On The Horizon Problem in Discrete Quantum Cosmology

Benjamin F. Dribus

Abstract

The horizon problem in the cosmology of the early universe involves uniformity of structure difficult to explain without a mechanism for interchange of information among spacetime regions that appear causally unrelated. This problem was a principal motivation for the inflationary hypothesis developed by Guth and Linde, and remains one of the most important empirical factors cited in support of this hypothesis. Inflation posits the contribution of a so-called *inflaton field* driving vast exponential expansion of spacetime during a very brief period in the early universe, leading to large separation of regions that were previously sufficiently close to interact. While the hypothetical inflaton field possesses certain unconventional properties, the mechanism otherwise fits into the conventional framework of relativistic cosmology, neither suggesting nor utilizing any essentially new ideas beyond general relativity and quantum field theory. Recent interest in alternatives to inflation is spurred partly by unforeseen implications of simple versions of the inflationary mechanism and partly by a desire to explain the origin and details of spacetime structure in a context more fundamental than that of fields on Lorentzian manifolds. In this paper, I examine a completely different method of approaching the horizon problem, via the hypothesis of discrete spacetime structure. Under this approach, early interchange of information leading to uniformity among distant regions need not involve spacetime expansion at all, although such expansion is favored on other grounds. Rather, this phenomena can be explained by an abrupt decrease in the sizes of causal horizons due to a *phase transition* in the sense of random graph theory, whereby the qualitative properties of a network of events changes suddenly under gradual change of a connectivity parameter. Numerical evidence is examined and results are proven for certain classes of finite universes.

1 General Horizon Problem

Figure 1 gives a schematic illustration of the horizon problem in the cosmological setting, independent of any specific geometric or topological assumptions except for the existence of a past boundary of the spacetime region involved, suggestively denoted "beginning of time" in the figure, and the absence of large-scale closed causal curves. The entire shaded area represents the spacetime region under consideration, with darker shading, curves, and labels representing additional information detailed below. Two spacelike-separated events x and y, represented by black nodes, are observable from a third event z. For simplicity, x and y are taken to belong to the *past causal horizon* of z; i.e., the boundary of its causal past. In relativistic language, x and y are connected to z by null curves. In general, however, the illustration only requires that x and y be on *or inside* this horizon, so that causal influence may reach z from x and y. This is the meaning of the statement that x and y are "observable from z." The *apparent* locations of x and y from the perspective of an observer at z may differ from their actual locations for a variety of reasons, including expansion or contraction of spacetime, gravitational lensing, variation in the speed of light, and spatial anisotropy. The apparent locations of x and y are represented by white nodes. In a general causal sense, z is a "common descendant" of x and y. One may ask if x and y also share any "common ancestors;" i.e., if there

exist any events observable from both x and y. In other words, do the pasts of x and y overlap, or is each past beyond the other's horizon?

In this particular case, the question of overlapping pasts is easily answered "yes," if one is privileged to view the entire diagram from "outside." The medium-shaded regions representing the causal pasts of x and y overlap in the dark-shaded region near the bottom of the diagram. Hence, this region represents common ancestors of x and y. However, the existence of such an overlap may not be obvious from the limited "internal" perspective of an observer at z. In fact, it appears from this perspective that x and y do *not* share common ancestors, based on knowledge about the "age of the universe" at z and local information about how the sizes of causal pasts scale in the immediate vicinity of z. The apparent causal past of z, and the apparent causal pasts of x and y, based on their apparent locations, are represented in the figure by the regions bounded by dashed lines. The apparent pasts of x and y are disjoint. In other words, it appears from the perspective of z as if observers at x and y could look back in time without observing any common events. Stated in terms of horizons, the entire past of y appears to be beyond the horizon of x, and the entire past of x appears to be beyond the horizon of y.

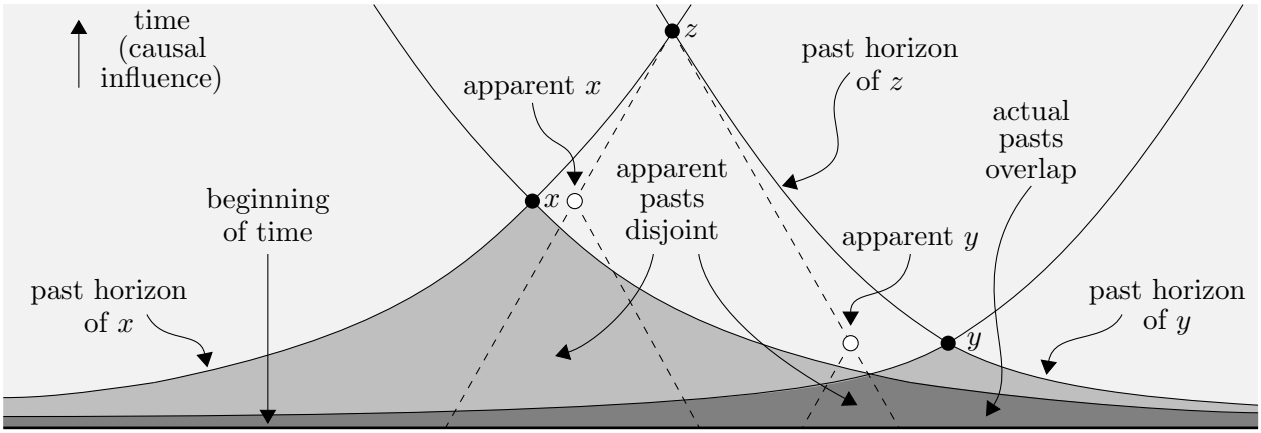

Figure 1: Horizon problem: apparent pasts of x and y are disjoint, suggesting no common history; actual pasts overlap due to markedly different early structure.

The erroneous impression that the pasts of x and y are disjoint, suggested by naïve interpretation of local data available near z, arises from the *incompleteness of information about global structure* encoded in this data. The question of whether the pasts of x and y overlap is a *nonlocal* question. What is missing, from the perspective of an observer at z, is direct evidence of a drastic change in the structure of the early universe as one approaches the "beginning of time," represented in the figure by the horizontal "stretching" of the causal pasts of x and y. Here, the "beginning" is drawn in a manner suggesting nontrivial spatial extent, and no information regarding expansion or contraction of spacetime, curvature, the speed of light, or anisotropy is specified. In this particular figure, the change in structure appears to involve a vast increase in the rate of information exchange; i.e., the speed of light, as one approaches the "beginning." However, the diagram is merely schematic, and other descriptions of this structural change are possible. For example, one may shrink the "beginning" to a point-like singularity represented by a single node; the resulting figure suggests an "inflationary epoch" rather than an increase in the speed of light.

Figure 2 illustrates a conventional inflationary explanation of the horizon problem. For simplicity, the entire spacetime region shown is taken to emanate from an initial singularity; i.e., a "Big-Bang-like" event. However, the same general picture applies to each "bubble" of spacetime in the

context of *eternal inflation,* in which a vast region of inflating spacetime produces myriad such bubbles of "ordinary" spacetime. The figure suggests a spatially finite universe; in particular, each horizontal slice appears to represent a finite "space" at a particular instant in time in a frame of reference analogous to the CMB frame. However, the "edges" of the diagram are included merely to indicate how the rate of expansion of spacetime varies; the same structural meaning applies to infinite inflationary models. Of course, these "edges" do not represent physical boundaries even in the finite case, but merely limitations of drawing on a two-dimensional page; if one compares the entire figure to an upside-down silhouette of the McMinnville "flying saucer," then a better schematic is given by the surface of the "saucer;" i.e., the surface of revolution of the "edges" about the vertical axis through the singularity.

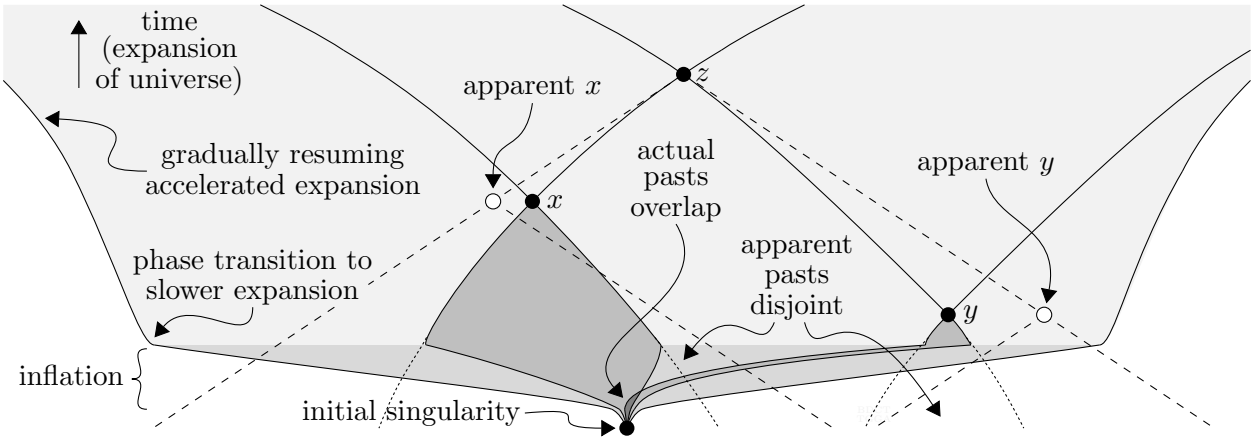

Figure 2: Inflationary view of the horizon problem.

Soon after the singularity, the rate of expansion increases drastically, with each spatial slice much larger than those a short time earlier. This extreme expansion is the inflationary epoch. The figure vastly understates its abruptness and brevity; the entire process is thought to have occurred between 10^{-36} and 10^{-32} seconds after the initial singularity, and to have increased the volume of space by a factor of perhaps 10^{60} or more. Following this explosive interval, a phase transition occurs and the spacetime decays into ordinary spacetime, with a much slower rate of expansion. This phase transition is indicated in the figure by the change from light to lighter shading and the nearly corner-shaped vertical turns in the "edge" curves representing the expansion rate. The apparent pasts of x, y, and z are represented by regions bounded by dashed lines, and the actual pasts of x and y are represented by medium-shaded regions. The dotted lines represent "adjusted apparent pasts," discussed below. Regardless of its varying rate, spacetime expansion implies that the actual past of each event is smaller than its apparent past. Hence, the apparent locations of the events x and y from the perspective of z are spatially more distant from each other than their actual locations. Although the apparent pasts of x and y are much larger than their actual pasts, they are disjoint. The comparatively small actual pasts overlap because their earliest parts are crammed together in a tiny region near the singularity. This description looks "backward in time," since it compares apparent and actual pasts. An alternative description is given by looking "forward in time" from the immediate future of the singularity. From this viewpoint, small adjacent spatial regions inside the overlap are blasted far apart by the inflationary process, producing regions that are much larger but much more widely separated. The gradual outward bending of the "edge" curves near the top of the figure represents the standard view that the expansion of spacetime resumed gradual acceleration around 5 billion years ago.

For the purposes of this paper, the fact of principal interest illustrated by Figures 1 and 2 is that *changes in the structure of the early universe can cause related events to appear unrelated from a remote future perspective.* How would one infer that such changes in structure had actually occurred? Why would an observer at event z, in either figure, conclude that events x and y might actually share common ancestors despite the fact that their apparent pasts are disjoint? Before answering this question, is useful to briefly distinguish alternative meanings of "apparent pasts" in this context, and to clarify that the horizon problem is not merely a result of ignoring obvious evidence that the most naïve interpretation of "apparent" structure is inaccurate. For example, the "apparent" location of the events x and y indicated by the white nodes in Figure 2 ignores the standard theory of spacetime expansion entirely, by assuming that the past of the event z is just as large as it would be in a static spacetime. The foremost evidence for such expansion is the redshifting of light from distant objects, which is a very robust and general phenomenon. On the basis of such knowledge, one might adjust the locations of x and y to their "correct" values, and adjust their apparent pasts by extrapolating backward in time the curves defining their actual pasts after the end of inflation. The resulting "adjusted apparent pasts" are represented in the figure by the dotted curves. Like the "naïve apparent pasts" represented by the dashed lines, they fail to overlap, so this adjustment does not enable the observer at z to detect the common history of x and y. A similar statement holds true for actual observations of redshift, so the horizon problem is not merely a trivial discrepancy between static and expanding models of spacetime.

A subtler and more significant reason why one might conclude that x and y share common ancestors despite contrary structural evidence is that regions near x and y observable from z might share similarities suggesting a common source of information. Not all possible types of similarities would qualify; for example, many basic similarities may be attributed directly to the common validity of universal physical laws. However, other types of similarities could reasonably be ascribed only to interactions. As an illustration, one could imagine an unrealistically powerful telescope revealing the presence of giraffes living on planets in the vicinities of both x and y. Regardless of apparent non-overlap of causal pasts suggested by large-scale structural observations from the perspective of z, it would be much more reasonable in this scenario to assume common history for x and y than to assume that the observed similarities arose by coincidence. The forgoing discussion provides sufficient background for a general statement of the horizon problem:

Horizon Problem: *In certain systems involving causal structure, particularly the observable universe at large scales, local details of spatially separated regions suggest common history, while limited data about global structure suggests the opposite.*

Two general responses to this problem are to either ascribe the *"local details of spatially separated regions"* to new mechanisms, or to hypothesize global structure much different than what the *"limited knowledge"* available directly suggests. In the context of early cosmology, the second response is much more reasonable, because the local details involved are analogous to the above illustration involving giraffes; they seem to resist explanation via any credible mechanism. The inflationary hypothesis of Guth and Linde and the alternative discussed in this paper are both examples of the second response. In the case of inflation, the new global structure involved is the brief, profound expansion of spacetime during the hypothesized inflationary epoch. The mechanism suggested in this paper is much different: it involves a steady decrease in the local connectivity of spacetime that naturally produces an abrupt global change in the sizes of causal horizons due to a particular type of discrete phase transition arising from purely combinatorial considerations.

2 Horizon Problem in Discrete Causal Theory

Figure 3 illustrates a discrete causal version of the horizon problem, analogous to the general schematic diagram from Figure 1. Spacetime events, called *elements* in this context, are represented by nodes, and direct causal relationships between pairs of spacetime events, called *relations,* are represented by edges connecting nodes, with influence flowing "up the page." A relation between two elements u and v is represented algebraically by the expression $u \prec v$, where the *precursor symbol* $\prec$ means that u directly influences v. The relationship $u \prec v$ is analogous to the familiar "less than" relation $n < n + 1$ between a pair of consecutive integers. Elements and relations are not regarded as residing *inside* some ambient "space" or "spacetime," represented, for instance, by the white background of the page. Rather, "space" and "time" are taken to be notions that emerge from the collective properties of families of elements and relations. In particular, the family of all elements in the figure is a set denoted by D, and the family of all relations between pairs of elements of D is a *binary relation* on D, denoted by the same precursor symbol $\prec$ used to denote individual relations such as $u \prec v$. The pair $(D, \prec)$, called a *directed set* or *directed graph,* serves as a discrete model of spacetime, whose structure is entirely specified by direct causal relationships. The fact that such discrete causal structure can serve as a realistic model of actual spacetime is a consequence of theorems of Stephen Hawking and David Malament, proven in the late 1970s, which demonstrate that large classes of relativistic spacetime manifolds can be approximated up to arbitrary precision by discrete directed sets. The assertion that spacetime geometry actually does emerge from causal structure, which I refer to as the *causal metric hypothesis,* was first suggested around 1980 by a number of different researchers, including 't Hooft, Myrheim, and Sorkin.

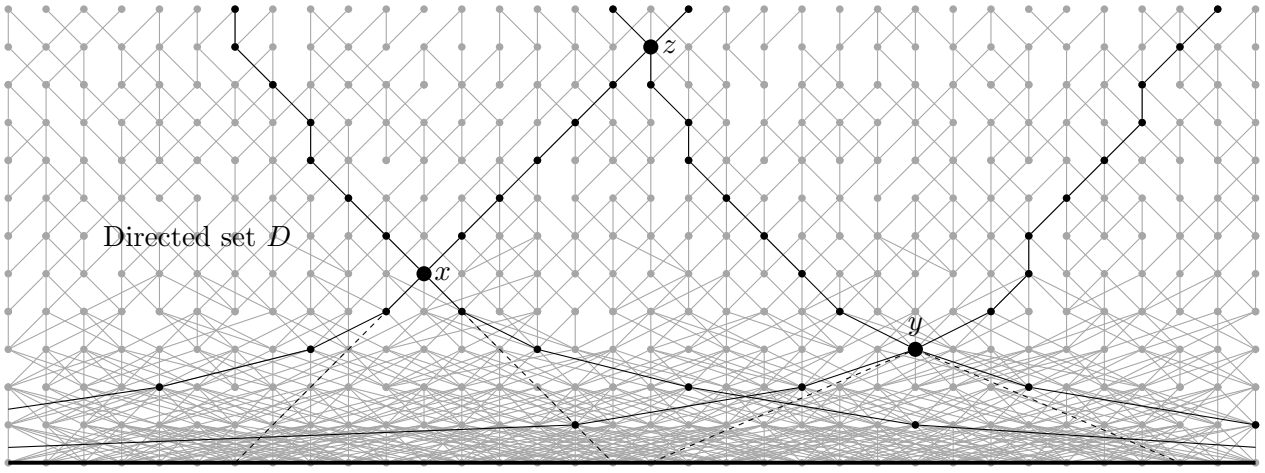

Figure 3: Discrete causal version of the setup from Figure 1: directed set D in which causal structure becomes sparser in the future direction.

The term "directed set" is more convenient than "directed graph," and will be used throughout the remainder of the paper, but the reader should be aware that this term is conventionally assigned a more specific meaning involving common upper bounds in purely mathematical settings. It is often useful to denote a directed set by just D, rather than $(D, \prec)$. Included in the figure are elements x, y, and z, which are analogous to the events labeled x, y, and z in Figure 1. The existence of an upward-directed path from x to z indicates that x influences z; in other words, x belongs to the causal past of z. In this case, the influence is indirect, since an entire sequence of individual relations is needed to reach from x to z. Similarly, y belongs to the causal past of z, and again the influence is indirect. Note that causal influence is represented only by *upward-directed* paths;

for example, one may trace a path from x to z and thence to y, but the second part of this path is downward-directed, so no causal influence passes from x to y. The convention of assigning the direction of influence as "upward" only applies to *acyclic* directed sets; i.e., directed sets which do not possess discrete analogues of relativistic closed causal curves.

It is necessary to explain here a subtlety of discrete causal structure arising from the distinction between *direct* and *indirect* influence, already mentioned above. If spacetime is continuous, then an event x never directly influences a future event z, because one may always specify an intervening event w. This is ultimately due to the *interpolative property* of the real-number continuum $\mathbb{R}$, the basic mathematical object from which most of the structural attributes of the manifolds used to model continuum spacetime arise. The interpolative property states that given any pair of real numbers x and z, where $x < z$, there exists a third real number w such that $x < w < z$. The simplest choice is to take w to be the average of x and z. If spacetime is discrete, however, direct influence *can* occur; i.e., there may exist causal relationships $x \prec z$ that *cannot* be "subdivided into closer relationships" $x \prec w$ and $w \prec z$. This is just one of several ways in which discrete spacetime models can encode a *broader* variety of interesting structure than their "larger" continuum counterparts! Failure of the interpolative property is already familiar in elementary settings; for example, there is no integer between integers n and $n+1$. Returning to the discrete causal setting, given two events x and z, and assuming that x influences z in some way, three basic scenarios are possible. First, x may influence z directly, without modification by any intervening event w. Second, x may influence z indirectly via one or more intervening events, without any direct influence. Third, both modes of influence may occur. These possibilities are illustrated in Figure 4.

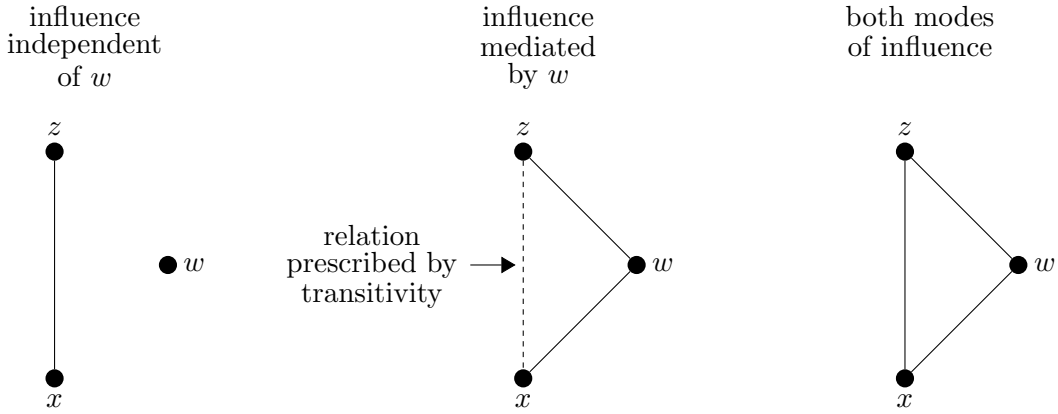

Figure 4: Possible modes of influence between related elements x and z in a directed set.

A directed set D is called *transitive* if any pair of "consecutive" relations $x \prec w$ and $w \prec z$ in D implies the existence of a "spanning relation" $x \prec z$, evoking vector addition, as shown in the right-hand diagram in Figure 4. Such a "spanning relation" is called *reducible,* because it may be roughly viewed as the "sum" or "product" of the two "shorter" relations $x \prec w$ and $w \prec z$. Transitivity is a natural property for familiar linearly-ordered number systems such as the integers, rational numbers, or real numbers. For example, given three integers l, m, and n, the relations $l < m$ and $m < n$ imply the relation $l < n$. Transitivity is also a natural property for causal relationships if one is uninterested in the distinction between direct and indirect causation. For example, the causal relationships of familial descent between grandparent and parent and between parent and child imply a corresponding relationship between grandparent and grandchild. However, in any setting where the distinction between direct and indirect causation is important, transitive relations are inadequate to model many structural scenarios. For example, transitive relations cannot distinguish

between the modes of influence illustrated in the middle and right-hand diagrams of Figure 4, because transitivity *prescribes* the existence of a reducible relation $x \prec z$ whenever there exist relations $x \prec w$ and $w \prec z$, regardless of whether there is actually direct influence between x and z. Relaxing the transitive property allows the freedom to add the relation $x \prec z$ if there is direct influence, and to leave out this relation if there is not. Hence, in this paper, individual relations $x \prec z$ always imply direct influence, and indirect influence is encoded by complex chains of relations. This is in contrast to causal set theory, which limits attention to transitive relations.

The directed set D illustrated in Figure 3 exhibits a degree of randomness, but is much less random than one would expect in a realistic setting. In particular, the 442 nodes representing elements of D appear to form a square lattice, in terms of how they are arranged on the page. By contrast, Figure 5 shows a random arrangement of 442 nodes in an equal-sized planar region, without specifying any relations. Such random "sprinklings" of elements are often used in causal set theory to construct causal sets that mimic large-scale properties of familiar geometries, such as the geometry of Minkowski spacetime $\mathbb{R}^{3+1}$, where the "3" in the superscript represents spatial dimensions and the "1" represents the temporal or causal dimension. In particular, one could construct a causal set from Figure 5 by taking the background of the page to represent a two-dimensional version $\mathbb{R}^{1+1}$ of Minkowski spacetime, and adding an *induced relation* from x to z whenever z belongs to the causal future of x. The figure shows three distinguished elements x, y, and z, along with their causal pasts and futures. In this case, pasts and futures are simply "light cones;" i.e., there is no deviation of light rays from their instantaneous directions at x, y, and z, because $\mathbb{R}^{1+1}$ is its "own tangent space." In particular, there is no distortion of causal pasts as one moves in the negative temporal direction, because Minkowski spacetime is time-symmetric. This is an obvious reason why four-dimensional Minkowski spacetime is a poor global model of the actual universe.

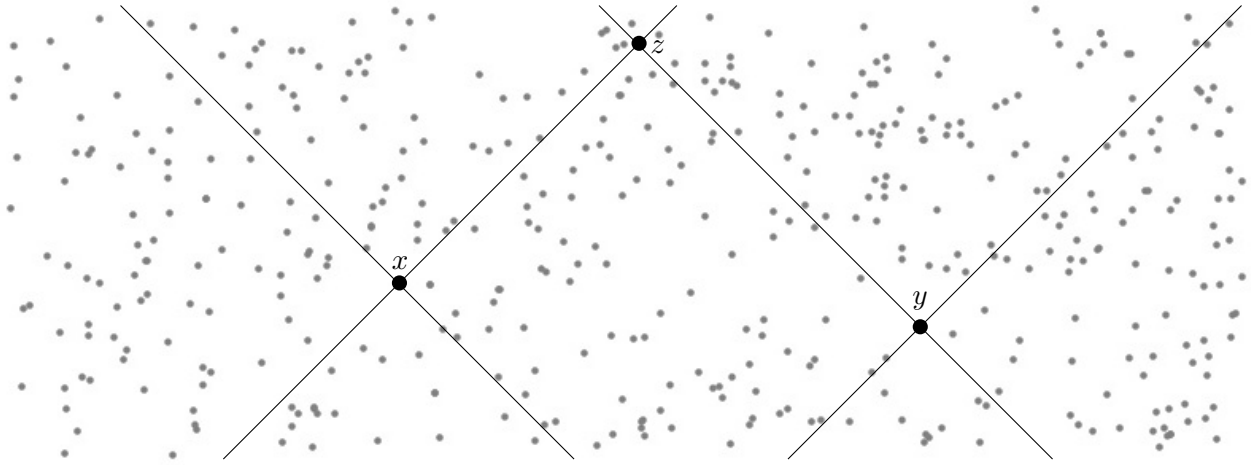

Figure 5: Elements "sprinkled" randomly in Minkowski spacetime.

Returning to the directed set D illustrated in Figure 3, the apparent lattice-like arrangement of the nodes representing its elements does not, by itself, imply a lattice-like *structure* for D, since the structural properties of D are completely determined by its elements and relations, without regard to how it is represented graphically. However, D is actually quite close to being "lattice-like" in its abstract structure, at least in certain important ways. For example, it exhibits an approximate version of locality, reflected by the fact that each element is unambiguously "close" to a few neighbors, and "far" from other elements of D, where distance is measured by the length of paths, not necessarily upward-directed, and where "unambiguously" means that if there is one

"short" path between two elements, then there are generally many other paths that are nearly as "short." It is also possible to identify "generations" in D, where the nth generation consists roughly of elements that may be reached by an upward directed path of length n from the bottom of the diagram. This means that D has an obvious "preferred frame of reference;" i.e., it is "very non-Lorentz invariant." This property might a priori elicit some concern from the perspective of recovering relativistic behavior, but is in fact irrelevant to the phenomena investigated in this paper, which is not frame-dependent. This topic is revisited below.

Because this paper is principally concerned with how the basic structure of spacetime may change as one moves from past to future, it is natural to consider relatively thin "slices" of the directed set D, and to compare how slices near the beginning of time differ from slices later in time. Figure 6 illustrates a slice Δ^3 near the middle of D, where the superscript 3 indicates that Δ^3 has three generations of elements. Taken by itself, Δ^3 is almost connected, though there is an isolated node in the eleventh column. This node is not causally related to any other element of Δ^3, so from a strict structural perspective, assigningthis node to the third generation of Δ^3 is an arbitrary choice. For most purposes, however, it will be reasonable to pretend that Δ^3 has 3 well-defined generations.

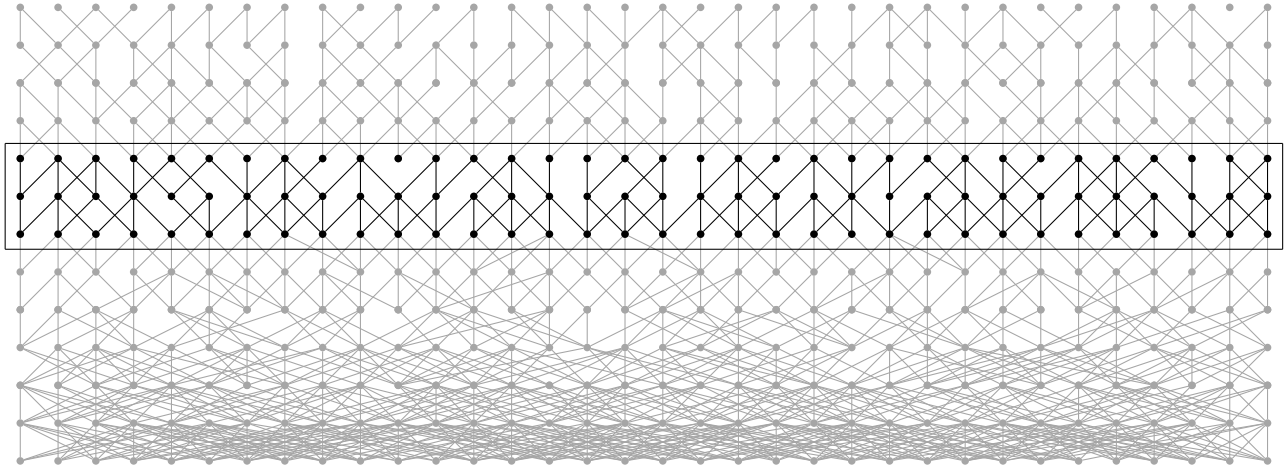

Figure 6: Slicing the directed set D from Figure 3 into layers.

In the context of the horizon problem, a question of obvious interest is how the sizes of causal pasts of elements scale as one moves back in time. More specifically, it is of interest to compare long-term scaling behavior to short-term scaling behavior, since significant differences between scaling behaviors over different time frames lead to misleading impressions about causal horizons. Returning to the general horizon problem illustrated in Figure 1, the reason why the causal pasts of x and y appear to be disjoint from the perspective of z is because causal pasts grow much more quickly as one moves back in time near the beginning of time. To be precise, it is actually the *relative* scaling behavior compared to the global expansion or contraction of spacetime that is pivotal in this context. For example, in the inflationary scenario illustrated in Figure 2, causal pasts grow more slowly as one approaches the initial singularity, but spacetime contracts so abruptly in this direction that the relative scaling behavior is qualitatively similar to that of Figure 1. Having sliced spacetime into layers, one may analyze short-term scaling behavior by examining a few consecutive layers, or long-term scaling behavior by examining many consecutive layers. In either case, the problem hinges on the question of how quickly causal pasts grow within individual layers as one moves back in time. For example, referring again to Figure 6, one may examine the sizes of the causal pasts of elements of the third generation of Δ^3. These sizes may be described as the numbers

of first and second-generation elements that influence a given third-generation element. A simpler, purely spatial, measure of size is given by confining attention to first-generation elements; i.e., by examining the sizes $|\sigma(x)|$ of the *initial cross sections* $\sigma(x)$ of the causal pasts of third-generation elements x of Δ^3. Figure 7 illustrates a few such cross sections.

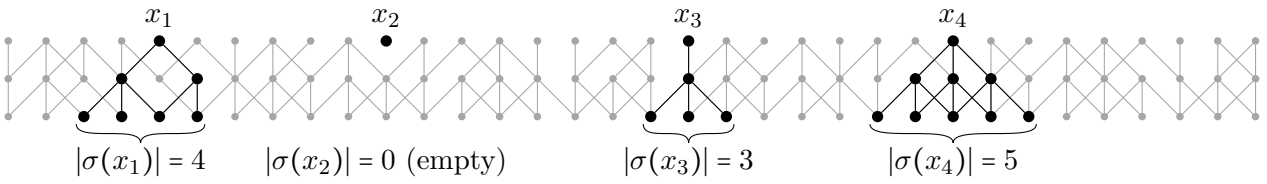

Figure 7: Initial cross sections in Δ^3.

Figure 8 illustrates a similar layer in a directed set induced by sprinkling elements into two-dimensional Minkowski spacetime $\mathbb{R}^{1+1}$.

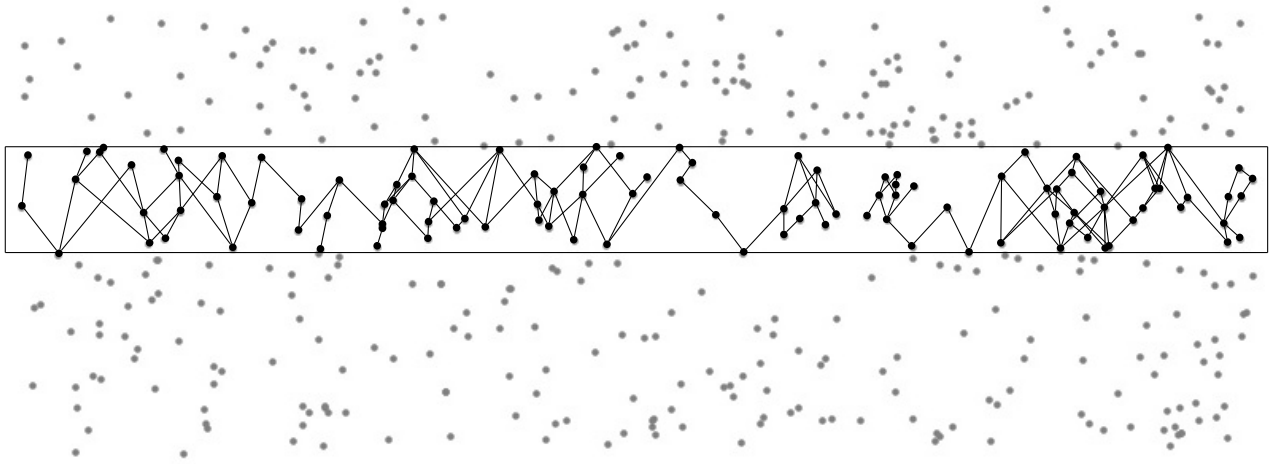

Figure 8: Layer in a directed set induced by sprinkling in Minkowski spacetime.

While such "random layers" generally do not possess well-defined generations, the same qualitative ideas introduced above still apply. In particular, one may examine the sizes of initial cross-sections of *maximal* elements in such a layer. The top diagram in figure 9 illustrates such cross-sections for a few selected elements x_1, x_2, x_3, and x_4 of the layer illustrated in figure 8. It is convenient for the purpose of visualization to rearrange the positions of the nodes illustrating the elements in this layer, as shown in the bottom diagram in the figure. This changes nothing about the structure of the layer, since this structure is completely encoded by the elements and relations themselves. This random layer appears somewhat more complicated than the more-regular layer Δ^3 illustrated in figure 7, but its qualitative properties appear rather similar. At present, we will assume that the similarities between such layers are sufficient that analysis of layers resembling Δ^3 will be at least somewhat relevant to describing the behavior of more random layers.

Before adopting a more precise viewpoint, I offer a qualitative description of how the sizes of causal horizons actually change between different layers of a universe like the one illustrated in Figure 6. In such a universe, the numbers relations in typical layers, or at least the numbers of relations per unit volume, decrease as one moves toward the future. As one would expect, the sizes of causal horizons within each layer decrease as well, but not in the same way as the numbers of relations. Instead, the these horizons tend to shrink abruptly when the number of relations reaches a certain threshold, even

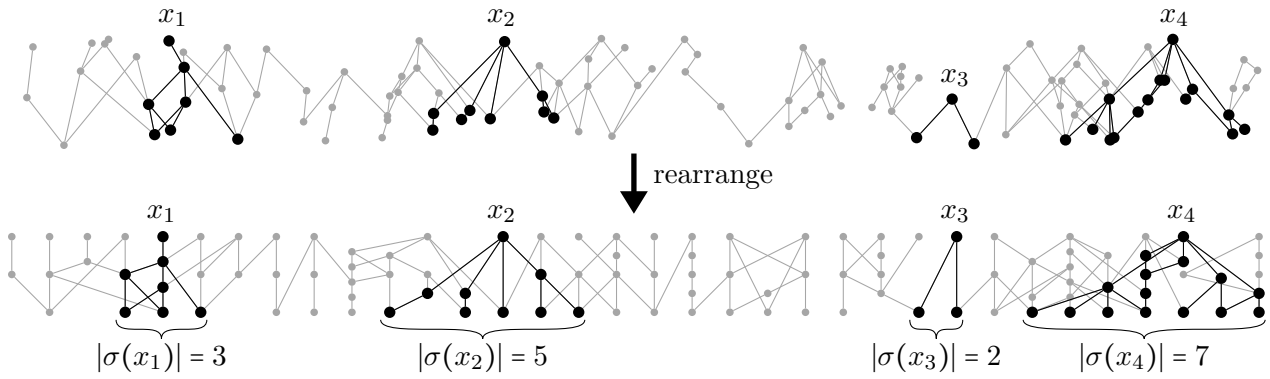

Figure 9: Initial cross sections in a random layer.

when the decrease in the number of relations is gradual and uniform. This abrupt change is called a *phase transition.* Figure 10 gives a rough illustration of this phenomenon. By itself, the figure is not very convincing, because the decrease in the number of relations in this particular universe is not at all uniform; the first few layers lose relations rapidly, while the last few layers possess roughly the same number of relations. In fact, the effect only becomes pronounced for somewhat larger discrete universes, which are more difficult to illustrate. However, the purpose of Figure 9 is merely to show what *does* happen in the generic case for future reference. The general behavioral description given above, including the presence of the phase transition, is justified quantitatively below.

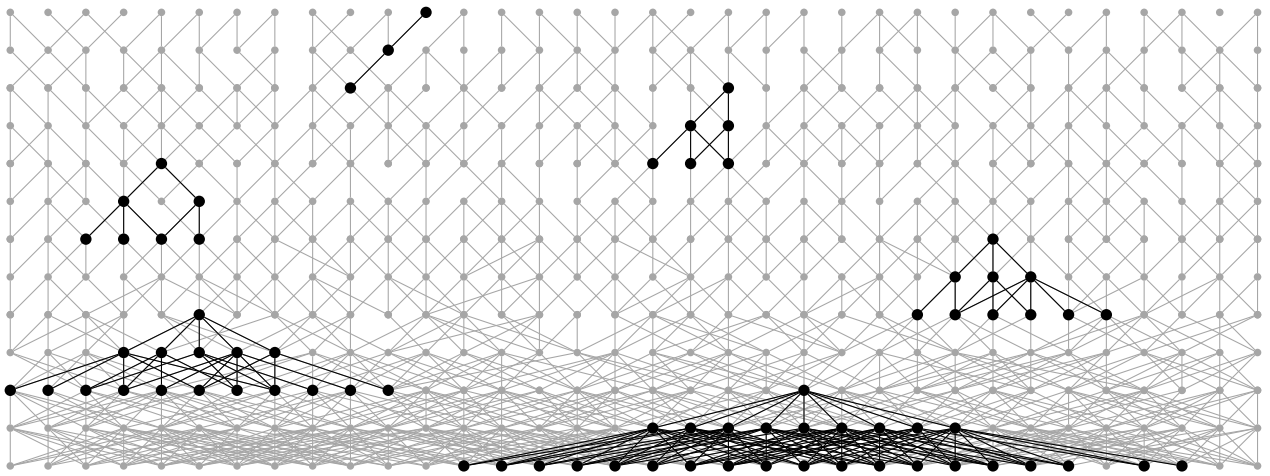

Figure 10: Decrease in sizes of causal horizons

It is useful to describe the decrease in sizes of causal horizons illustrated in Figure 10 in terms of a single number; namely, the probability that a randomly-chosen element in the first generation of a given layer belongs to the causal past of a randomly chosen element in the third generation. This probability may then be plotted as a function of the "layer number" as one moves from past to future. The left-hand diagram in Figure 11 illustrates how this probability function decreases for the universe illustrated in Figure 10. This universe may be partitioned into six layers of three generations each, where the third generation of layer n coincides with the first generation of layer $n + 1$. For future reference, it is useful to allow for the possibility of earlier and later layers, since more relations could be added to the bottom layer, and more relations could be removed from the top layer. The right-hand diagram in Figure 11 illustrates how this probability function decreases

for the universe illustrated in Figure 10 illustrates what happens in the generic case for a universe whose initial layers include all or most possible relations. In this situation, the probability that a randomly-chosen element in the first generation belongs to the causal past of a randomly chosen element in the third generation is very close to unity. Decreasing the number of relations generally does not change the sizes of causal horizons much until a certain threshold is reached, at which point a phase transition occurs. This is the part of the diagram illustrated by the gray region. Following the phase transition, the probability function levels out to a low value which may be zero or nonzero depending on whether the decrease in the number of relations continues, or stops at some point.

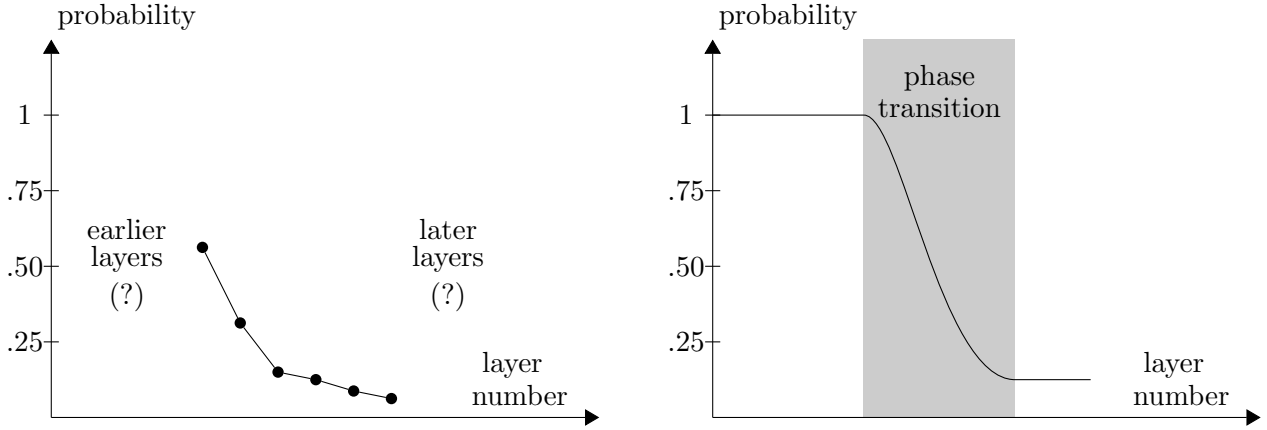

Figure 11: Probability function for the universe shown in Figure 10; generic case showing phase transition (gray).

For large universes, the phase transition illustrated in Figure 11 becomes very abrupt; the graph is nearly a step function. It is this sudden change in structure that produces effects similar to an inflationary epoch in the context of the horizon problem.

3 Some Toy Universes

Preliminary Remarks

The type of phase transition described qualitatively at the end of Section 3 may be demonstrated via a simple class of "toy universes" whose layers each consist of three generations of elements, like those illustrated in Figures 6 and 10. An individual such layer is illustrated in Figure 12. I denote such a layer by Δ_N^3, where the symbol Δ is intended to evoke the idea that each layer "increments" the corresponding universe. In this context, it is convenient to add or remove relations between pairs of elements of Δ_N^3 without renaming it; hence, Δ_N^3 is always assumed to possess three rows and n columns, but may have anywhere from 0 to $2N^2$ relations. The latter number, which is the maximal number of relations allowed in the present setting, is achieved by connecting each element of the first generation to each element of the second generation, and each element of the second generation to each element of the third generation. These choices are made for convenience, since the resulting universes suffice to illustrate the desired effect. The scope of validity of the resulting analysis can be appreciated only after further explanation.

A few specific features of Δ_N^3 are worth pointing out before proceeding. First, Δ_N^3 does not have

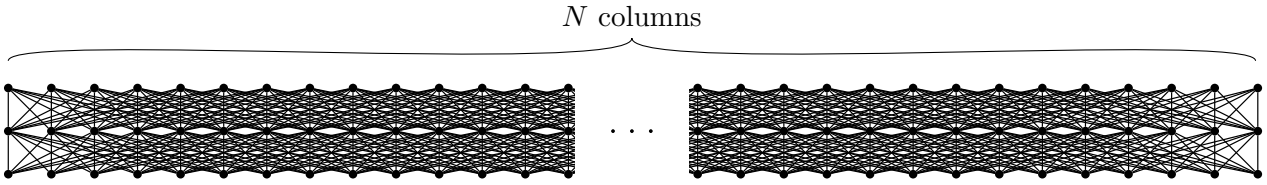

Figure 12: Layer Δ^3_N with three rows and N columns of elements.

any relations "skipping" from the first generation to the third generation. Such relations could be added, and relations of this type are physically relevant, but their presence would only enhance the demonstrated effects. More generally, Δ^3_N is deliberately chosen to produce, in a sense, the most *unfavorable* type of nontrivial universe for observing the desired phase transitions. The justification for this assertion is explained below. This choice allows certain important conclusions of the study to be applied in a broad setting. Second, when all the relations of Δ^3_N are removed, the resulting set of elements does not intrinsically encode any nontrivial causal structure. In this case, the choice to regard different elements of Δ^3_N as belonging to different generations is nothing but an arbitrary bookkeeping device. The same statement applies in an appropriate sense when Δ^3_N has only a few relations: certain elements will not admit any intrinsic identification as "top," "bottom," or "middle" elements. Consequently, the naïve causal interpretation of certain universes constructed from copies of Δ^3_N will "break down" for layers from which most of the relations have been removed. For practical purposes, this type behavior results from working with very small universes, and from carrying certain processes involving the removal of relations into terminal behavior outside the regime of physical interest. However, it is always useful to spell out the limits of validity in the interpretation of the mathematical objects involved in such an analysis.

Third, all of the universes constructed from copies of Δ^3_N differ from the closest discrete analogues relativistic spacetime in that they are not "Lorentz invariant;" i.e., they possess a preferred frame of reference, defined by taking each generation of elements to be a constant-time section. As it turns out, the behavior of interest in studying these universes does not depend on Lorentz invariance in any way, but this is not obvious at the present stage of discussion. In fact, the class of directed sets that do exhibit "discrete Lorentz invariance," such as causal sets defined by "sprinkling" elements randomly into a Lorentzian manifold, is a very restrictive family that is in some ways non-ideal for studying quantum gravity. The type of behavior demonstrated in this paper is much more general. For example, it does not depend on the emergent dimension of the objects involved. In this context, it is worth noting that the objects Δ^3_N are not intrinsically two dimensional, despite the manner in which they are represented graphically. In fact, when all of the edges of Δ^3_N are included, each generation resembles an $(N-1)$-dimensional simplex, in the sense that every element is related in the same way to every other element. In particular, the apparent "edge behavior" evident in Figure 12, in which the "far left" and "far right" elements appear to occupy special positions, is entirely illusory. It is possible for a recognizable notion of dimension emerge following the removal of certain families of relations of Δ^3_N, but most such removal processes will not exhibit any stable dimension, integer or otherwise. This reflects the more general fact that a very broad variety of universes may be built from copies of Δ^3_N. Fourth, one type of uniformity that most of the examples described here *do* exhibit is constant volume, where volume is understood to mean the number of elements in each generation. In other words, most of the toy universes studied here are built from copies of Δ^3_N for a *constant* choice of N. This choice of focus is made for purely for convenience, since constant-volume universes exhibit the desired behavior. In fact, this is another way in which these universes are chosen to be "worst case" examples for demonstrating phase transitions, since combinatorial considerations favor phase transitions in expanding universes. This statement is

verified quantitatively below.

Of particular interest in the present context are universes constructed by beginning with a copy of Δ_N^3 containing all its possible relations, then adding successive layers given by removing families of these relations. In other words, each layer of the resulting universe has fewer relations than the previous layer, and no relation removed from a given layer is added back in to a subsequent layer. The reason for focusing on universes of this type is that their structure may be analyzed by carrying out random removal processes on single copies of Δ_N^3. Any individual removal process is arbitrary, and is therefore not of particular interest; what is important is the *average structural properties* of universes obtained by carrying out random removal processes many times. The property of principal interest in the context of the horizon problem is the average sizes of initial cross sections of causal pasts, measured locally in each layer. More specifically, one has the following question:

Question: *After removing X relations from Δ_N^3, what is the probability $P_N(X)$ that there remains a path from a randomly-chosen element x in the first generation of Δ_N^3 to a randomly-chosen element z in the third generation of Δ_N^3?*

It turns out that this question may be answered precisely via a counting argument involving the inclusion-exclusion principle, as described below. The answer is:

$$P_N(X) = \frac{1}{\binom{2N^2}{X}} \sum_{n=1}^{N} (-1)^{n+1} \binom{N}{n} \binom{2N^2 - 2n}{X}.$$

It is sometimes convenient to work with the complementary probability

$$\begin{aligned} Q_N(X) &:= 1 - P_N(X) \\ &= 1 - \frac{1}{\binom{2N^2}{X}} \sum_{n=1}^{N} (-1)^{n+1} \binom{N}{n} \binom{2N^2 - 2n}{X}, \end{aligned}$$

which simplifies to

$$Q_N(X) = \frac{1}{\binom{2N^2}{X}} \sum_{n=0}^{N} (-1)^{n} \binom{N}{n} \binom{2N^2 - 2n}{X}. \tag{3.1}$$

$Q_N(X)$ may be understood qualitatively as the probability of a path between randomly-chosen x and z after *adding* $2N^2 - X$ relations to an empty copy of Δ_N^3, where "empty" means "having no relations" in this context. Mathematically, $Q_N(X)$ is the Nth entry of the *binomial transform* of the sequence

$$\left\{ \binom{2N^2 - 2n}{X} \right\}, \qquad n = 0, 1, 2, ..., N.$$

The theory of binomial transforms enables precise analysis of $P_N(X)$ and $Q_N(X)$ as N becomes large. In the context of cosmology, N might easily be 10^{100} or larger. Hence, the limiting behavior of $P_N(X)$ and $Q_N(X)$ as $N \to \infty$ is of particular interest. It turns out that there is an N-dependent critical value X_c of X, namely, $X_c = 2(N^2 - N^{\frac{3}{2}})$, at which a phase transition occurs in this limiting

behavior. Working with $P_N(X)$, this phase transition is described quantitatively by the result that

$$\lim_{N\to\infty} P_N\big(2(N^2 - N^\alpha)\big) = \begin{cases} 0 & \text{if } 0 < \alpha < \frac{3}{2} \\ \frac{e-1}{e} & \text{if } \alpha = \frac{3}{2} \\ 1 & \text{if } \frac{3}{2} < \alpha < 2 \end{cases}.$$

What this means in the context of removing relations from a copy of Δ^3_N for a large fixed value of N is that almost every element x in the first generation of Δ^3_N remains connected to almost every element z in the third generation until nearly X_c relations are removed. Beginning around this point, the probability that a randomly-chosen x and z are connected drops rapidly to zero as more relations are removed. The descriptor "rapidly" is relative to the maximal number $2N^2$ of relations in Δ^3_N. The left and right-hand diagrams of Figure 13 illustrate random removal processes for Δ^3_{10} and Δ^3_{20}, respectively, implemented via a Python script provided in the Appendix. The details of these diagrams are described below.

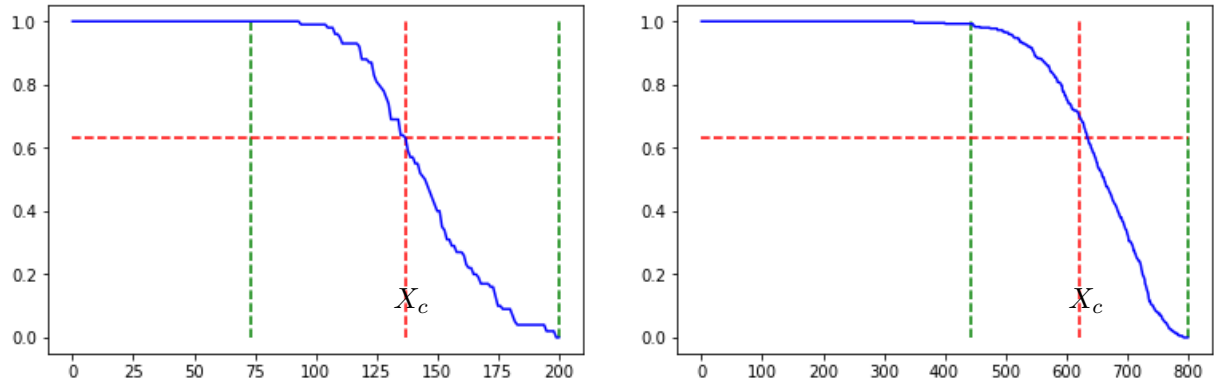

Figure 13: Phase transitions for Δ^3_{10} and Δ^3_{20}.

Given the relatively "tiny" sizes of 10 and 20 elements per generation compared to cosmologically relevant sizes of 10^{10} or 10^{50} or 10^{100}, it is not surprising that the drop-offs in these diagrams do not clearly evoke the type of drastic behavior one would associate with an inflationary epoch. However, the drop-off in the right-hand diagram is already significantly sharper, in a relative sense, than the drop-off in the left-hand diagram. At the beginning of the removal process, there are 10 different paths between a given x and z, and all these paths must be severed before x and z become disconnected. Hence, it is impossible for any pair to be disconnected before the first 10 relations are removed, and very unlikely that any pair will be disconnected before a significantly larger number of relations is removed. In this particular random trial, 79 relations are removed before the first pair (x, z) is disconnected. Overall, 99% of the drop-off occurs over a range of about 110 removals, or about 55% of the overall process. In the right-hand diagram, 99% of the drop-off occurs over a range of about 320 removals, or about 40% of the overall process. For sufficiently large values of N, a corresponding drop-off in connectivity will occur, on average, within an arbitrarily short interval in a relative sense. Figure 14 illustrates similar processes for slightly larger universes.

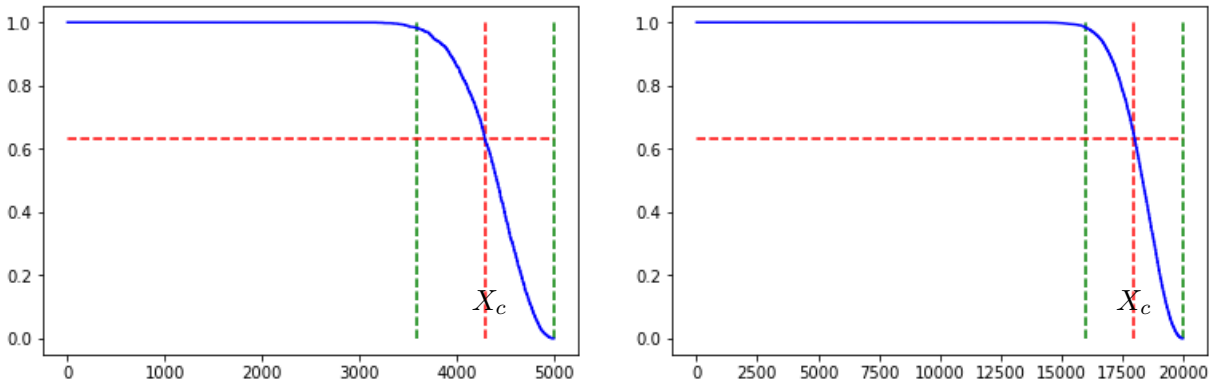

Figure 14: Phase transitions for Δ^3_{50} and Δ^3_{100}.

Case $N = 2$

The case $N = 2$ yields the simplest nontrivial removal processes. The quantity of principal interest governing these processes is the probability function $P_2(X)$, illustrated in Figure 16 below. The "full" version of Δ^3_2, including all possible relations, is illustrated on the left-hand side of Figure 15. The middle part of the figure illustrates a particular random removal process, where X denotes the number of relations removed at each stage. The nine directed sets appearing in this process are regarded as layers; i.e., copies of Δ^3_2 with varying numbers of relations which will be used to build a larger universe. The right-hand part of the figure shows a universe built by "stacking up" these layers, where the third generation of layer n is identified with the first generation of layer $n+1$. This particular universe becomes disconnected at the $X = 4$ layer, which means that its interpretation in terms of causal structure breaks down at this stage. However, this degenerate behavior is a consequence of the very small number of elements in each generation. For larger values of N, such wholesale disconnection will not occur until relatively late in the removal process, at a point already considered to be beyond the regime of physical interest in such idealized models.

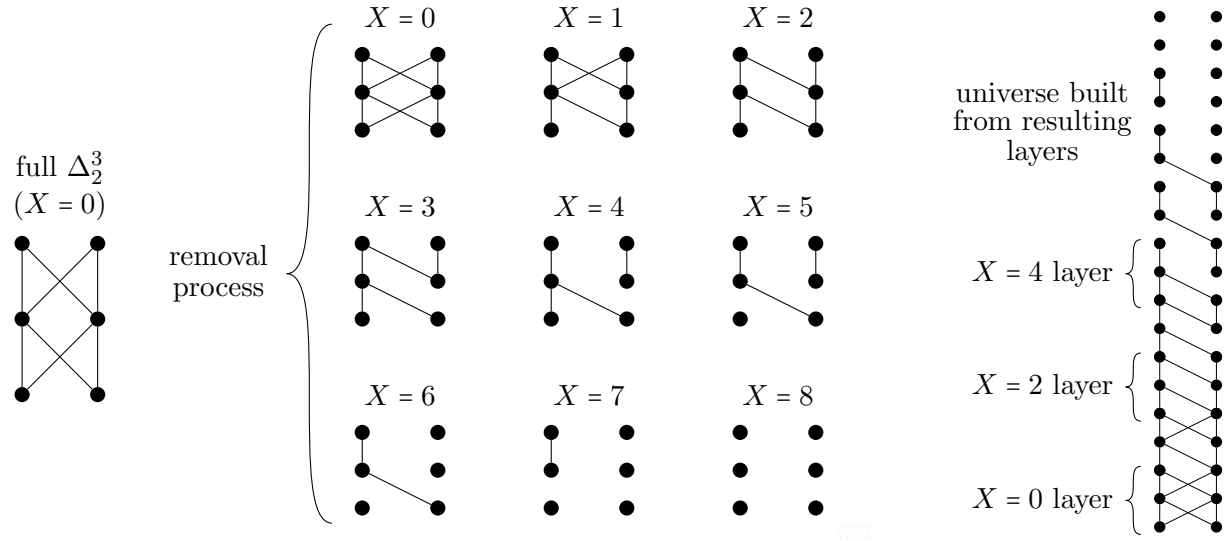

Figure 15: $N = 2$

There are $2N^2 = 8$ edges in the bottom layer; this is the "full" version of Δ_2^3, where $X = 0$. Inside a given layer, elements of the first generation, i.e., "bottom elements," are called *minimal elements,* while elements of the third generation, i.e., "top elements," are called *maximal elements.* A path from a minimal element to a maximal element, if it exists, consists of a chain of two consecutive relations joined together, and is therefore called called a 2-*chain.* The probability that both relations in a specifically-chosen 2-chain c are among a random choice of K edges is

$$P(\text{choosing } c) = \frac{\begin{pmatrix}\text{ways to choose the other } K-2\\ \text{relations besides those in } c\end{pmatrix}}{\begin{pmatrix}\text{total ways to choose } K\\ \text{out of } 2N^2 \text{ edges}\end{pmatrix}}$$

$$= \frac{\binom{2N^2-2}{K-2}}{\binom{2N^2}{K}} = \frac{\binom{6}{K-2}}{\binom{8}{K}} = \frac{6!}{(K-2)!(8-K)!}\frac{(K!)(8-K)!}{8!} = \frac{K(K-1)}{56}.$$

Similarly, the probability that all four relations belonging to a given pair of 2-chains c_1, c_2 is among a random choice of K edges is

$$\frac{\binom{2N^2-4}{K-4}}{\binom{2N^2}{K}} = \frac{\binom{4}{K-4}}{\binom{8}{K}} = \frac{4!}{(K-4)!(8-K)!}\frac{(K!)(8-K)!}{8!} = \frac{K(K-1)(K-2)(K-3)}{1680}.$$

Any minimal element x in the bottom generation is connected to any maximal element z by two 2-chains. After removing $8-K$ edges, K edges remain. The probability that x remains connected to z may be computed via the inclusion-exclusion principle, described in words as follows:

$$P_2(8-K) = \frac{\begin{pmatrix}\text{ways to include } c_1\\ \text{among } K \text{ chosen edges}\end{pmatrix} + \begin{pmatrix}\text{ways to include } c_2\\ \text{among } K \text{ chosen edges}\end{pmatrix} - \begin{pmatrix}\text{ways to include}\\ \text{both } c_1, c_2\end{pmatrix}}{\begin{pmatrix}\text{total ways to choose } K\\ \text{out of } 2N^2 \text{ edges}\end{pmatrix}}$$

$$= \frac{2\begin{pmatrix}\text{ways to include a specific one of}\\ c_1 \text{ or } c_2 \text{ among } K \text{ chosen edges}\end{pmatrix} - \begin{pmatrix}\text{ways to include}\\ \text{both } c_1, c_2\end{pmatrix}}{\begin{pmatrix}\text{total ways to choose } K\\ \text{out of } 2N^2 \text{ edges}\end{pmatrix}}$$

$$= \frac{2\binom{6}{K-2} - \binom{4}{K-4}}{\binom{8}{K}},$$

which works out to

$$P_2(8-K) = \begin{cases}\frac{1}{28}\left(K^2 - K\right) & \text{if } K \le 3\\ \frac{1}{1680}\left(-K^4 + 6K^3 + 49K^2 - 54K\right) & \text{if } K > 3.\end{cases}$$

Note that when $K = 3$, both polynomials share the common value 3/14. The derivatives are *not* equal at $K = 3$; they are 5/28 and 7/40, respectively. So the function P is a polynomial spline of order 0. While it is useful conceptually to work out these quantities in terms of the number K of remaining edges, the removal process underlying the computation suggests that the more natural variable is the number X of *removed* edges, which is equal to $2N^2 - K$ in the general case, and equal to $8 - K$ in the particular example $N = 2$. Fortunately, the choice to work with X turns out to be more convenient mathematically as well. This is the reason for writing $P_2(8-K)$ in the description of the inclusion-exclusion principle above. Rewriting P_2 in terms of X gives

$$P_2(X) = P_2(8-K) = \frac{2\binom{6}{K-2} - \binom{4}{K-4}}{\binom{8}{K}} = \frac{2\binom{6}{X} - \binom{4}{X}}{\binom{8}{X}},$$

which works out to

$$P_2(X) = \begin{cases} 2\left(\frac{8-X}{8}\right)\left(\frac{7-X}{7}\right) - \left(\frac{8-X}{8}\right)\left(\frac{7-X}{7}\right)\left(\frac{6-X}{6}\right)\left(\frac{5-X}{5}\right) & \text{if } X < 5 \\ 2\left(\frac{8-X}{8}\right)\left(\frac{7-X}{7}\right) & \text{if } X \geq 5. \end{cases} \tag{3.2}$$

This may be expanded to yield

$$P_2(X) = \begin{cases} \frac{1}{1680}\left(-X^4 + 26X^3 - 191X^2 + 166X + 1680\right) & \text{if } X < 5 \\ \frac{1}{28}\left(X^2 - 15X + 56\right) & \text{if } X \geq 5 \end{cases}$$

However, there does not seem to be any advantage in performing this expansion; the pattern is much clearer and the computations much easier using Equation (4.2). The values of $P_2(X)$ for $X = 0, 1, 2, ..., 8$, using a constant denominator of 70, are

$$\frac{70}{70}, \frac{70}{70}, \frac{60}{70}, \frac{45}{70}, \frac{29}{70}, \frac{15}{70}, \frac{5}{70}, \frac{0}{70}, \frac{0}{70}$$

Figure 16 illustrates the behavior of $P_2(X)$. The black nodes indicate the values of $P_2(X)$ for the integers $X = 0, 1, 2, ..., 8$, which are the only values that are meaningful in this context. In particular, note that the values indicated by the dark red portion of the curve between 0 and 1 are greater than 1, which makes no sense for a probability, while the values of the values indicated by the dark blue portion of the curve between 7 and 8 are negative, which again makes no sense. However, the extension of $P_2(X)$ to a continuous polynomial spline interpolating between pairs of integer values is nevertheless useful for the purposes of analysis. The graph of this polynomial spline is the union of the dark red and dark blue polynomial curves in the figure. The dark red curve is the graph of the polynomial $P_2^1(X) = \frac{1}{1680}\left(-X^4 + 26X^3 - 191X^2 + 166X + 1680\right)$ for $0 \leq X < 5$, while the dark blue curve is the graph of the polynomial $P_2^2(X) = \frac{1}{28}\left(X^2 - 15X + 56\right)$ for $5 \leq X \leq 8$. The superscripts in $P_2^1(X)$ and $P_2^2(X)$ denote the first and second pieces of the spline. These two polynomials share the common value of 3/14 at $X = 5$. The light red curve shows the extension of $P_1^2(X)$ outside its interval of validity, while the light blue curve shows the corresponding extension of $P_2^2(X)$.

The probability function $P_2(X)$ gives the *expected* fraction of third-generation elements reachable from a given first-generation element at each stage in a random removal process. However, different particular instances of such processes will produce different functions. In particular, each value of a particular such function will be of the form $m/4$ for some integer $0 \leq m \leq 4$, since $2^2 = 4$ is the number of ways to pair a first-generation element with a third-generation element. Averaging the functions for many such removal processes produces a function that approaches $P_2(X)$ in the limit.

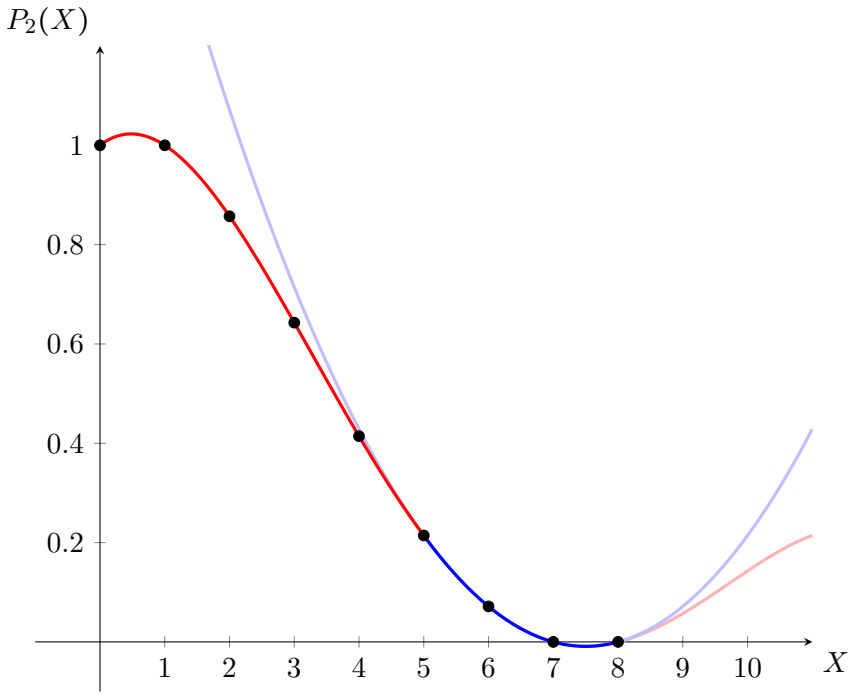

Figure 16: $P_2(X)$.

Case $N = 3$

Using $N = 3$ as an example, there are $2N^2 = 18$ edges in D. The probability that a given 2-chain c_1 is among a choice of K edges is

$$\frac{\binom{2N^2-2}{K-2}}{\binom{2N^2}{K}} = \frac{\binom{16}{K-2}}{\binom{18}{K}} = \frac{16!}{(K-2)!(18-K)!}\frac{(K!)(18-K)!}{18!} = \frac{K(K-1)}{306}$$

The probability that a given pair of 2-chains c_1, c_2 is among a choice of K edges is

$$\frac{\binom{2N^2-4}{K-4}}{\binom{2N^2}{K}} = \frac{\binom{14}{K-4}}{\binom{18}{K}} = \frac{14!}{(K-4)!(18-K)!}\frac{(K!)(18-K)!}{18!} = \frac{K(K-1)(K-2)(K-3)}{73440}$$

The probability that a given triple of 2-chains c_1, c_2, c_3 is among a choice of K edges is

$$\frac{\binom{2N^2-6}{K-6}}{\binom{2N^2}{K}} = \frac{\binom{12}{K-6}}{\binom{18}{K}} = \frac{12!}{(K-6)!(18-K)!}\frac{(K!)(18-K)!}{18!} = \frac{K(K-1)(K-2)(K-3)(K-4)(K-5)}{13366080}$$

Even for such a small choice of N, the denominators on the right-hand sides of the previous three computations are unwieldy and unilluminating, and it does not seem to be useful to expand the expressions involving K in the numerators. It will be easier below to work with the variable $X = 2N^2 - K = 18 - K$. Any minimal element x in D is connected to any maximal element z by three 2-chains. After removing $18 - K$ edges, K edges remain. The probability that x remains connected to z may be computed via the inclusion-exclusion principle:

$$P(18-K) = \frac{3\left(\begin{array}{c}\text{ways to include one of } c_1,\ c_2\\ \text{or } c_3 \text{ among } K \text{ chosen edges}\end{array}\right) - 3\left(\begin{array}{c}\text{ways to include}\\ \text{two of } c_1, c_2, c_3\end{array}\right) + 1\left(\begin{array}{c}\text{ways to include all}\\ \text{three of } c_1, c_2, c_3\end{array}\right)}{\left(\begin{array}{c}\text{total ways to choose } K\\ \text{out of } 2N^2 \text{ edges}\end{array}\right)}$$

$$= \frac{3\binom{16}{K-2} - 3\binom{14}{K-4} + 1\binom{12}{K-6}}{\binom{18}{K}},$$

Changing variables to $X = 18 - K$ yields

$$P(X) = \frac{3\binom{16}{X} - 3\binom{14}{X} + 1\binom{12}{X}}{\binom{18}{X}},$$

which works out to

$$P(X) = \begin{cases} 3\left(\frac{18-X}{18}\right)\left(\frac{17-X}{17}\right) - 3\left(\frac{18-X}{18}\right)\left(\frac{17-X}{17}\right)\left(\frac{16-X}{16}\right)\left(\frac{15-X}{15}\right) + \left(\frac{18-X}{18}\right)\left(\frac{17-X}{17}\right)\left(\frac{16-X}{16}\right)\left(\frac{15-X}{15}\right)\left(\frac{14-X}{14}\right)\left(\frac{13-X}{13}\right) & \text{if } X < 13 \\ 3\left(\frac{18-X}{18}\right)\left(\frac{17-X}{17}\right) - 3\left(\frac{18-X}{18}\right)\left(\frac{17-X}{17}\right)\left(\frac{16-X}{16}\right)\left(\frac{15-X}{15}\right) & \text{if } 13 \le X < 15 \\ 3\left(\frac{18-X}{18}\right)\left(\frac{17-X}{17}\right) & \text{if } X \ge 15 \end{cases} \tag{3.3}$$

Figure 16 illustrates the behavior of $P(X)$. The black nodes indicate the values of $P(X)$ for the integers $X = 0, 1, 2, ..., 18$. The polynomial spline for $P(X)$ is indicated by the union of the dark red, dark green, and dark blue polynomial curves in the figure.

4 3-Generation Universes

Probability Function

In the general 3-generation case with N elements per generation, any minimal element x in the bottom generation of a full copy of Δ^3_N is connected to any maximal element z by N different 2-chains $c_1, ..., c_N$. After removing $2N^2 - K$ edges, K edges remain. The probability that x remains connected to z may be computed via the inclusion-exclusion principle, described in words as follows:

$$P_N(2N^2 - K) =$$

$$\frac{N\left(\begin{array}{c}\text{ways to include one of } c_1, ..., c_N\\ \text{among } K \text{ chosen edges}\end{array}\right) - \binom{N}{2}\left(\begin{array}{c}\text{ways to include}\\ \text{two of } c_1, ..., c_N\end{array}\right) + \binom{N}{3}\left(\begin{array}{c}\text{ways to include}\\ \text{three of } c_1, ..., c_N\end{array}\right) - ...}{\left(\begin{array}{c}\text{total ways to choose } K\\ \text{out of } 2N^2 \text{ edges}\end{array}\right)}$$

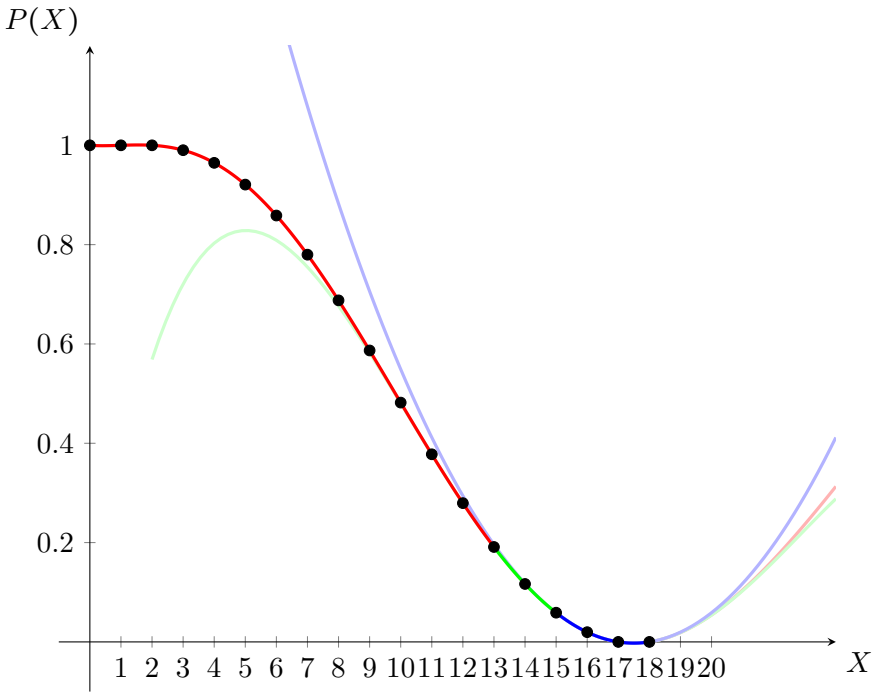

Figure 17: Three by three.

$$= \frac{1}{\binom{2N^2}{K}} \sum_{n=1}^{N} (-1)^{n+1} \binom{N}{n} \binom{2N^2 - 2n}{K - 2n}.$$

Let $X = 2N^2 - K$. By symmetry of binomial coefficients,

$$\binom{2N^2}{K} = \binom{2N^2}{2N^2 - K} = \binom{2N^2}{X}$$

and

$$\binom{2N^2 - 2n}{K - 2n} = \binom{2N^2 - 2n}{(2N^2 - 2n) - (K - 2n)} = \binom{2N^2 - 2n}{X},$$

So the probability becomes

$$P_N(X) = \frac{1}{\binom{2N^2}{X}} \sum_{n=1}^{N} (-1)^{n+1} \binom{N}{n} \binom{2N^2 - 2n}{X}.$$

The complementary probability $Q_N(X) = 1 - P_N(X)$ is

$$Q_N(X) = 1 - \frac{1}{\binom{2N^2}{X}} \sum_{n=1}^{N} (-1)^{n+1} \binom{N}{n} \binom{2N^2 - 2n}{X}$$

$$= 1 + \frac{1}{\binom{2N^2}{X}} \sum_{n=1}^{N} (-1)^{n} \binom{N}{n} \binom{2N^2 - 2n}{X}$$

(changing $(-1)^{n+1}$ to $(-1)^n$) and finally

$$Q_N(X) = \frac{1}{\binom{2N^2}{X}} \sum_{n=0}^{N} (-1)^{n} \binom{N}{n} \binom{2N^2 - 2n}{X},$$

since the $n = 0$ term is

$$\frac{1}{\binom{2N^2}{X}}(-1)^0\binom{N}{0}\binom{2N^2}{X} = 1.$$

5 4-Generation Universes

In the general 4-generation case with N elements per generation, there are a total of $3N^2$ relations. Any minimal element w in the bottom generation of a full copy of Δ_N^4 is connected to any maximal element z by N^2 different 3-chains $c_{11}, ..., c_{NN}$. This is because there are N choices for the second-generation element x_k and N choices for the third-generation element y_k. After removing $3N^2 - K$ relations, K relations remain. The probability that w remains connected to z is complicated by the fact that different 3-chains from w to z are not necessarily disjoint. For example, there are N different chains $c_{11}, c_{12}, ..., c_{1N}$ all sharing the relation $w \prec x_1$. However, the probability of selecting a particular 3-chain c among a choice of K edges is straightforward:

$$P(\text{choosing } c) = \frac{\begin{pmatrix}\text{ways to choose the other } K-3\\ \text{relations besides those in } c\end{pmatrix}}{\begin{pmatrix}\text{total ways to choose } K\\ \text{out of } 3N^2 \text{ edges}\end{pmatrix}}$$

$$= \frac{\binom{3N^2-3}{K-3}}{\binom{3N^2}{K}}.$$

A pair of chains c_{ij} and c_{IJ} either

1. Coincide if $(i, j) = (I, J)$. (three total relations)
2. Share a single relation if $i = I$ XOR $j = J$. (five total relations)
3. Are disjoint if $i \neq I$ and $j \neq J$. (six total relations)

In case 2, the probability of selecting all five relations belonging to two 3-chains c, c' sharing a single relation among a choice of K edges is

$$P(\text{choosing } c \text{ and } c') = \frac{\begin{pmatrix}\text{ways to choose the other } K-5\\ \text{relations besides those in } c, c'\end{pmatrix}}{\begin{pmatrix}\text{total ways to choose } K\\ \text{out of } 3N^2 \text{ edges}\end{pmatrix}}$$

$$= \frac{\binom{3N^2-5}{K-5}}{\binom{3N^2}{K}}.$$

There are $2N$ different relations that could be shared by c, c' (N from w to x_k and N from y_l to z), and for each such relation, there are $\binom{N}{2}$ choices of c, c'. Altogether, then, there are $2N\binom{N}{2}$ such pairs c, c'.

Similarly, in case 3,

$$P(\text{choosing } c \text{ and } c') = \frac{\binom{3N^2-6}{K-6}}{\binom{3N^2}{K}}.$$

There are $\binom{N}{2}$ ways to choose second-generation elements for c, c', and $\binom{N}{2}$ ways to choose third generation elements. Hence, there are $2\binom{N}{2}^2$ such pairs c, c', these choices may be ordered in two different ways.

A triple of chains $c_{i_1 j_1}, c_{i_2 j_2}, c_{i_3 j_3}$ has a more complex list of possible relationships:

1. All three coincide (three total relations)
2. Two coincide and share a single relation with the third (five total relations)
3. Two coincide and share no relations with the third (six total relations)
4. Three share a single common relation (seven total relations)
5. One shares a relation with another, and a different relation with a third (seven total relations)
6. Two share a single relation (eight total relations)
7. All are disjoint (nine total relations)