabla^2 \phi_0$ of the field evaluated at the origin can be recovered in

the $r_c \rightarrow 0$ limit from the discrete average $\bar{\phi}$ of the field over the neighbors of the 0 node in the random geometric graph. That is,

$$\mathcal{D}\phi_0 = \frac{8}{r_c^2} (\bar{\phi} - \phi_0) \quad (5)$$

can be understood as a discrete Laplacian associated with our random geometric graph. Yet this continuous recovery is possible, and the discrete Laplacian converges to its continuous father,

$$\lim_{r_c \rightarrow 0} \mathcal{D}\phi_0 = \nabla^2 \phi_0, \quad (6)$$

only if both the error terms in (4) are small in the limit. They are both small if

$$\frac{1}{\rho^{1/4}} \ll r_c \ll 1. \quad (7)$$

This condition is fulfilled in the continuum limit $\rho \rightarrow \infty$ if r_c is chosen to be $r_c \sim \rho^{-\delta}$ with $0 < \delta < 1/4$. This limit means that we are sampling the space more and more densely ($\rho \rightarrow \infty$), while links in our random geometric graph become increasingly microscopic ($r_c \rightarrow 0$), thus approximating distances in the space with growing precision. There is also the optimal value of $\delta = 1/8$ making the scaling of both error terms in (4) with ρ the same $\mathcal{O}(\rho^{-1/4})$, so Eq. (4) becomes

$$\nabla^2 \phi_0 = \mathcal{D}\phi_0 + \mathcal{O}(\rho^{-1/4}). \quad (8)$$

To validate this result in simulations, we choose the scalar field $\phi(x, y) = e^{ax+by}$, for which $\nabla^2 \phi_0 = a^2 + b^2$, and then sample 100 values of a and b selected uniformly at random within the interval $[0, 3]$. For each pair of values (a, b) , we generate random geometric graphs at different densities ρ , with the connection radius scaling as $r_c \sim \rho^{-1/8}$. We then compute the discrete Laplacian in the generated graphs according to Eq. (5), and compare the results of these computations with the continuous value $a^2 + b^2$. The outcomes of these experiments shown in Fig. 2 confirm the expected: the discrete Laplacian converges to the continuous one as ρ increases.

These results provide a simple illustration that differential operators on smooth manifolds can be approximated using discrete random samplings of the manifold, and that under appropriate scaling conditions, such approximations can become exact in the continuum limit when the sampling density approaches infinity.

The key idea behind the approximation used in the discrete Laplacian in Eq. (5) is fairly straightforward. Indeed, since the Laplacian operator is a measure of how fast the field changes around a point, the properly

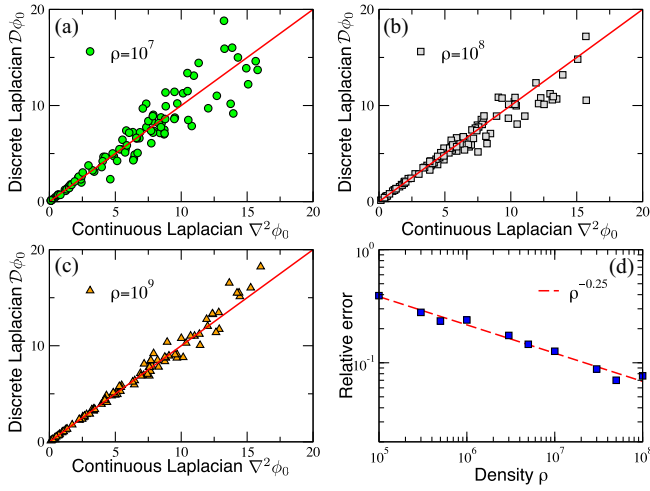


FIG. 2. Laplacian simulations. (a)–(c) The values of the discrete Laplacian $\mathcal{D}\phi_0$ in Eq. (5) acting on the field $\phi = e^{ax+by}$ at $x = y = 0$ vs the continuous value $\nabla^2\phi_0 = a^2 + b^2$ for different point densities ρ . For each value of a and b , a separate realization of the Poisson point process within the disk $r < r_c = 0.75\rho^{-1/8}/\sqrt{a^2 + b^2}$ is sampled. (d) The average relative error of the discrete vs the continuous Laplacian, defined as $\langle |(\mathcal{D}\phi_0/\nabla^2\phi_0) - 1| \rangle$, for different values of ρ . For each value of ρ , the relative error is averaged of 100 random values of a , b , and Poisson point process realizations. The red dashed line is the predicted scaling behavior $\rho^{-1/4}$ from Eq. (8).

rescaled difference of the values of the field at a node and its neighbors in a random geometric graph converges to the continuous Laplacian.

However, while in Riemannian geometry, this convergence is relatively easy to prove, and the intrinsic non-locality of Lorentzian geometry complicates things significantly. We discuss these complications next.

III. NONLOCAL D'ALEMBERTIANS IN CAUSAL SETS

In this section we show how Lorentzian nonlocality leads to the lack of convergence in discrete nonlocal d'Alembertians applied to nonlocal fields.

Causal sets built on top of Lorentzian manifolds are akin to random geometric graphs in Riemannian manifolds. The vertices in both cases are realizations of Poisson point processes on a manifold, but edges in causal sets reflect causality instead of spatial proximity: two vertices in a causal set are linked if they are timelike-separated in the spacetime manifold. We will refer to vertices in causal sets as *events*, and will assume that links implied by transitivity are removed.

The Laplace operator in Lorentzian manifold is the d'Alembert operator, which in natural units is $\square \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu = -\partial_t^2 + \nabla^2$ in Minkowski spacetime. While this operator is perfectly local when defined in continuous spacetime, its satisfactory definition in a causal set that

discretely samples this spacetime is not immediately clear. Indeed, if Minkowski spacetime is sampled at finite density ρ , each event in the sampling is directly connected to an infinite number of other events lying approximately on the hyperboloid of constant proper time $\tau_c \propto \rho^{-1/(d+1)} = t_P$, where d is the dimension of the spatial part of the metric tensor, and t_P is defined to be the Planck time. Therefore, while these neighbors of the event are at a small temporal distance from it, they can be arbitrarily far apart among themselves in terms of spacelike distance. This observation suggests that discrete versions of Lorentzian manifolds are intrinsically nonlocal, and thus, operators acting on them should also be nonlocal.

Nonlocal d'Alembertians acting on an analytic scalar field ϕ at a given event a_0 were first introduced in [22] for the $\mathbb{M}^{1+1}$ Minkowski spacetime, extended later to $\mathbb{M}^{3+1}$ in [23], and generalized further in a number of papers [24–27], where they appear as a basis for a definition of action in causal sets. The main goal behind such definitions is to come up with discrete d'Alembertians that are linear combinations of the scalar field evaluated at specific events in the past of a_0 , and that satisfy the following key requirements [26]:

- (1) *Invariance*: a discrete d'Alembertian B must be Lorentz invariant.
- (2) *Convergence*: the expectation of B 's values averaged over Poisson sprinklings of increasing densities ρ must converge to the continuous d'Alembertian $\square$ in the continuum limit where the sprinkling density goes to infinity:

$$\lim_{\rho \rightarrow \infty} \langle B_\rho \rangle = \square. \quad (9)$$

The existing B definitions satisfy all these requirements, although to satisfy the last one—the convergence in the continuum limit—they require that the field ϕ to which B is applied must have a compact support. This requirement is not unreasonable as it finds many physical motivations. For example, in the Algebraic Quantum Field Theory (AQFT) formulation of Quantum Field Theory (QFT), one often smears the fields with test functions of compact support, and it is in this context that the d'Alembertian is perhaps of most relevance. Another consideration justifying this requirement is that any physical measurement deals only with a finite region of spacetime or is equivalently constrained by cosmological (Hubble) scales. Yet, if applied to a noncompact field (e.g., matter in a whole noncompact universe), such nonlocal d'Alembertians may not converge.

To illustrate this point, let us consider the two-dimensional (2D) Minkowski spacetime for simplicity; similar results hold in other dimensions. In two dimensions, the simplest nonlocal d'Alembertian operator B_ρ acting on the scalar field ϕ at a_0 is defined as [22]

$$(B_\rho^{(2)}\phi)(a_0) \equiv \rho \left[-2\phi(a_0) + 4 \left(\sum_{\mathbf{x} \in I_0(a_0)} \phi(\mathbf{x}) - 2 \sum_{\mathbf{x} \in I_1(a_0)} \phi(\mathbf{x}) + \sum_{\mathbf{x} \in I_2(a_0)} \phi(\mathbf{x}) \right) \right], \quad (10)$$

where $I_n(a_0)$ is the set of events $\mathbf{x}$ in the past of a_0 such that the number of events in the Alexandrov set between a_0 and $\mathbf{x}$ is n . Given the symmetry of Minkowski spacetime, the same definition can be made using events in the future of a_0 instead of its past.

Equation (10) defines a random variable that depends on a particular realization of the Poisson point process. The average of this random variable over Poisson point process realizations is given by

$$\langle (B_\rho^{(2)}\phi)(a_0) \rangle = \rho \left[-2\phi(a_0) + 4\rho \left(\int_{\mathbf{x} \in \text{Past}(a_0)} \sqrt{-g} d\mathbf{x} e^{-\rho V(\mathbf{x}|a_0)} \left(1 - 2\rho V(\mathbf{x}|a_0) + \frac{1}{2}(\rho V(\mathbf{x}|a_0))^2 \right) \phi(\mathbf{x}) \right) \right], \quad (11)$$

where the integral extends over the past light cone of a_0 and $V(\mathbf{x}|a_0)$ is the volume of the Alexandrov interval between a_0 and $\mathbf{x}$. We next rewrite Eq. (11) as

$$\langle (B_\rho^{(2)}\phi)(a_0) \rangle = \rho \left[-2\phi(a_0) + 4\rho \left(1 + 2\rho \frac{d}{d\rho} + \frac{1}{2}\rho^2 \frac{d^2}{d\rho^2} \right) I(\rho) \right], \quad (12)$$

where

$$I(\rho) = \int_{\mathbf{x} \in \text{Past}(a_0)} \sqrt{-g} d\mathbf{x} e^{-\rho V(\mathbf{x}|a_0)} \phi(\mathbf{x}). \quad (13)$$

Next, without loss of generality, we set a_0 to be the origin, $a_0 = (0, 0)$, and thanks to the symmetry of Minkowski spacetime, we use the future of a_0 instead of its past. The future light cone (t, x) of a_0 can be foliated by hyperbolas with hyperbolic coordinates (τ, χ) given by

$$t = \tau \cosh \chi, \quad (14)$$

$$x = \tau \sinh \chi, \quad (15)$$

where $\tau \in (0, \infty)$ is the proper time from a_0 to the hyperbola and $\chi \in (-\infty, \infty)$ is its spatial coordinate. Let us now define the dimensionless proper time

$$\hat{\tau} = \sqrt{\frac{\rho}{2}} \tau, \quad (16)$$

and the integral

$$\hat{I}(\rho, \chi) \equiv \frac{2}{\rho} \int_0^\infty d\hat{\tau} \hat{\tau} e^{-\hat{\tau}^2} \phi \left(\sqrt{\frac{2}{\rho}} \hat{\tau} \cosh \chi, \sqrt{\frac{2}{\rho}} \hat{\tau} \sinh \chi \right). \quad (17)$$

With these notations, the integral $I(\rho)$ in Eq. (12) becomes $I(\rho) = \int_{-\infty}^\infty \hat{I}(\rho, \chi) d\chi$, so Eq. (12) can be rewritten as

$$\langle (B_\rho^{(2)}\phi)(a_0) \rangle = \rho \left[-2\phi(a_0) + 4\rho \int_{-\infty}^\infty d\chi \left(1 + 2\rho \frac{d}{d\rho} + \frac{1}{2}\rho^2 \frac{d^2}{d\rho^2} \right) \hat{I}(\rho, \chi) \right]. \quad (18)$$

We next use this expression to analyze the d'Alembertian operator $B_\rho^{(2)}$ acting upon simple test fields.

- (1) *Example 1: Constant scalar field.* If the scalar field ϕ is constant, $\phi(t, x) = \phi_0$, then the continuous d'Alembertian applied to this field yields zero,

$$\square \phi = 0. \quad (19)$$

To see what we get from the discrete d'Alembertian, observe that $\hat{I}(\rho, \chi) = \phi_0/\rho$, substituting which into Eq. (18), we get

$$\left(1 + 2\rho \frac{d}{d\rho} + \frac{1}{2}\rho^2 \frac{d^2}{d\rho^2} \right) \hat{I}(\rho, \chi) = 0, \quad (20)$$

leading to

$$\langle (B_\rho^{(2)}\phi)(a_0) \rangle = -2\rho\phi_0. \quad (21)$$

That is, instead of being zero, the discrete nonlocal d'Alembertian diverges in the limit $\rho \rightarrow \infty$, unless $\phi_0 = 0$. A similar divergence takes place with any scalar field that converges to a constant in the limit $x \rightarrow \pm\infty$, or is a growing function of x , or does not go to 0 at large x sufficiently fast.

- (2) *Example 2: Scalar field as a function of $\tau^2 = t^2 - x^2$.* As an example of an even worse divergence, let us consider the scalar field given by

$$\phi(t, x) = \tilde{\phi}(\tau^2) = e^{-\tau^2}, \quad (22)$$

for which the continuous d'Alembertian yields

$$\square\phi(0, 0) = 4. \quad (23)$$

Turning to the discrete d'Alembertian, Eq. (17) yields

$$\hat{I}(\rho, \chi) = \frac{1}{2 + \rho}, \quad (24)$$

which does not depend on χ , so Eq. (18) says that

$$\langle (B_\rho^{(2)}\phi)(a_0) \rangle = -2\rho + \frac{16\rho^2}{(2 + \rho)^3} \int_{-\infty}^{\infty} d\chi = \infty. \quad (25)$$

That is, in this case, the discrete d'Alembertian diverges even for any finite ρ due to the divergence of the integral over χ . A similar divergence holds for any field that is a function of τ^2 .

The observed lack of the convergence of the discrete nonlocal d'Alembertian to the continuous one in the examples above arises from the fact that the considered test fields have noncompact supports. The convergence is restored if the field is set to zero outside of a finite-length interval.

To see this explicitly, let us consider the constant field $\phi(t, x) = \phi_0$ for $x \in [-L, L]$, while at $|x| > L$ the field is set to 0. In this case, the integral $I(\rho)$ becomes

$$I(\rho) = \frac{2\sqrt{\pi}\phi_0}{\rho} \int_0^{L\sqrt{\frac{\rho}{2}}} e^{x^2} \text{Erfc}(x) dx, \quad (26)$$

where $\text{Erfc}(x)$ is the complementary error function. Plugging this expression into Eq. (12), we get

$$\langle (B_\rho^{(2)}\phi)(a_0) \rangle = \rho\phi_0[-2 + \sqrt{\pi}s(3 + 2s^2)e^{s^2}\text{Erfc}(s) - 2s^2], \quad (27)$$

where $s = L\sqrt{\rho/2}$. In the limit $\rho \rightarrow \infty$, this expression is asymptotically

$$\langle (B_\rho^{(2)}\phi)(a_0) \rangle \sim \frac{1}{L^4\rho}. \quad (28)$$

Therefore, for any fixed L , the discrete nonlocal d'Alembertian converges to zero, as it should on a constant field.

This result holds for an interval of any size L as long as it is finite. In view of this observation, one may hope that the limit $\rho \rightarrow \infty$ could lead to the correct continuum limit for fields with noncompact supports. However, even if L is

taken to be arbitrarily large, the contribution to the nonlocal d'Alembertian from the region $|x| > L$ remains non-negligible. Consequently, the order of limits matter: if we first take $\rho \rightarrow \infty$ and then $L \rightarrow \infty$, then we get 0 according to Eq. (28), but if we first take $L \rightarrow \infty$ and then $\rho \rightarrow \infty$, then we get ∞ according to Eq. (21). In the former case, we are dealing with constant fields with compact supports, while in the latter case, our constant field has a noncompact support.

A related important observation, which is typical for situations like here, is that we can always do some mathematical tricks that do not make any physical sense to obtain any desirable outcome. For instance, we can take the limit $\rho \rightarrow \infty$ while simultaneously tending L to ∞ as some growing function of ρ . If we do this, then we can choose this function to be such that $\langle (B_\rho^{(2)}\phi)(a_0) \rangle$ converges to any arbitrary value. Similar observations can be made for fields that depend on τ^2 .

These observations show that discrete nonlocal d'Alembert operators converge to the correct continuum limit only on *local* fields—that is, fields with compact supports. This restriction is a form of locality requirement. In the next section, we construct a *local* d'Alembertian operator that converges to the correct continuum limit at $\rho \rightarrow \infty$ on any field, local or nonlocal.

IV. LOCAL D'ALEMBERTIAN IN CAUSAL SETS

Here we construct a d'Alembert operator for causal sets that converges to the correct continuum limit on any field. We prove this convergence for causal sets sprinkled into Minkowski spacetimes. Our operator is *local* in the sense that it operates on local neighborhoods of an event in a causal set. These neighborhoods are defined in Sec. IV B.

The main challenge behind their definition is that it cannot refer to any information about the continuous manifold onto which the causal set is sprinkled. Indeed, a general causal set is not obtained by sprinkling into any spacetime. The causal sets that we deal with in this section are sprinkled into Minkowski spacetime just to prove the d'Alembertian convergence. But since a causal set may not be (directly) associated with any continuous spacetime, a causal set d'Alembertian definition may use only information contained in the causal set itself. Therefore, no part of the operator definition, including the definition of our local neighborhoods, is allowed to use any manifold information either. Fortunately, the definition of our local neighborhoods relies only on proper temporal and spatial distances, both of which can be reliably estimated from the causal set structure alone according to the past results that we review in Sec. IV A.

With the local neighborhoods defined in Sec. IV B, we then show, in Secs. IV C and IV D, how they can be used to define discrete temporal and spatial derivatives, again using only the causal set information. Combining the two

derivatives, we arrive at the definition of our discrete d'Alembertian in Sec. IV E, whose expected deviation from the continuous d'Alembertian, we show, converges to zero in the continuum limit. We discuss some important additional conditions for this convergence to hold in Sec. IV F, and confirm our theoretical results in simulations in Sec. IV G.

For our purposes here, we will treat a general causal set $\mathcal{C}$ as a directed acyclic graph in which links implied by transitivity are removed. That is, if $x, y \in \mathcal{C}$ and $x < y$, meaning that x is in the past of y , then there is a directed link from x to y if and only if there is no other element $z \in \mathcal{C}$ such that $x < z < y$.

A. Measuring proper times and lengths

1. Temporal distances

The measurement of temporal distances in causal sets is relatively straightforward. The proper time $\tau_{\mathcal{C}}(a, b)$ between $a < b$ in $\mathcal{C}$ is proportional to the length $n_{\mathcal{C}}(a, b)$ of the longest chain from a to b , because this chain represents the geodesic between the two events [18]. The convergence of this discrete proper time definition to its continuum limit has been proven for causal sets that are Poisson sprinkled into Minkowski spacetimes $\mathbb{M}^{d+1}$.

Specifically, let us define the Planck time t_P via the sprinkling density ρ as

$$t_P = \rho^{-1/(d+1)}, \quad (29)$$

and the discrete proper time as

$$\tau_{\mathcal{C}}(a, b) \equiv \alpha_d t_P n_{\mathcal{C}}(a, b), \quad (30)$$

where α_d is a constant that depends on the spacetime dimension only. Then it was shown in [18,31,32] that

$$\tau_{\mathcal{C}}(a, b) = \tau_{\mathbb{M}^{d+1}}(a, b) + t_P^{1-\beta_d} \zeta_d, \quad (31)$$

where $\tau_{\mathbb{M}^{d+1}}(a, b)$ is the proper time between $a < b$ in $\mathbb{M}^{d+1}$, $0 \leq \beta_d < 1$, and ζ_d is a random variable with bounded fluctuations and negative average. The values of constants α_d and β_d are known exactly only for $d = 1$. They are $\alpha_1 = 1/\sqrt{2}$ [33] and $\beta_1 = 1/3$ [34]. In higher dimensions, the approximate values of β_d are known to be $\beta_2 \approx 1/4$, $\beta_3 \approx 1/6$, and $\beta_d \approx 0$ for $d \geq 4$ [35], while our numeric experiments in Appendix A suggest that $\alpha_2 = 1/\sqrt{2}$ as well. Even though the exact values of β_d are not known for $d > 1$, the fact that $\beta_d < 1$ in Eq. (31) implies that

$$\lim_{\rho \rightarrow \infty} \tau_{\mathcal{C}}(a, b) = \tau_{\mathbb{M}^{d+1}}(a, b), \quad (32)$$

so that proper times can be recovered with arbitrary precision simply by counting links in $\mathcal{C}$.

2. Spatial distances

The measurement of spatial distances in causal sets is much more challenging, yet the problem has been recently solved. In [21], we introduced an estimator of the distance $d_{\mathbb{M}^{d+1}}(a, b)$ between spacelike separated events a and b in $\mathbb{M}^{d+1}$, based on how much their causal pasts overlap. Given an event $c \in \text{Past}(a, b)$ in the common past of a and b , we define the causal overlap as

$$\mathcal{O}(a, b) \equiv \frac{V[C]}{\min(V[A], V[B]) + V[C]}, \quad (33)$$

where $V[\cdot]$ denotes spacetime volume and the regions A, B, C are defined in terms of Alexandrov intervals $I(x, y) = \text{Past}(x) \cap \text{Future}(y)$ as

$$A = I(a, c) \setminus I(b, c), \quad (34)$$

$$B = I(b, c) \setminus I(a, c), \quad (35)$$

$$C = I(a, c) \cap I(b, c). \quad (36)$$

By definition, $\mathcal{O}(a, b) \in [0, 1]$ and $\mathcal{O}(a, b) = 1$ only if a and b coincide.

The point c in the common past is arbitrary, but in $\mathbb{M}^{d+1}$ one can always choose it such that a and b lie at the same proper time τ_c from c . By placing c at the origin, we place a and b on the same constant-proper-time hyperboloid, so we work in a frame where a and b are simultaneous. In this case, the spatial distance $d_{\mathbb{M}^{d+1}}(a, b)$ is a monotonic function of the overlap $\mathcal{O}_{\mathbb{M}^{d+1}}(a, b)$, and for τ_c much larger than the spatial separation, one finds the asymptotic relation

$$d_{\mathbb{M}^{d+1}}(a, b) = \frac{2}{c_d} \lim_{\tau_c \rightarrow \infty} \tau_c (1 - \mathcal{O}_{\mathbb{M}^{d+1}}(a, b)), \quad (37)$$

where c_d is a dimension-dependent constant; see Ref. [21] for further details. Therefore, in flat spacetime the spatial distance between two spacelike separated events can be read off from the size of the overlap of their past light cones, using any common past point c taken sufficiently far away.

This estimator extends naturally to causal sets via the number-volume correspondence: the volumes $V[A], V[B], V[C]$ in Eq. (33) are replaced by the numbers of elements $N[A], N[B], N[C]$ in the corresponding regions, so that the causal overlap is estimated as

$$\mathcal{O}_{\mathcal{C}}(a, b) \equiv \frac{N[C]}{\min(N[A], N[B]) + N[C]}. \quad (38)$$

For causal sets obtained from a Poisson sprinkling of $\mathbb{M}^{d+1}$, the expectation value of $\mathcal{O}_{\mathcal{C}}(a, b)$ converges to the

continuum overlap $\mathcal{O}_{\mathbb{M}^{d+1}}(a, b)$, so the corresponding distance extracted from Eq. (37) converges to the Minkowski distance. In particular, as shown in [21], this remains true even if a and b are separated by a distance of the order of the Planck length.

B. Defining local neighborhoods

1. Temporal neighborhoods

We define our temporal neighborhoods based on inertial reference frames in $\mathcal{C}$, which are geodesic timelike curves, with one of the events in the geodesic chosen as the origin.

A geodesic timelike curve in a causal set is a numerable sequence of ordered events

$$\mathbb{A} = \{a_i \in \mathcal{C} | \dots < a_{-1} < a_0 < a_{+1} < \dots\} \quad (39)$$

such that for any pair of events $a_j < a_k \in \mathbb{A}$, the sequence $\{a_j, a_{j+1}, \dots, a_k\}$ is a maximal chain connecting the two events. Let us now define a geodesic segment $\mathbb{A}_n$ to be a finite subset of events in $\mathbb{A}$,

$$\mathbb{A}_n = \{a_{-n} < \dots < a_{-1} < a_0 < a_1 < \dots < a_n\}, \quad (40)$$

consisting of $2n + 1$ events, n of which are in the past of a_0 , and the other n are in its future, so $n_{\mathcal{C}}(a_0, a_n) = n_{\mathcal{C}}(a_{-n}, a_0) = n$. An example of such a segment $\mathbb{A}_n$ is shown in orange in Fig. 3, where Mr. White a_0 is at the origin. The boundary elements $a_{\pm n}$ of $\mathbb{A}_n$ are Mr. Blue.

Equations (30) and (31) say that the proper time from a_0 to $a_{\pm n}$ is $\tau_{\mathbb{M}^{d+1}}(a_0, a_{\pm n}) \approx \alpha_d t_P n$, with an error term $\sim t_P^{1-\beta_d}$ that goes to zero in the limit $t_P \rightarrow 0$. Using this, we show in Appendix B that there exists a reference frame $R_{\mathbb{A}}$ in $\mathbb{M}^{d+1}$ in which the event a_0 is at the origin, while the temporal and spatial radial coordinates of $a_{\pm n}$ are

$$t(a_{\pm n}) = \pm \alpha_d t_P n + \mathcal{O}(t_P^{1-\beta_d}), \quad (41)$$

$$r(a_n) = r(a_{-n}) \sim \sqrt{t_P^{2-\beta_d} n}. \quad (42)$$

We now want our segment $\mathbb{A}_n$ to be *local*. Our definition of what it means for it to be *local* is that the temporal and spatial coordinates of its boundaries $a_{\pm n}$ must go to zero in the continuum limit. In view of the last two equations, this requirement is

$$t_P n \ll 1, \quad (43)$$

$$t_P^{2-\beta_d} n \ll 1. \quad (44)$$

However, since $\beta_d < 1$, the former requirement implies the latter.

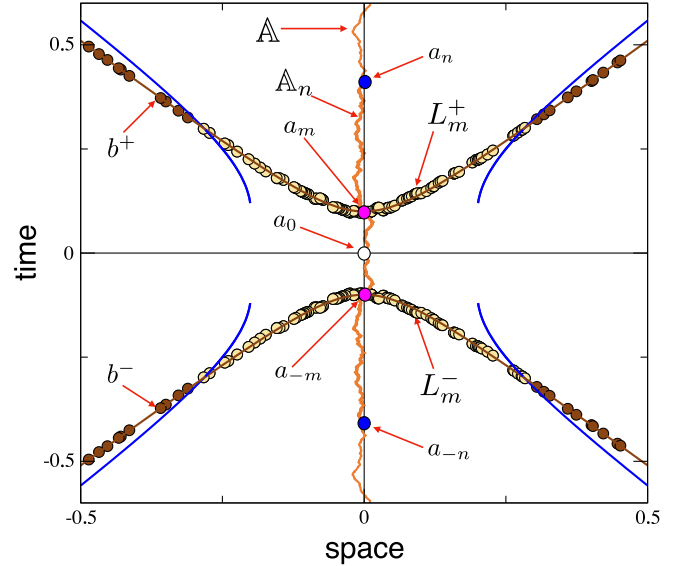


FIG. 3. Defining local neighborhoods. The 100,000 events (not shown) are sprinkled in the $[-0.5, +0.5]^2$ box in $\mathbb{M}^2$ to illustrate the key elements of the local discrete d'Alembertian construction in Sec. IV. In this sprinkling, the point density is thus $\rho = 10^5$ and $t_P = 3.16 \times 10^{-3}$. Event a_0 is at the origin; geodesic $\mathbb{A}$ passing through it is in orange; events $a_{\pm n}$ and $a_{\pm m}$ are blue and pink ($n = 160, m = 40$); the sets of $b^{\pm}$ events that are m links away to the future and past from a_0 are in brown, if they are at proper spatial distances larger than 0.2 from $a_{\pm m}$, and in blond, if they are closer than 0.2 to $a_{\pm m}$; the brown curves are the curves at constant proper time $\alpha_2 t_P m$ from a_0 in $\mathbb{M}^2$; and the blue curves are the curves at constant proper spatial distance 0.2 from $a_{\pm m}$ in $\mathbb{M}^2$. The thick orange events form the set $\mathbb{A}_n$, while the blond events form the sets $L_m^{\pm}$ with $N_m^{\pm} = 100$, so the full local neighborhood $\mathcal{N}_{n,m}(a_0)$ in Eq. (48) is the union of thick orange and blond.

2. Spatial neighborhoods

The basic idea behind our definition of local neighborhoods is locality not only in time, but also in space. That is, we want to define a set of events that are at approximately the same small proper time from a_0 , but also lie within a small spatial distance from one of such events.

Let us consider all events that are m links to the future of a_0 in $\mathcal{C}$ —the brown, blond, and pink events in Fig. 3—and let us denote them all by b^+ . By construction, they are spacelike separated among themselves, and are at proper time

$$\tau_m = \alpha_d t_P m + \mathcal{O}(t_P^{1-\beta_d}) \quad (45)$$

from a_0 , which follows from Eqs. (30) and (31). Henceforth, to make sure that the error term in the last equation is indeed negligible, we require that

$$m \gg t_P^{-\beta_d} \quad \text{or} \quad \tau_m \gg t_P^{1-\beta_d}. \quad (46)$$

Out of the infinite number of events in the set b^+ , we now select one event $a_m \in \mathbb{A}_n$, $m \leq n$ —Mr. Pink in

Fig. 3—which is the event at the intersection of the geodesic segment $\mathbb{A}_n$ and the set b^+ . Next, we sort the b^+ events in the order of increasing proper spatial distance from the reference point a_m , and we define the set L_m^+ —Mr. Blond in Fig. 3—to be the set of N_m^+ events b^+ that are spatially closest to a_m :

$$L_m^+ \equiv \{b_i^+ \in \mathcal{C} | d(b_1^+, a_m) < d(b_2^+, a_m) < \dots < d(b_{N_m^+}^+, a_m)\}. \quad (47)$$

This definition makes sense because it is possible to correctly measure proper lengths at all scales based only on the causal set structure [21].

In Appendix C we calculate N_m^+ as a function of the radius l_c of L_m^+ , defined as the proper distance from a_m to the boundary event in L_m^+ . The sets b^- and L_m^- are defined similarly, except that they are to the past of a_0 .

3. Local neighborhoods

Finally, for fixed integers n and m , chosen to satisfy the convergence requirements that we will derive in Eqs. (53) and (83), we define the *local neighborhood* of event a_0 to be the union of the three sets $\mathbb{A}_n$, L_m^+ , and L_m^- . In other words, once n and m are chosen, the set of events used to compute the d'Alembertian at a_0 is

$$\mathcal{N}_{n,m}(a_0) \equiv \mathbb{A}_n \cup L_m^+ \cup L_m^-. \quad (48)$$

We stress that in what follows, n and m are held fixed. We do *not* take any union over different values of m , so the construction relies only on this finite neighborhood $\mathcal{N}_{n,m}(a_0)$, and not on any foliation of the spacetime. A schematic depiction of $\mathcal{N}_{n,m}(a_0)$ in the continuum two-dimensional Minkowski spacetime is shown in Fig. 3.

The subsets $L_m^\pm$ are made *local* by an appropriate scaling of their sizes $N_m^\pm$, which we determine in Sec. IV D 4. With this choice, the support of the d'Alembertian remains localized around a_0 while still allowing for a bounded statistical error.

We stress that this definition is Lorentz invariant since it is based only on Lorentz-invariant quantities—proper times and lengths—and that these quantities are measured using only the causal set structure.

C. Measuring temporal derivatives $\partial_t^2 \phi$

In this section we introduce the following discrete second-order time derivative operator in causal sets acting on a smooth scalar field ϕ :

$$\mathcal{D}_t \phi(a_0) = \frac{\phi(a_n) + \phi(a_{-n}) - 2\phi(a_0)}{(\alpha_d t_P n)^2}, \quad (49)$$

and show that it converges to the continuous second-order time derivative in the continuum limit:

$$\lim_{t_P \rightarrow 0} \mathcal{D}_t \phi(a_0) = \frac{\partial^2 \phi}{\partial t^2} \Big|_{a_0}. \quad (50)$$

To show this, we first evaluate the scalar field ϕ at the boundary events $a_{\pm n}$ of the local temporal segment $\mathbb{A}_n$ from Sec. IV B 1 via the Taylor series expansion around the origin event a_0 in the reference frame $R_{\mathbb{A}}$:

$$\begin{aligned} \phi(a_{\pm n}) &\approx \phi(a_0) + \frac{\partial \phi}{\partial t} \Big|_{a_0} t(a_{\pm n}) + \nabla \phi|_{a_0} \cdot \mathbf{r}(a_{\pm n}) \\ &\quad + \frac{1}{2} \frac{\partial^2 \phi}{\partial t^2} \Big|_{a_0} t^2(a_{\pm n}), \end{aligned} \quad (51)$$

where $t(a_{\pm n})$ and $\mathbf{r}(a_{\pm n})$ are the temporal and spatial coordinates of $a_{\pm n}$ in $R_{\mathbb{A}}$. Using this expansion, we now evaluate the sum $\phi(a_n) + \phi(a_{-n})$ appearing in Eq. (49). This sum involves the first-order time derivative term with factor $[t(a_n) + t(a_{-n})]$, which is $\sim t_P^{1-\beta_d}$ according to Eq. (41). The spatial gradient term in the sum comes with factor $\sim \sqrt{t_P^{2-\beta_d} n}$ according to Eq. (42). Finally, the factor $[t^2(a_n) + t^2(a_{-n})]/2$ of the second-order time derivative term in the sum is $\approx (\alpha_d t_P n)^2$ according to Eq. (41). Among these three factors— $\sim t_P^{1-\beta_d}$, $\sim \sqrt{t_P^{2-\beta_d} n}$, and $\sim t_P^2 n^2$ —the last one is dominating if $n \gg t_P^{-\frac{2+\beta_d}{3}}$. One can also check that higher-order derivative terms are negligible. Among those, the largest is the fourth-order time derivative. It is $t_P^4 n^4$, which is $\ll t_P^2 n^2$ according to Eq. (43). Combining these considerations, we conclude that

$$\frac{\partial^2 \phi}{\partial t^2} \Big|_{a_0} = \mathcal{D}_t \phi(a_0) + \mathcal{O}(t_P^{\frac{2+\beta_d}{2}} n^{-\frac{3}{2}}) + \mathcal{O}(t_P^2 n^2), \quad (52)$$

where the first and second error terms come, respectively, from the gradient and fourth-order time derivative terms in the Taylor series expansion of $\phi(a_0)$. Both error terms go to zero in the continuum limit $t_P \rightarrow 0$ if

$$t_P^{-\frac{2+\beta_d}{3}} \ll n \ll t_P^{-1}, \quad (53)$$

which is always possible since $\beta_d < 1$ (cf. Sec. IV A 1). There is also the optimal scaling of n as a function of t_P that minimizes the total error in Eq. (52). This scaling is

$$n \sim t_P^{-\frac{6+\beta_d}{7}}, \quad (54)$$

in which case Eq. (52) becomes

$$\frac{\partial^2 \phi}{\partial t^2} \Big|_{a_0} = \mathcal{D}_t \phi(a_0) + \mathcal{O}(t_P^{\frac{2}{7}(1-\beta_d)}). \quad (55)$$

D. Measuring spatial derivatives $\nabla^2\phi$

The last part of our discrete d'Alembertian construction is the definition of the discrete second-order spatial derivative in causal sets. The key idea behind this definition is similar to the one we used to construct the discrete Riemannian Laplacian in Sec. II. Specifically, we simply average the field ϕ over all the events in the spatial neighborhoods $L_m^\pm$ from Sec. IV B 2,

$$\overline{\phi^\pm} = \frac{1}{N_m^\pm} \sum_{i=1}^{N_m^\pm} \phi(b_i^\pm), \quad (56)$$

and then relate the difference

$$\overline{\phi^+} + \overline{\phi^-} - 2\phi(a_0) \quad (57)$$

to $\nabla^2\phi$ and the second-order temporal derivative at a_0 .

To do so, we consider the field value $\phi(b_i^+)$ at event $b_i^+ \in L_m^+$ as the Taylor series expansion of the field at a_0 :

$$\begin{aligned} \phi(b_i^+) &\approx \phi(a_0) + \left. \frac{\partial\phi}{\partial t} \right|_{a_0} t_i^+ + \nabla\phi|_{a_0} \cdot \mathbf{r}_i^+ \\ &+ \frac{1}{2} \partial_{\mu\nu}^2 \phi|_{a_0} x_i^{+\mu} x_i^{+\nu} + \mathcal{O}(x_i^{+\mu} x_i^{+\nu} x_i^{+\lambda}), \end{aligned} \quad (58)$$

and similarly for $b_i^- \in L_m^-$, where $t_i^\pm$ and $\mathbf{r}_i^\pm$ are the temporal and spatial coordinates of $b_i^\pm$ in $R_\mathbb{A}$, and $x_i^\pm = (t_i^\pm, \mathbf{r}_i^\pm)$. This expansion is possible if

$$t_i^\pm \ll 1 \quad \text{and} \quad r_i^\pm \ll 1 \quad (59)$$

for all $i = 1, \dots, N_m^\pm$, which imposes some conditions on m that we determine in Sec. IV D 4.

Assuming these conditions are satisfied, let us see what happens when we start averaging the terms on the right-hand side of Eq. (58) over all events $b_i^+ \in L_m^+$ to obtain the average field $\overline{\phi^+}$ in Eq. (56). The resulting averages are population averages over the set of events in L_m^+ . However, since these events are random, their population averages are also random. They are random variables that depend on a particular realization of the Poisson point process. As in the previous sections, we denote averages over the sprinkling process by brackets $\langle \cdot \rangle$, whereas overbars denote population averages.

1. First-order derivatives

We start by analyzing the population average of the term with the first-order temporal derivative $\partial_t\phi$ in Eq. (58). It comes with the factor

$$\overline{t^+} \equiv \frac{1}{N_m^+} \sum_{i=1}^{N_m^+} t_i^+. \quad (60)$$

The population average $\overline{t^+}$ is a random variable whose average $\langle \overline{t^+} \rangle$ over the Poisson point process is $\langle t^+ \rangle$, where t^+ is the temporal coordinate of an event chosen uniformly at random from the random set L_m^+ and $\langle t^+ \rangle$ is its average. Since events in L_m^+ are uncorrelated, the variance of $\overline{t^+}$ is given by

$$\text{Var}(\overline{t^+}) = \frac{\text{Var}(t^+)}{N_m^+}. \quad (61)$$

Combining these observations, we write that

$$\overline{t^+} = \langle t^+ \rangle + \mathcal{O}\left(\sqrt{\frac{\text{Var}(t^+)}{N_m^+}}\right). \quad (62)$$

Next, due to the rotational symmetry of $\mathbb{M}^{d+1}$, the population averages of terms with odd powers of any spatial coordinate are random variables with zero mean over the Poisson point process. These include, in the first place, the term with the field gradient $\nabla\phi$ in the right-hand side of Eq. (58), which comes with the factor

$$\overline{\mathbf{r}^+} \equiv \frac{1}{N_m^+} \sum_{i=1}^{N_m^+} \mathbf{r}_i^+. \quad (63)$$

The average coordinate $\langle \mathbf{r}^+ \rangle$ of a random node in L_m^+ is $\mathbf{0}$, and therefore $\langle \overline{\mathbf{r}^+} \rangle = \mathbf{0}$. Since events in L_m^+ are uncorrelated, the variance of each individual component of $\overline{\mathbf{r}^+}$ —let it be the x component for concreteness—is given by

$$\text{Var}(\overline{x^+}) = \frac{\langle x^{+2} \rangle}{N_m^+}, \quad (64)$$

where x^+ is the x -coordinate of a random node in L_m^+ , and where we have used that $\langle x^+ \rangle = 0$. Invoking the symmetry of $\mathbb{M}^{d+1}$ again, we also note that the second moment of any spatial coordinate is the same and equal to $\langle x^{+2} \rangle$. Therefore, all the components of vector $\overline{\mathbf{r}^+}$ are of the order

$$\overline{x^+} \sim \mathcal{O}\left(\sqrt{\frac{\langle x^{+2} \rangle}{N_m^+}}\right). \quad (65)$$

2. Second-order derivatives

We now focus on the terms with the second-order derivatives $\partial_{\mu\nu}^2\phi$ of the field in Eq. (58). Consider first the second-order temporal derivative $\partial_t^2\phi$. This term comes with the factor

$$\overline{t^{+2}} \equiv \frac{1}{N_m^+} \sum_{i=1}^{N_m^+} t_i^{+2}. \quad (66)$$

This is a random variable whose average $\langle \overline{t^{+2}} \rangle$ over the Poisson point process is $\langle t^{+2} \rangle$. Therefore, this term dominates the gradient term with factors in Eq. (65) if $\langle t^{+2} \rangle \gg \sqrt{\langle x^{+2} \rangle / N_m^+}$ or, equivalently,

$$N_m^+ \gg \frac{\langle x^{+2} \rangle}{\langle t^{+2} \rangle^2}. \quad (67)$$

This is an important condition on the number of events in L_m^+ : if it holds, the second-order derivative term dominates the gradient term in Eq. (58), so the latter becomes a subleading error term. The evaluation of this condition calls for the evaluation of $\langle t^{+2} \rangle$ and $\langle x^{+2} \rangle$, which we do in Appendix D, a gist of which appears in Sec. IV D 4, where we collect all other error terms. We also note at this point that we do not have to worry about the domination of the second-order derivative terms over the first-order time derivative one, since the latter, coming with factor $\sim \langle t^+ \rangle$, will be canceled in Eq. (57) by its past counterpart from L_m^- , coming with factor $\sim \langle t^- \rangle = -\langle t^+ \rangle$, as we further discuss in Sec. IV D 4 as well.

We now move to the second-order spatial derivatives in Eq. (58). Let us consider the $\partial_x^2 \phi$ term. It comes with the factor

$$\overline{x^{+2}} \equiv \frac{1}{N_m^+} \sum_{i=1}^{N_m^+} x_i^{+2}, \quad (68)$$

whose sprinkling average $\langle \overline{x^{+2}} \rangle$ is $\langle x^{+2} \rangle$. Therefore, according to Eq. (65), this term dominates the gradient term in Eq. (58) if $\langle x^{+2} \rangle \gg \sqrt{\langle x^{+2} \rangle / N_m^+}$ or, equivalently,

$$N_m^+ \gg \frac{1}{\langle x^{+2} \rangle}, \quad (69)$$

which is another condition that N_m^+ must satisfy. We note that according to Appendix D, the values of $\langle x^{+2} \rangle$ and $\langle t^{+2} \rangle$ are of the same order of magnitude, so that both the second-order temporal and spatial derivative terms in Eq. (58) are equally important.

Finally, the second-order derivative terms with mixed derivatives such as $\partial_t \partial_x \phi$ or $\partial_x \partial_y \phi$ come with factors like

$$\overline{t^+ x^+} \equiv \frac{1}{N_m^+} \sum_{i=1}^{N_m^+} t_i^+ x_i^+, \quad (70)$$

$$\overline{x^+ y^+} \equiv \frac{1}{N_m^+} \sum_{i=1}^{N_m^+} x_i^+ y_i^+, \quad (71)$$

and similarly for the rest of the mixed terms. Once again the symmetry of $\mathbb{M}^{d+1}$ implies that $\langle \overline{t^+ x^+} \rangle = 0$ and $\langle \overline{x^+ y^+} \rangle = 0$, and their variances

$$\text{Var}(\overline{t^+ x^+}) = \frac{\langle t^{+2} x^{+2} \rangle}{N_m^+}, \quad (72)$$

$$\text{Var}(\overline{x^+ y^+}) = \frac{\langle x^{+2} y^{+2} \rangle}{N_m^+} \quad (73)$$

are negligible compared to the variance Eq. (64) of the gradient term because $\langle t^{+2} x^{+2} \rangle \ll \langle x^{+2} \rangle$ and $\langle x^{+2} y^{+2} \rangle \ll \langle x^{+2} \rangle$ if t_i^+ , x_i^+ , and y_i^+ are $\ll 1$ for all events b_i^+ in L_m^+ ($i = 1, \dots, N_m^+$). That is, all these terms are negligible with respect to the gradient term in Eq. (58).

3. Higher-order derivatives

The higher-order derivative terms in Eq. (58) come with factors that are population averages of higher powers of temporal and spatial coordinates, or only spatial coordinates, of events in L_m^+ . Let us analyze one class of such terms as an example. Let their factors be

$$\overline{x^{+l} t^{+k}} \equiv \frac{1}{N_m^+} \sum_{i=1}^{N_m^+} x_i^{+l} t_i^{+k} \quad \text{with } l+k \geq 3. \quad (74)$$

The symmetry of $\mathbb{M}^{d+1}$ implies that the ensemble average of terms with odd l values is zero. Similarly, due to the symmetry between the sets L_m^+ and L_m^- , all terms with odd k values have zero ensemble average once averaged over the two sets. This implies that the contribution of these terms to the sum $\overline{\phi^+} + \overline{\phi^-}$ reduces to their statistical error, through their variance

$$\text{Var}(\overline{x^{+l} t^{+k}}) = \frac{\text{Var}(x^{+l} t^{+k})}{N_m^+}, \quad (75)$$

which is negligible compared to the variance Eq. (64) of the gradient term because $\text{Var}(x^{+l} t^{+k}) < \langle x^+ \rangle$.

The only higher-order terms with a nonvanishing ensemble average are those with even values of both l and k . However, thanks to our locality requirement $t_i^+ \ll 1$ and $x_i^+ \ll 1$ for all $i = 1, \dots, N_m^+$, these terms are subdominant with respect to the second-order derivative terms. Therefore, out of these higher-order terms, we consider only the largest fourth-order terms in our recollection of all the terms in the next section.

4. Collecting all the terms and scalings

Here we first recollect all the scaling requirements to m , τ_m , and $N_m^\pm$. The first ones are the locality requirements $t_i^+ \ll 1$ and $x_i^+ \ll 1$ in Eq. (59) for the temporal and spatial coordinates of all events b_i^+ in the local spatial neighborhood L_m^+ ($i = 1, \dots, N_m^+$). We show in Appendix C that these requirements impose the following constraints on τ_m and L_m^+ 's radius l_c (the proper length from a_m to the farthest event in L_m^+):

TABLE I. The terms in the Taylor series expansion of the field in Eq. (58) up to the fourth order, and the expectations and error corrections of their factors emerging from the averaging in Eq. (56). The shown values assume the scaling relations in Eqs. (78)–(83). The terms highlighted in the bold are the ones contributing to the final expression in Eq. (87). The third- and fourth-order terms with zero ensemble averages are omitted.

Term	Factor's population average	Ensemble average	Error correction
$t_i \partial_i \phi$	$\overline{t^+}$	$\langle t^+ \rangle \sim \tau_m$	$\mathcal{O}(\frac{\tau_m}{\sqrt{N_m^+}})$
$x_i \partial_x \phi, y_i \partial_y \phi, z_i \partial_z \phi$	$\overline{x^+}, \overline{y^+}, \overline{z^+}$	0	$\mathcal{O}(\frac{\tau_m}{\sqrt{N_m^+}})$
$t_i^2 \partial_i^2 \phi$	$\overline{t^{+2}}$	$\langle t^{+2} \rangle \sim \tau_m^2$	$\mathcal{O}(\frac{\tau_m^2}{\sqrt{N_m^+}})$
$t_i x_i \partial_i \partial_x \phi, t_i y_i \partial_i \partial_y \phi, t_i z_i \partial_i \partial_z \phi$	$\overline{t^+ x^+}, \overline{t^+ y^+}, \overline{t^+ z^+}$	0	$\mathcal{O}(\frac{\tau_m^2}{\sqrt{N_m^+}})$
$x_i y_i \partial_x \partial_y \phi, y_i z_i \partial_y \partial_z \phi, z_i x_i \partial_z \partial_x \phi$	$\overline{x^+ y^+}, \overline{y^+ z^+}, \overline{z^+ x^+}$	0	$\mathcal{O}(\frac{\tau_m^2}{\sqrt{N_m^+}})$
$x_i^2 \partial_x^2 \phi, y_i^2 \partial_y^2 \phi, z_i^2 \partial_z^2 \phi$	$\overline{x^{+2}}, \overline{y^{+2}}, \overline{z^{+2}}$	$\langle x^{+2} \rangle \sim \tau_m^2$	$\mathcal{O}(\frac{\tau_m^2}{\sqrt{N_m^+}})$
$t_i^3 \partial_i^3 \phi$	$\overline{t^{+3}}$	$\langle t^{+3} \rangle \sim \tau_m^3$	$\mathcal{O}(\frac{\tau_m^3}{\sqrt{N_m^+}})$
$t_i x_i^2 \partial_i \partial_x^2 \phi, t_i y_i^2 \partial_i \partial_y^2 \phi, t_i z_i^2 \partial_i \partial_z^2 \phi$	$\overline{t^+ x^{+2}}, \overline{t^+ y^{+2}}, \overline{t^+ z^{+2}}$	$\langle t^+ x^{+2} \rangle \sim \tau_m^3$	$\mathcal{O}(\frac{\tau_m^3}{\sqrt{N_m^+}})$
$t_i^4 \partial_i^4 \phi$	$\overline{t^{+4}}$	$\langle t^{+4} \rangle \sim \tau_m^4$	$\mathcal{O}(\frac{\tau_m^4}{\sqrt{N_m^+}})$
$t_i^2 x_i^2 \partial_i^2 \partial_x^2 \phi, t_i^2 y_i^2 \partial_i^2 \partial_y^2 \phi, t_i^2 z_i^2 \partial_i^2 \partial_z^2 \phi$	$\overline{t^{+2} x^{+2}}, \overline{t^{+2} y^{+2}}, \overline{t^{+2} z^{+2}}$	$\langle t^{+2} x^{+2} \rangle \sim \tau_m^4$	$\mathcal{O}(\frac{\tau_m^4}{\sqrt{N_m^+}})$
$x_i^2 y_i^2 \partial_x^2 \partial_y^2 \phi, y_i^2 z_i^2 \partial_y^2 \partial_z^2 \phi, z_i^2 x_i^2 \partial_z^2 \partial_x^2 \phi$	$\overline{x^{+2} y^{+2}}, \overline{y^{+2} z^{+2}}, \overline{z^{+2} x^{+2}}$	$\langle x^{+2} y^{+2} \rangle \sim \tau_m^4$	$\mathcal{O}(\frac{\tau_m^4}{\sqrt{N_m^+}})$
$x_i^4 \partial_x^4 \phi, y_i^4 \partial_y^4 \phi, z_i^4 \partial_z^4 \phi$	$\overline{x^{+4}}, \overline{y^{+4}}, \overline{z^{+4}}$	$\langle x^{+4} \rangle \sim \tau_m^4$	$\mathcal{O}(\frac{\tau_m^4}{\sqrt{N_m^+}})$

$$\tau_m \ll 1 \quad \text{and} \quad l_c \ll \sqrt{\tau_m}. \quad (76)$$

We also have the basic asymptotic requirement

$$\tau_m \gg t_P^{1-\beta_d} \quad (77)$$

from Eq. (46), which we use in Appendix C to obtain an approximate expression for N_m^+ as a function of τ_m and l_c . Since we have some freedom in choosing how l_c scales with respect to τ_m , we show in Appendix C that we can use this freedom to let l_c scale proportionally to τ_m ,

$$l_c \sim \tau_m, \quad (78)$$

without violating any scaling requirements. This setting simplifies all the expressions significantly. In particular, the expression for N_m^+ simplifies to

$$N_m^+ \sim t_P^{-(d+\beta_d)} \tau_m^d. \quad (79)$$

Yet we also have the critical requirements in Eqs. (67) and (69). They say that the number N_m^+ of events in the local neighborhood L_m^+ must be sufficiently high to ensure that the ensemble averages of terms involving first-order derivatives are negligible compared to those involving second-order derivatives. These conditions involve $\langle t^{+2} \rangle$ and $\langle x^{+2} \rangle$. We calculate these quantities in Appendix D and

show that under the conditions in Eqs. (77) and (78), they simplify to

$$\langle t^{+2} \rangle \sim \langle x^{+2} \rangle \sim \tau_m^2. \quad (80)$$

Therefore, the requirements in Eqs. (67) and (69) become simply $N_m^+ \gg 1/\tau_m^2$, which in view of Eq. (79) is equivalent to

$$\tau_m \gg t_P^{\frac{d+\beta_d}{d+2}}. \quad (81)$$

This requirement is stronger than the one in Eq. (77), given the values of β_d summarized in Sec. IVA 1. Therefore, under the assumption in Eq. (78), the final allowed scalings of τ_m and m satisfying all the requirements above are

$$t_P^{\frac{d+\beta_d}{d+2}} \ll \tau_m \ll 1, \quad (82)$$

$$t_P^{\frac{2-\beta_d}{d+2}} \ll m \ll t_P^{-1}, \quad (83)$$

which can always be satisfied in the continuum limit $t_P \rightarrow 0$.

We now assume that the scaling relations in Eqs. (78)–(83) are enforced, and under this assumption, we summarize in Table I all the results in Sec. IV D. The table recollects all the terms, up to the fourth order, appearing in

Eq. (58), along with their population-averaged and ensemble-averaged factors and their error corrections emerging from Eq. (56). In the table, we use the fact that, as a consequence of the results in Appendix D, $\langle x^{+l} t^{+k} \rangle \sim \tau_m^{l+k}$ if l is even, and $\langle x^{+l} t^{+k} \rangle = 0$ if l is odd due to the rotational symmetry.

The leading terms in Table I are marked in bold. To see what terms are leading, we observe that in the scaling regime of Eqs. (78)–(83),

$$\tau_m \gg \tau_m^2 \gg \frac{\tau_m}{\sqrt{N_m^+}}, \quad (84)$$

$$\tau_m^2 \gg \tau_m^4 \gg \frac{\tau_m^2}{\sqrt{N_m^+}}, \quad (85)$$

yet which one of the two terms,

$$\tau_m^4 \text{ versus } \frac{\tau_m}{\sqrt{N_m^+}}, \quad (86)$$

is leading is not known *a priori* since this depends on a particular scaling of τ_m as a function of t_P .

The fact that the terms with odd powers of t^+ in Table I are not in bold is because the table reports only the averages $\overline{\phi^+}$ over events $b_i^+ \in L_m^+$, but we have to add $\overline{\phi^+}$ to $\overline{\phi^-}$ according to Eq. (57). Upon this addition, due to the symmetry between L_m^+ and L_m^- (we set $N_m^- = N_m^+$), the expectations with odd powers of t^+ cancel out, e.g., $\langle t^+ \rangle + \langle t^- \rangle = 0$. Therefore, the contributions from these terms to $\overline{\phi^+} + \overline{\phi^-}$ reduce to their error corrections, which are subleading.

Recollecting all these observations and leading terms in Table I, we finally conclude that

$$\begin{aligned} \langle t^{+2} \rangle \frac{\partial^2 \phi}{\partial t^2} \Big|_{a_0} + \langle x^{+2} \rangle \nabla^2 \phi|_{a_0} \\ = \overline{\phi^+} + \overline{\phi^-} - 2\phi(a_0) + \mathcal{O}\left(\frac{\tau_m}{\sqrt{N_m^+}}\right) + \mathcal{O}(\tau_m^4). \end{aligned} \quad (87)$$

As a reminder, this equation assumes that all the scaling relations in Eqs. (78)–(83) are satisfied. As evident from Table I, the first error term in Eq. (87) arises from terms involving first-order derivatives, whereas the second error term comes from fourth-order derivatives. In general, the dominant error is the larger of these two terms. However, as is also the case with Eq. (52), there is an optimal scaling of m for which both error terms are of the same order of magnitude. This scaling is

$$m \sim t_P^{\frac{6-\beta_d}{d+6}}, \quad (88)$$

which lies in the regime of Eq. (83). With this scaling of m , Eq. (87) becomes

$$\begin{aligned} \langle t^{+2} \rangle \frac{\partial^2 \phi}{\partial t^2} \Big|_{a_0} + \langle x^{+2} \rangle \nabla^2 \phi|_{a_0} = \overline{\phi^+} + \overline{\phi^-} - 2\phi(a_0) \\ + \mathcal{O}(t_P^{\frac{4(d+\beta_d)}{d+6}}), \end{aligned} \quad (89)$$

whereas its contribution to the error of the d'Alembertian becomes

$$\mathcal{O}(t_P^{\frac{2(d+\beta_d)}{d+6}}). \quad (90)$$

E. Measuring the d'Alembertian

We finally have everything in place to define our discrete d'Alembertian. Combining Eqs. (49), (55), and (89), we get

$$\begin{aligned} \square \phi|_{a_0} &= \left(-\frac{\partial^2}{\partial t^2} + \nabla^2 \right) \phi \Big|_{a_0} \\ &= -\left(1 + \frac{\langle t^{+2} \rangle}{\langle x^{+2} \rangle} \right) \frac{\phi(a_n) + \phi(a_{-n}) - 2\phi(a_0)}{(\alpha_d t_P n)^2} \\ &\quad + \frac{\overline{\phi^+} + \overline{\phi^-} - 2\phi(a_0)}{\langle x^{+2} \rangle} + \mathcal{O}(t_P^{\frac{2}{d}(1-\beta_d)}), \end{aligned} \quad (91)$$

where the error term is the maximum of the two optimal errors in Eqs. (55) and (90).

Using the scaling relations in Eqs. (78)–(83), we show in Appendix D that $\langle x^{+2} \rangle$ can be set equal to τ_m^2/C for any constant $C > 0$, in which case $\langle t^{+2} \rangle = (d/C + 1)\tau_m^2$, so Eq. (91) becomes

$$\square \phi(a_0) = \mathcal{B}\phi(a_0) + \mathcal{O}(t_P^{\frac{2}{d}(1-\beta_d)}), \quad (92)$$

where our discrete d'Alembert operator $\mathcal{B}$ is defined as the following sum of the discrete second-order temporal and spatial derivatives:

$$\mathcal{B} = -(C + d + 1)\mathcal{D}_t + C\mathcal{D}_s, \quad \text{where} \quad (93)$$

$$\mathcal{D}_t \phi(a_0) = \frac{\phi(a_n) + \phi(a_{-n}) - 2\phi(a_0)}{(\alpha_d t_P n)^2}, \quad (94)$$

$$\mathcal{D}_s \phi(a_0) = \frac{\overline{\phi^+} + \overline{\phi^-} - 2\phi(a_0)}{(\alpha_d t_P m)^2}. \quad (95)$$

Equation (92) implies the desired convergence:

$$\lim_{t_P \rightarrow 0} \mathcal{B}\phi(a_0) = \square \phi(a_0). \quad (96)$$

It is important to note that the definition of the discrete d'Alembertian in Eq. (93) is not unique. Here, we have chosen a symmetric version that uses information from

both the future and past of event a_0 . Another option is to define a retarded version that uses information only from the past of a_0 . We could also define the discrete temporal derivative $\mathcal{D}_t$ by averaging over all the events in $\mathbb{A}_n$, similar to the averaging over all the events in $L_m^\pm$ in the definition of the discrete spatial derivative $\mathcal{D}_s$. We also note that the averaging in the definition of $\mathcal{D}_s$ appears unavoidable as it is needed to suppress the contributions from the first-order derivative terms in Eq. (58).

F. Choosing the geodesic segment $\mathbb{A}_n$

Our derivation in the previous section is based on the assumption that the field ϕ admits a Taylor expansion within a small neighborhood δx^μ of a given event x_0^μ . That is, we assume that in a given reference frame we can expand ϕ to the first order as

$$\phi(x_0^\mu + \delta x^\mu) \approx \phi(x_0^\mu) + \partial_\nu \phi(x_0^\mu) \delta x^\nu + \mathcal{O}(\delta^2 x^\mu). \quad (97)$$

Assuming $\partial_\nu \phi(x_0^\mu) < \infty$, this approximation holds for sufficiently small displacements δx^μ from the event x_0^μ .

Consider now another reference frame related to the first one via a Lorentz boost with Lorentz factor γ . In the boosted frame, the first-order derivatives are linear combinations of the original $\partial_\nu \phi(x_0^\mu)$, with coefficients that are proportional to γ . Since $\gamma > 1$ can be arbitrarily large, there exist frames in which these derivatives become arbitrarily large. In the continuum, this is not a problem because we can always choose the neighborhood to be small enough for the Taylor expansion to remain well defined in all boosted frames. In the discrete settings, however, the neighborhood size is bounded from below by the Planck scale t_P . Consequently, there will be geodesic segments $\mathbb{A}_n$ for which the expansions in Eqs. (51) and (58) are not well defined. This happens whenever $\gamma t_P \gg 1$.

Therefore, we need a criterion to decide whether a given geodesic segment $\mathbb{A}_n$ is suitable for evaluating the d'Alembertian of the field ϕ within our approach. Assume that the segment is indeed suitable, so that the first-order partial derivatives are bounded and satisfy $|\partial_\nu \phi| \lesssim t_P^{-1}$. Then, using the results of Sec. IV C, we can write

$$\phi(a_n) - \phi(a_{-n}) \approx 2t(a_n) \frac{\partial \phi}{\partial t} \Big|_{a_0} + \frac{1}{3} t^3(a_n) \frac{\partial^3 \phi}{\partial t^3} \Big|_{a_0}. \quad (98)$$

Since the partial derivatives are bounded, it follows that

$$\Delta \phi_n = |\phi(a_n) - \phi(a_{-n})| \sim n t_P \ll 1. \quad (99)$$

This inequality can be interpreted as a sufficient condition for the segment suitability both in the $t_P \rightarrow 0$ limit ($\Delta \phi_n \rightarrow 0$) and for a finite sprinkling with $t_P > 0$

($\Delta \phi_n < 1$), the latter providing a practical way to select an appropriate geodesic segment in a given causal set.

We stress that this condition uses only information available in the causal set, so it is Lorentz invariant. Since the construction of the local neighborhood is Lorentz invariant and the d'Alembertian acting on the field is also Lorentz invariant, any admissible geodesic segment in $\mathcal{C}$ with bounded variation of the field leads to the same estimate of the d'Alembertian. Different segment choices affect only the prefactor in the error terms in Eq. (92), but not their scaling with the Planck time t_P .

We also note that the discussed constraints on the choice of geodesics bear some conceptual similarities with the nonlocal d'Alembertian constraints in Sec. III. There, one is restricted to work only with fields with compact support; here, we can work with any fields but only in frames in which the field is slowly varying, simply because the Taylor expansion does not hold in other frames.

G. Numerical simulations

Finally, we perform numerical simulations to confirm the convergence of our discrete d'Alembertian in Eq. (93) to its continuous brother.

To do so, we apply the discrete d'Alembertian to causal sets in $\mathbb{M}^{2+1}$ generated at increasing densities ranging from $\rho = 10^4$ to $\rho = 3 \times 10^6$. For each density ρ , we sprinkle events uniformly at random into a rectangular box in $\mathbb{M}^{2+1}$ with temporal coordinate $t \in [-1, +1]$ and spatial coordinates $x, y \in [-1/2, +1/2]$. The event a_0 , at which the d'Alembertian of the field is evaluated, is at the origin. To select event a_m , we first need to fix m for each value of ρ . To do so, we follow Eq. (88), and set $m \propto \rho^{23/96}$. The event a_m is then m links to the future of a_0 . There are many such events, but we select a_m to be the one with the smallest radial coordinate to ensure that the boundaries of the local spatial neighborhood L_m^+ do not cross the boundaries of our simulation box. To define L_m^+ , we set $l_c = 2\tau_m/3$, where $\tau_m = m/\sqrt{2\rho}^{1/3}$. According to Eq. (C4) in Appendix C, this choice bounds the radial coordinates of events in L_m^+ to $\sqrt{80}\tau_m/9$, again ensuring that all these events lie within our simulation box. The past spatial neighborhood L_m^- is set up symmetrically, while for the temporal neighborhood, we set $n = m$, $a_n = a_m$, and $a_{-n} = a_{-m}$. Finally, to evaluate the constant C in Eq. (93), we observe that according to Appendix D, if $d = 2$, then

$$\langle t^{+2} \rangle = \frac{1}{3} (\cosh^2 \chi_c + \cosh \chi_c + 1) \tau_m^2, \quad (100)$$

$$\langle x^{+2} \rangle = \frac{1}{6} (\cosh^2 \chi_c + \cosh \chi_c - 2) \tau_m^2. \quad (101)$$

It follows then from Eq. (C2) that our choice of $l_c = 2\tau_m/3$ sets $\cosh \chi_c = 11/9$, so that $\langle t^{+2} \rangle = 301\tau_m^2/243$ and $\langle x^{+2} \rangle = 29\tau_m^2/243$, that is, $C = 243/29$.

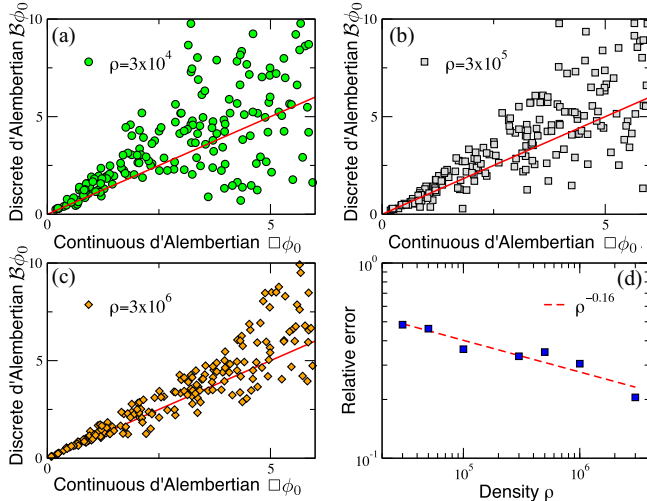


FIG. 4. D'Alembertian simulations. (a)–(c) The values of the discrete d'Alembertian $\mathcal{B}\phi(a_0)$ in Eq. (93) acting on the field $\phi = e^{-a(t^2-x^2-y^2)}$ at $t = x = y = 0$ in the $(2+1)$ -dimensional Minkowski spacetime $\mathbb{M}^{2+1}$ vs the continuous value $\square\phi(a_0) = 6a$ for three different point densities ρ and $m = (2/5)\sqrt{2/ap^{23/96}}$. For each value of ρ , 200 random values of $a \in [0, 1]$ are used, and for each value of a , a separate realization of the Poisson point process is sampled. (d) The average relative error of the discrete vs continuous d'Alembertian, defined as $\langle |(\mathcal{B}\phi(a_0)/\square\phi(a_0)) - 1| \rangle$, for different values of ρ . For each value of ρ , the relative error is averaged of 200 random values of a and sprinkling realizations—one sprinkling for each a . The red dashed line is the fit $\rho^{-0.16}$, close to the analytical prediction $\mathcal{O}(\rho^{-2(d+\beta_d)/[(d+6)(d+1)]}) = \mathcal{O}(\rho^{-3/16}) \approx \mathcal{O}(\rho^{-0.19})$ from Eq. (90). Equations (90) vs (92) yields the error scaling in this case because the field has the zero first-order derivatives at the origin, so Eq. (52) contributes error only via its second error term, which is equivalent to the error term in Eq. (90) since $n = m$.

With these settings, we evaluate our discrete d'Alembertian in Eq. (93) acting on the field

$$\phi(t, x, y) = e^{-a\tau^2} = e^{-a(t^2-x^2-y^2)}, \quad (102)$$

on which nonlocal d'Alembertians from Sec. III do not converge. The continuous d'Alembertian applied to this field at the origin yields $\square\phi(a_0) = 6a$. We sample multiple values of a uniformly at random in $[0, 1]$, and for each sampled value and each density ρ , we generate a different realization of the Poisson point process in the simulation box, and measure $\mathcal{B}\phi(a_0)$ in each such sprinkling.

Figure 4 shows the results. Figures 4(a)–4(c) juxtapose the measured values of the discrete d'Alembertian $\mathcal{B}\phi(a_0)$ against the continuous values $\square\phi(a_0) = 6a$ for 200 different values of a and three different point densities ρ . We observe that at higher densities, the discrete d'Alembertian values are concentrated more tightly around the continuous values. Figure 4(d) illustrates this further by showing the

relative error between the discrete and continuous values. As predicted by our analysis in the previous sections, this error goes to zero polynomially as a function of the point density, implying the convergence.

V. CONCLUDING REMARKS

Lorentzian geometry differs fundamentally from Riemannian geometry. In Riemannian geometry, the triangle inequality ensures that pairs of points that are close to a third point are also close to each other. In Lorentzian geometry, two points at a given proper time from a third point can be arbitrarily far apart from each other spatially. This observation suggests that causal set theory is inherently nonlocal, and that fundamental differential operators in the theory, including the d'Alembertian, should be nonlocal as well.

In this work, we demonstrated that causal set theory can, in fact, be formulated as a local theory, with discrete differential operators defined via a careful construction of local neighborhoods in a causal set. Based on this insight, we defined a local discrete d'Alembert operator, which converges the continuous d'Alembertian in the continuum limit, which we have shown both analytically and numerically.

In addition to convergence, this operator has several other nice properties. First, since it operates on local neighborhoods, it converges on any field, as opposed to nonlocal operators that converge only on local fields that have compact supports. Second, our operator is self-averaging because when it acts on any field, the deviation of its values from their expectations in a single realization of the sprinkling process converges to zero in the continuum limit $\rho \rightarrow \infty$. Finally, our operator is manifestly Lorentz-invariant because everything in its definition, including the local neighborhood, is formulated solely in terms of Lorentz-invariant quantities, such as proper times and proper lengths, which can be measured using only the causal set structure.

Generalizing, these results demonstrate that discrete approximations to continuous objects whose definitions are based on the notion of locality—by these we mean differential operators in the first place—are possible even in inherently nonlocal theories, of which causal set theory is a good example. It is thus plausible that similar ideas might help defining and proving convergence of discrete versions of more sophisticated locality-based objects, like spacetime curvature appearing in the Einstein field equations.

ACKNOWLEDGMENTS

We thank David Meyer for a thorough reading of our manuscript and Yasaman Yazdi for interesting discussions on a preliminary version of this work. We also acknowledge Rafael Sorkin, Fay Dowker, and an anonymous referee for very helpful comments leading to Sec. IV F. M. B. acknowledges support from Grant No. PID2022-137505NB-C22

funded by MCIN/AEI/10.13039/501100011033. M. B. acknowledges the ICREA Academia award, funded by the Generalitat de Catalunya. D. K. acknowledges NSF Grants No. IIS-1741355 and No. CCF-2311160.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

APPENDIX A: NUMERICAL EVALUATION OF CONSTANT α_2

Equation (31) allows us to estimate the value of the constant α_d in higher dimensions. By taking the average of $\tau_{\mathbb{M}^{d+1}}$ for a fixed number of causal set hops n and an arbitrary parameter α^* , we can write

$$\frac{\rho^{1/(d+1)}}{n} \langle \tau_{\mathbb{M}^{d+1}} | n \rangle - \alpha^* = \alpha_d - \alpha^* - \frac{\rho^{d_d} \langle \zeta_d \rangle}{n}. \quad (\text{A1})$$

The left-hand side of this equation can be directly measured from numerical simulations. Therefore, the empirical value of α_d will be the value of α^* that makes the left-hand side behave as a perfect power law proportional to n^{-1} . Figure 5 shows results for causal sets generated by sprinkling events in $\mathbb{M}^3$ at density $\rho = 3 \times 10^6$. As can be seen, the best power law is obtained when $\alpha^* = \alpha_2 = 1/\sqrt{2}$. Lower values of α^* lead to curves that bend upwards, whereas higher values result in curves decaying faster than a power law.

Unfortunately, in the case of $\mathbb{M}^4$, the maximum attainable value of n grows as $\rho^{1/4}$. Therefore, to explore large values of n , we need to reach densities $\rho \sim 10^9$, which is

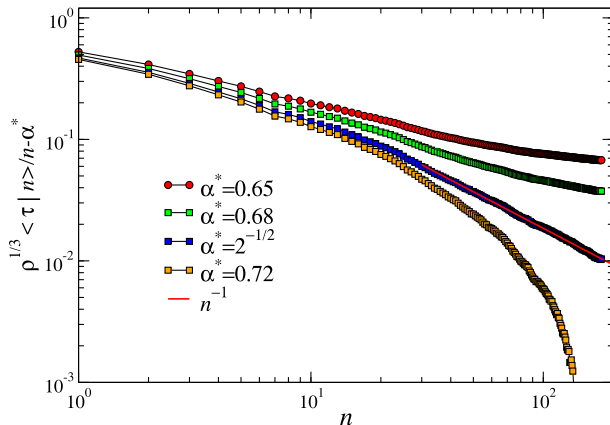


FIG. 5. Estimation of parameter α_2 .

beyond our current computational capability. However, preliminary results at a density of $\rho \sim 10^6$ support our conjecture that $\alpha_4 = 1/\sqrt{2}$.

APPENDIX B: FLUCTUATIONS OF GEODESIC PATHS IN $\mathcal{C}$ AND THE DEFINITION OF THE REFERENCE FRAME R_A

Let a and b be two timelike related events in $\mathbb{M}^{d+1}$. Without loss of generality, we work in the 2D timelike plane spanned by the origin and these events, and place event a on the temporal axis at $(\tau_a, \mathbf{0})$, so that its temporal coordinate coincides with the proper time from the origin to a . We place event b at temporal coordinate $t_b < \tau_a$ and spatial radial offset r_b , so that the proper time from the origin to b satisfies

$$\tau_b^2 = t_b^2 - r_b^2. \quad (\text{B1})$$

If we now go to a new reference frame in which b is at the origin, the coordinates of a in this frame are (t_{ab}, r_b) , and the proper time between a and b is

$$\tau_{ab}^2 = t_{ab}^2 - r_b^2, \quad t_{ab} = \tau_a - t_b. \quad (\text{B2})$$

We thus have three proper times, τ_a , τ_b , and τ_{ab} , and three coordinate relations. Eliminating t_b and t_{ab} from these relations gives an expression for the spatial offset r_b purely in terms of the three proper times. A straightforward calculation yields

$$r_b^2 = \frac{1}{4\tau_a^2} [(\tau_a + \tau_b + \tau_{ab})(\tau_a + \tau_b - \tau_{ab}) \times (\tau_a - \tau_b + \tau_{ab})(\tau_a - \tau_b - \tau_{ab})]. \quad (\text{B3})$$

This is the standard expression for the spatial separation in terms of the three edge lengths of the timelike “triangle” formed by the origin, a , and b .

We now assume that events a , b , and the origin of coordinates are three events of a geodesic path in $\mathcal{C}$. Equation (31) relates the continuum proper time between two events in $\mathbb{M}^{d+1}$ to the proper time measured in the causal set by counting the number of hops in the geodesic path connecting them, up to an error of order $t_p^{1-\beta_d}$. To simplify notation, we denote by

$$\tilde{\tau} \equiv \alpha_d t_p n \quad (\text{B4})$$

the proper time in $\mathcal{C}$ associated with a geodesic path of n hops, so that the continuum proper time can be written as

$$\tau = \tilde{\tau} + \delta\tau, \quad (\text{B5})$$

where $\delta\tau$ is the correction term of order $t_p^{1-\beta_d}$ from Eq. (31). The key point is that if a and b lie on the same geodesic

path in $\mathcal{C}$, then the causal-set proper times along that path add:

$$\tilde{\tau}_a = \tilde{\tau}_b + \tilde{\tau}_{ab}. \quad (\text{B6})$$

In the continuum, an exact geodesic would also satisfy $\tau_a = \tau_b + \tau_{ab}$, which would make $r_b = 0$ in Eq. (B3). Thus, any nonzero r_b for a point b on the causal-set geodesic originates from the small corrections $\delta\tau$ that break this exact equality.

Substituting $\tau_i = \tilde{\tau}_i + \delta\tau_i$ ($i = a, b, ab$) into Eq. (B3) and using the additivity relation (B6) to simplify the result, we obtain

$$\begin{aligned} r_b^2 &= \frac{2\tilde{\tau}_a\tilde{\tau}_b\tilde{\tau}_{ab}}{(\tilde{\tau}_a + \delta\tau_a)^2} \left(1 + \frac{\delta\tau_a + \delta\tau_b + \delta\tau_{ab}}{2\tilde{\tau}_a}\right) \\ &\times \left(1 + \frac{\delta\tau_a + \delta\tau_b - \delta\tau_{ab}}{2\tilde{\tau}_b}\right) \left(1 + \frac{\delta\tau_a - \delta\tau_b + \delta\tau_{ab}}{2\tilde{\tau}_{ab}}\right) \\ &\times (\delta\tau_a - \delta\tau_b - \delta\tau_{ab}). \end{aligned} \quad (\text{B7})$$

Here $\delta\tau_b$ and $\delta\tau_{ab}$ are independent correction terms, while $\delta\tau_a$ is constrained by Eq. (B6) and by the corresponding relation for the continuum proper times.

Assuming that the relative corrections are small,

$$\frac{\delta\tau_i}{\tilde{\tau}_i} \ll 1, \quad i = a, b, ab, \quad (\text{B8})$$

we can expand Eq. (B7) to first order in the $\delta\tau_i$. To leading order we obtain

$$r_b^2 \approx 2\tilde{\tau}_b \left(1 - \frac{\tilde{\tau}_b}{\tilde{\tau}_a}\right) (\delta\tau_a - \delta\tau_b - \delta\tau_{ab}). \quad (\text{B9})$$

Since $\tilde{\tau}_b \leq \tilde{\tau}_a$, the factor in parentheses is bounded between 0 and 1. Using that each $\delta\tau_i$ is of order $t_P^{1-\beta_d}$, we conclude that for a generic event b on the causal-set geodesic but different from a ,

$$r_b \sim \sqrt{\tilde{\tau}_b t_P^{1-\beta_d}}. \quad (\text{B10})$$

Combining this with the relation between the proper and coordinate times of event b , we find

$$t_b = \tilde{\tau}_b + O(t_P^{1-\beta_d}). \quad (\text{B11})$$

Finally, we identify event b with a_0 , event a with a_n , and the origin of coordinates with event a_{-n} . Applying a translation that places event a_0 at the origin of coordinates while keeping the temporal axis parallel to the original one leads to a reference frame $R_{\mathbb{A}}$ where the temporal coordinates of events $a_{\pm n}$ satisfy

$$t(a_{\pm n}) = \pm \alpha_d t_P n + O(t_P^{1-\beta_d}), \quad (\text{B12})$$

whereas their spatial offsets are

$$r(a_n) = r(a_{-n}) \sim \sqrt{\alpha_d t_P^{2-\beta_d} n}. \quad (\text{B13})$$

APPENDIX C: NUMBER OF EVENTS IN $L_m^\pm$

Let us compute the expected number of events $N_m^\pm$ at a proper distance $l < l_c$ from the event $a_m \in L_m^+$. Suppose that a_m is located at the position $(\tau_m, \mathbf{0})$ in $R_{\mathbb{A}}$. Thus, the region of $\mathbb{M}^{d+1}$ with proper length $l < l_c$ is bounded by the surface $r^2 = l_c^2 + (t - \tau_m)^2$, which in hyperbolic coordinates becomes

$$\cosh \chi = \frac{l_c^2 + \tau^2 + \tau_m^2}{2\tau\tau_m}. \quad (\text{C1})$$

The proper time of events in L_m^+ is also τ_m . Therefore, such events have a hyperbolic coordinate $\chi < \chi_c$, where χ_c is the solution of

$$\cosh \chi_c = 1 + \frac{l_c^2}{2\tau_m^2}. \quad (\text{C2})$$

Therefore, setting an upper bound for the proper length is equivalent to setting an upper bound for the coordinate χ , which translates to the following upper bounds for the temporal and spatial coordinates in the set L_m^+ :

$$t_c^+ = \tau_m \cosh \chi_c = \tau_m + \frac{l_c^2}{2\tau_m}, \quad (\text{C3})$$

$$x_c^+ = \tau_m \sinh \chi_c = l_c \sqrt{1 + \frac{l_c^2}{4\tau_m^2}}. \quad (\text{C4})$$

From these equations, we see that the locality condition $t_i^+ \ll 1$ and $x_i^+ \ll 1$ for $i = 1, \dots, N_m^+$ implies that

$$\tau_m \ll 1 \quad \text{and} \quad l_c \ll \sqrt{\tau_m}. \quad (\text{C5})$$

We note that we can choose any scaling of l_c as long as these conditions are satisfied.

The volume element of $\mathbb{M}^{d+1}$ in hyperbolic coordinates is

$$dV = \tau^d d\tau \sinh^{d-1} \chi d\chi d\Omega_{d-1}, \quad (\text{C6})$$

where $d\Omega_{d-1}$ is the volume element of the $(d-1)$ -sphere. The expected number of events from L_m^+ within an infinitesimal neighborhood of $(\tau, \chi, \Omega_{d-1})$ is then

$$dN^+ = \rho P_d^+(\tau) \tau^d d\tau \sinh^{d-1} \chi d\chi d\Omega_{d-1}, \quad (\text{C7})$$

where $P_d^+(\tau)$ is the probability that an event at proper time τ from the origin belongs to the set L_m^+ . The total number of expected events in L_m^+ is then given by

$$N_m^+ = \rho \Omega_{d-1} \int_0^\infty P_d^+(\tau) \tau^d d\tau \int_0^{\chi_c} \sinh^{d-1} \chi d\chi. \quad (\text{C8})$$

The probability $P_d^+(\tau)$ is not known for arbitrary m . However, as follows from Eqs. (30) and (31), in the regime $m \gg t_P^{1-\beta_d}$ ($\tau_m \gg t_P^{1-\beta_d}$), the proper time of events from L_m^+ is $\tau_m + O(t_P^{1-\beta_d})$. In this regime, we expect $P_d^+(\tau)$ to be a function sharply peaked at τ_m , with fluctuations of the order $O(t_P^{1-\beta_d})$, so that the integral over proper times can be approximated as

$$\int_0^\infty P_d^+(\tau) \tau^d d\tau \propto \tau_m^d t_P^{1-\beta_d}, \quad (\text{C9})$$

and that the number of events in L_m^+ becomes

$$N_m^+ = \sigma_d t_P^{-(d+\beta_d)} \tau_m^d {}_2F_1\left(\frac{1}{2}, \frac{d}{2}, \frac{d+2}{2}, -\sinh^2 \chi_c\right) \sinh^d \chi_c, \quad (\text{C10})$$

where

$$\sinh \chi_c = \sqrt{\cosh^2 \chi_c - 1} = \frac{l_c}{\tau_m} \sqrt{1 + \frac{l_c^2}{4\tau_m^2}}, \quad (\text{C11})$$

${}_2F_1()$ is the Gauss hypergeometric function, and σ_d a constant that can be measured from numerical simulations.

Let us now set l_c to be proportional to the proper time τ_m , so that χ_c becomes a constant. This setting satisfies the requirements in Eq. (C5). With this setting, the number of nodes in L_m^+ scales as

$$N_m^+ \sim t_P^{-(d+\beta_d)} \tau_m^d. \quad (\text{C12})$$

We note that $N_m^+ \gg 1$ in the regime $t_P^{1-\beta_d} \ll \tau_m \ll 1$.

APPENDIX D: EVALUATION OF $\langle t^{+2} \rangle$ AND $\langle x^{+2} \rangle$

In hyperbolic coordinates, the temporal coordinate of an event in L_m^+ is $t^+ = \tau^+ \cosh \chi$, where τ^+ and χ are statistically independent random variables, the former in

the range $(0, \infty)$ and the latter in $(0, \chi_c)$, where χ_c is given by Eq. (C2). The second moment of t^+ can then be computed as

$$\langle t^{+2} \rangle = \frac{\int_0^\infty P_d^+(\tau) \tau^{d+2} d\tau \int_0^{\chi_c} \cosh^2 \chi \sinh^{d-1} \chi d\chi}{\int_0^\infty P_d^+(\tau) \tau^d d\tau \int_0^{\chi_c} \sinh^{d-1} \chi d\chi}. \quad (\text{D1})$$

In the regime $\tau_m \gg t_P^{1-\beta_d}$ we obtain

$$\langle t^{+2} \rangle = f_t(\chi_c) \tau_m^2, \quad \text{where} \quad (\text{D2})$$

$$f_t(\chi_c) = \frac{{}_2F_1\left(-\frac{1}{2}, \frac{d}{2}, \frac{d+2}{2}, -\sinh^2 \chi_c\right)}{{}_2F_1\left(\frac{1}{2}, \frac{d}{2}, \frac{d+2}{2}, -\sinh^2 \chi_c\right)}. \quad (\text{D3})$$

We note that $f_t(\chi) > 1$ for $\chi > 0$. Using a similar calculation, it is easy to show that the variance of t^+ scales as $\text{Var}(t^+) \sim \langle t^{+2} \rangle$.

Finally, the second moment of individual space coordinates follows from the identity $t^{+2} = \tau^{+2} + r^{+2}$ and the symmetry of $\mathbb{M}^{d+1}$, so that

$$\langle x^{+2} \rangle = \frac{1}{d} (\langle t^{+2} \rangle - \tau_m^2), \quad (\text{D4})$$

which, if $\tau_m \gg t_P^{1-\beta_d}$, is

$$\langle x^{+2} \rangle = f_x(\chi_c) \tau_m^2 = \frac{1}{d} (f_t(\chi_c) - 1) \tau_m^2. \quad (\text{D5})$$

If $l_c \sim \tau_m$, then χ_c is a constant, and

$$\langle t^{+2} \rangle \sim \langle x^{+2} \rangle \sim \tau_m^2. \quad (\text{D6})$$

By choosing the constant $C > 0$ in $l_c = C\tau_m$, we can fix the constant C_x in

$$\langle x^{+2} \rangle = C_x \tau_m^2 \quad (\text{D7})$$

to any positive value (e.g., 1), in which case $\langle t^{+2} \rangle$ becomes

$$\langle t^{+2} \rangle = C_t \tau_m^2, \quad \text{where} \quad (\text{D8})$$

$$C_t = C_x d + 1. \quad (\text{D9})$$

-
- [1] S. Surya, *Living Rev. Relativity* **22**, 5 (2019).
 [2] L. Bombelli, J. Lee, D. Meyer, and R. D. Sorkin, *Phys. Rev. Lett.* **59**, 521 (1987).

- [3] R. D. Sorkin, *Int. J. Theor. Phys.* **36**, 2759 (1997).
 [4] D. Rideout and R. D. Sorkin, *Phys. Rev. D* **61**, 024002 (1999).

- [5] R. D. Sorkin, in *Lectures on Quantum Gravity*, edited by A. Gomberoff and D. Marolf (Springer, New York, 2005), pp. 305–327.
- [6] R. D. Sorkin, *Causal Sets: Discrete Gravity*, in *Lectures on Quantum Gravity* (Springer, New York, 2005), pp. 305–327.
- [7] J. Henson, *The Causal Set Approach to Quantum Gravity, Approaches to Quantum Gravity: Toward a New Understanding of Space, Time and Matter* (Cambridge University Press, Cambridge, UK, 2009).
- [8] F. Dowker and S. Surya, *Classical Quantum Gravity* **23**, 1381 (2006).
- [9] S. Surya, *Gen. Relativ. Gravit.* **44**, 2149 (2012).
- [10] S. W. Hawking, A. R. King, and P. J. McCarthy, *J. Math. Phys. (N.Y.)* **17**, 174 (1976).
- [11] D. Malament, *J. Math. Phys. (N.Y.)* **18**, 1399 (1977).
- [12] S. Major, D. Rideout, and S. Surya, *J. Math. Phys. (N.Y.)* **48** (2007).
- [13] G. Brightwell, J. Henson, and S. Surya, *Classical Quantum Gravity* **25**, 105025 (2008).
- [14] S. Surya, *Classical Quantum Gravity* **29**, 132001 (2012).
- [15] M. Saravani and S. Aslanbeigi, *Classical Quantum Gravity* **31**, 205013 (2014).
- [16] A. Belenchia, D. M. Benincasa, and F. Dowker, *Classical Quantum Gravity* **33**, 245018 (2016).
- [17] L. Machet and J. Wang, *Classical Quantum Gravity* **38**, 015010 (2020).
- [18] G. Brightwell and R. Gregory, *Phys. Rev. Lett.* **66**, 260 (1991).
- [19] D. Rideout and P. Wallden, *Classical Quantum Gravity* **26**, 155013 (2009).
- [20] A. Eichhorn, S. Surya, and F. Versteegen, *Classical Quantum Gravity* **36**, 105005 (2019).
- [21] M. Boguñá and D. Krioukov, *Phys. Rev. D* **110**, 024008 (2024).
- [22] R. D. Sorkin, *Does Locality Fail at Intermediate Length-Scales* (Cambridge University Press, Cambridge, England, 2009), pp. 26–43.
- [23] D. M. T. Benincasa and F. Dowker, *Phys. Rev. Lett.* **104**, 181301 (2010).
- [24] F. Dowker and L. Glaser, *Classical Quantum Gravity* **30**, 195016 (2013).
- [25] L. Glaser, *Classical Quantum Gravity* **31**, 095007 (2014).
- [26] S. Aslanbeigi, M. Saravani, and R. D. Sorkin, *J. High Energy Phys.* **06** (2014) 024.
- [27] F. Dowker, R. Liu, and D. Lloyd-Jones, *Classical Quantum Gravity* **42**, 205006 (2025).
- [28] M. Belkin and P. Niyogi, *J. Comput. Syst. Sci.* **74**, 1289 (2008); *Learning Theory* 2005.
- [29] M. Hein, J.-Y. Audibert, and U. v. Luxburg, *J. Mach. Learn. Res.* **8**, 1325 (2007).
- [30] M. Hamidouche, L. Cottatellucci, and K. Avrachenkov, *J. Theor. Probab.* **36**, 46 (2023).
- [31] B. Bollobás and G. Brightwell, *Trans. Am. Math. Soc.* **324**, 59 (1991).
- [32] E. Bachmat, *Contemp. Math.* **458**, 347 (2008).
- [33] B. Logan and L. Shepp, *Adv. Math.* **26**, 206 (1977).
- [34] J. Baik, P. Deift, and K. Johansson, *J. Am. Math. Soc.* **12**, 1119 (1999).
- [35] M. Lässig, *J. Phys. Condens. Matter* **10**, 9905 (1998).