(Quantum) spacetime as a statistical geometry of lumps in random networks

Manfred Requardt
Institut für Theoretische Physik, Universität Göttingen, Bunsenstrasse 9, 37073 Göttingen, Germany
E-mail: requardt@theorie.physik.uni-goettingen.de

Received 11 May 1999, in final form 30 March 2000

Abstract. In the following we undertake to describe how macroscopic spacetime (or rather, a microscopic protoform of it) is supposed to emerge as a superstructure of a web of lumps in a stochastic discrete network structure. As in preceding work (mentioned below), our analysis is based on the working philosophy that both physics and the corresponding mathematics have to be genuinely discrete on the primordial (Planck scale) level. This strategy is concretely implemented in the form of cellular networks and random graphs. One of our main themes is the development of the concept of physical (proto)points or lumps as densely entangled subcomplexes of the network and their respective web, establishing something like (proto)causality. It may perhaps be said that certain parts of our programme are realizations of some early ideas of Menger and more recent ones sketched by Smolin a couple of years ago. We briefly indicate how this two-storey concept of quantum spacetime can be used to encode the (at least in our view) existing non-local aspects of quantum theory without violating macroscopic spacetime causality!

PACS number: 0460

1. Introduction

In recent papers [1–4] we developed some facets of an extensive programme we formulated therein, i.e. to reconstruct ordinary continuum physics or mathematics as a kind of coarse-grained limit from a much more primordial and genuinely discrete substrate, including, in particular, a discrete theory of (proto)spacetime as the universal receptacle or substratum of all physical processes.

A corresponding philosophy is presently held by a substantial minority of workers in quantum gravity and we commented on some of the various approaches, at least as far as we are aware of them, in the foregoing papers. We therefore refer the interested reader to these papers for references we do not mention in the following just for the sake of brevity. As an exception we mention only the early and prophetic remarks made by Penrose in, for example, [5] about the surmised combinatorial substratum underlying our continuous spacetime, the ideas of Smolin, sketched at the end of [6], because they are surprisingly close to our working philosophy and the work of ’t Hooft [7] which is based on the model system of cellular automata (the more rigid and regular relatives of our dynamic cellular networks introduced in the following).

Our personal working philosophy is that spacetime at the very bottom (i.e. near or below the notorious Planck scale) resembles or can be modelled as an evolving information processing
cellular network, consisting of elementary modules (with, typically, simple internal discrete state spaces) interacting with each other via dynamical bonds which transfer the elementary pieces of information among the nodes. That is, the approach shares the combinatorial point of view in fundamental spacetime physics which has been initiated by, for example, Penrose. It is a crucial and perhaps characteristic extra ingredient of our framework that the bonds (i.e. the elementary interactions) are not simply dynamical degrees of freedom (as with the nodes their internal-state space is assumed to be simple) but can a fortiori, depending on the state of the local network environment, be switched on or off, i.e. can temporarily be active or inactive! This special ingredient of the dynamics hopefully allows the network to perform geometric phase transitions into a new ordered phase having a certain two-storey structure to be explained below. This conjectured geometric order we view as a kind of discrete protospace-time carrying metric, causal and dimensional structures. The main part of this paper consists of a description and analysis of these structural elements and their interplay.

In [1] we dealt primarily with dimensional concepts on such discrete and irregular spaces. It furthermore became apparent that there exist close ties to the theory of fractal sets. Papers [3, 4], on the other hand, are, among other things, devoted to the development of several possible versions of discrete analysis and discrete differential geometry, respectively, discrete functional analysis with certain zones of contact with non-commutative geometry.

As has been beautifully reviewed by Isham in various papers (see, e.g., [8]) one could, among several possible attitudes, adopt the perhaps most radical working philosophy in quantum gravity and speculate that both quantum theory and gravity are merely secondary and derived aspects or, expressed in more physical terms, so-called effective theories of an underlying more primordial theory of a markedly combinatorial flavour. However, a theory comprising quantum theory and gravitation as emergent subtheories should, first of all, provide a framework in which both the emergence of something we may consider as a proto form of classical spacetime, respectively, quantum vacuum can be expressed or discussed, most notably the emergence of the continuum from the discrete and the concept of physical spacetime points (having an internal dynamical structure) and their intricate web. In a forthcoming paper [9] we will (try to) show how quantum theory emerges as a coarse-grained stochastic theory within such a framework.

Some months ago, after the preparation of a first draft, we were kindly informed by El Naschie that a very early source where a couple of related ideas can be found is the contribution of Menger in [10], perhaps better known from his research on topological dimension or fractal sets such as the Menger sponge. In this essay he entertains very interesting ideas about the necessity of a new geometry of the microcosmos based on the geometry of lumps and the concept of a statistical metric space. We note in passing that some of the concepts we develop in the following seem to be very much in accord with this point of view. Quite remarkable in this respect is also Einstein’s open-minded attitude towards the possibility of a more primordial discrete spacetime theory expressed at the end of the same volume. As to this particularly important point see also his many utterances compiled by Stachel in [11], p 27ff.

We think that our discrete framework may exactly realize some of the ideas he had in mind. (In this context we also want to mention the Cantorian spacetime approach of El Naschie et al who try to model microspace as a particular type of (random) fractal (see, e.g., [12]). Another interesting early source with a possible bearing on our approach may also be Neumann’s concept of continuum geometry (geometry without points) which is briefly described in [13] (the original source can be found in [14]). Furthermore, there exist surprising links to other seemingly unrelated areas of current research as we learnt very recently, a catchword being small-world networks (see [15]). Some of the mechanisms that are effective in our cellular network environment are also observed there.
After having nearly completed this final version we received a message from Roy who informed us that the framework sketched by Menger has been developed further by Menger himself together with several co-workers and later by other groups (see, e.g., [16]). The state of affairs is described in his monograph [17] and the more mathematical aspects are developed in [18]. From these monographs one may see that the early ideas of Menger have meanwhile unfolded into various branches in their own right and it is not easy at the moment to relate all of them with our own working philosophy. Anyhow, the situation looks very promising and interesting: we will deal with the geometric web of lumps (we also called them cliques) in section 4 (which is the central section of this paper) and make some remarks concerning the metric aspects in subsection 4.3. Another interesting paper we came across only very recently is [19], which also employs the random graph idea to calculate certain graph characteristics and relates them to properties of space. The results belong, however, more to the context of our paper [1], which deals with dimensional concepts on networks and the like.

Some clarifying comments are perhaps in order at this point. The modelling of the depth structure of spacetime as a cellular network consisting of nodes and bonds should not necessarily be understood in a plain bodily sense. One should rather consider it as a representation or emulation of the main characteristics of the physical scenario. In this connection we want to mention the following interesting and purely discrete approaches in [20, 21]. There may, in particular, exist a variety of superficially different systems, the logical structure of which can nevertheless be encoded in roughly the same abstract underlying network model. It is our belief that such a discrete network, governed by a relatively simple but cleverly chosen dynamical law, is capable of generating most if not all of the phenomena and emergent laws on which our ordinary continuum physics is usually grounded. That such a hypothesis is not entirely far-fetched may, for example, be inferred from the emerging complexity of such a simple cellular automaton model as the famous game of life created by Conway. A typical example is the geometry of lumps envisaged by Menger. Take as lumps the hypothetical infinitesimal grains of space or spacetime which cannot be further resolved (be it in a practical or principle sense). Let them overlap according to a certain rule so that they can interact or exchange information. Draw a node for each such lump and a bond for each two lumps which happen to overlap. In combinatorial topology such a combinatorial complex is called the nerve of the set system (cf, e.g., [22]). In a next step one may encode the respective strengths of interaction or degrees of overlap in a valuation of the corresponding bonds, yielding a cellular network of the kind we have in mind. A fortiori one can make these mutual overlaps into dynamical variables, i.e. let them change during the course of evolution.

It may seem, at first glance, that our bottom-up approach, which starts almost from first principles, has the disadvantage, if compared with other more top-down-oriented approaches such as, for example, string theory or loop quantum gravity, of being rather arbitrary. These other frameworks are based (at least to a large part) on more or less continuum concepts and ideas theoretical physics has become accustomed to and some discrete behaviour (e.g. spin networks) can be detected only at the end of the road, if at all. Furthermore, in the case of, for example, string theory, the quantum principles (whatever that means exactly) are taken for granted down to arbitrarily fine scales (at least as far as we can see) and both rely on such approved concepts as functional integrals or sum over configurations. To such a potential criticism we would like to reply in the following way. For one, it is far from obvious that the (typically Euclidean) functional integral philosophy does still hold sway in, for example, the Planck regime, as it is (implicitly) based on some kind of action principle which, on one hand, may have its roots in classical physics. On the other hand, it may turn
out that such concepts are only the surface aspect of a deeper, more hidden reality. Such a possibility can of course not be precluded, but we do not want to base our approach on such a heuristic principle from the outset. Quite the contrary, we want to deduce quantum theory (and by the same token the functional integral framework) as an effective theory from our presumably more fundamental theory. How this may work we undertake to show in the above-mentioned paper [9]. We expect, however, that these approaches may merge on some intermediate level.

We would like to add yet another, as we think, important remark. Quite a few researchers in this field expect physics at the Planck scale to be quite different from the kind of physics we are accustomed to (take, for example, the interesting remarks in [23]). The situation may be comparable to the physics near or at the critical point in, say, renormalization theory where different microscopic theories (the so-called universality class) lead to basically the same macroscopic behaviour as the microscopic details happen to be washed out in the coarse-graining process. This of course does not mean that the search for such an underlying microscopic theory is futile or useless. The lesson is rather that one should be prepared to employ perhaps other guiding principles in the quest for these presumed more primordial laws and concentrate, at least in a first step, rather on a whole class of possible primordial model theories instead of a single theory of everything. As we will show in the mentioned companion paper, quantum behaviour may just emerge as such a coarse-grained stochastic behaviour of a whole class of discrete microscopic model theories. In other words, one of our more heuristic guiding principles is the following. We consider a class of microscopic (discrete) theories to be worth studying if they have the propensity to generate, for example, quantum behaviour or a kind of protospacetime in some classical limit. It would be even nicer if, at an intermediate level, they develop links to, say, string theory or loop quantum gravity. That there are relations to non-commutative geometry was already shown in the above-mentioned papers.

2. The cellular network environment

In this section we briefly describe the general framework within which our investigation will take place. Among other things we introduce a simple type of local dynamical network law (for more details see [3, 24]). A certain class of relatively simple cellular networks is the following.

We emulate the underlying substratum of our world, or, more specifically, of our spacetime (quantum) vacuum (containing, however, in addition all the existing quantum and macro objects as extended excitation patterns!) by what we call a cellular network. This discrete structure consists of elementary nodes, $n_i$, which interact (or exchange information) with each other via bonds, $b_{ik}$, playing the role of (in this context) not further reducible (abstract) elements. The possible internal structure of the nodes (modules) or bonds (interaction channels) is emulated by discrete internal-state spaces carried by the nodes/bonds. The node set is assumed to be large but finite or countable. The bond $b_{ik}$ is assumed to connect the nodes $n_i, n_k$. The internal states of the nodes/bonds are denoted by $s_i, J_{ik}$, respectively. As our philosophy is to generate complex behaviour out of simple models we typically make simple choices for them, one of which is, for example,

$$s_i \in q\mathbb{Z}, \quad J_{ik} \in \{-1, 0, +1\}$$

with $q$ an elementary quantum of information.

As in our approach the bond states are dynamical degrees of freedom which, a fortiori, can be switched off or on, the wiring, that is the pure geometry, of the network is also an
emergent, dynamical property and is not given in advance. Consequently, the nodes, bonds are typically not arranged in a more or less regular array, a lattice say, with a fixed near-/far-order. This implies that geometry will become to some extent a relational (Machian) concept and is no longer an ideal element (cf the discussion in the above-mentioned forthcoming paper [9]).

On the other hand, as in cellular automata, the node and bond states are updated (for convenience) in discrete clock time steps, \( t = z \tau \) and \( z \in \mathbb{Z} \) where \( \tau \) is an elementary clock time interval. This updating is given by some local dynamical law (examples are given below). In this context local means that the node/bond states are changed at each clock time step according to a prescription with the input as the overall state of a certain neighbourhood (in some topology) of the node/bond under discussion. We want, however, to emphasize that \( t \) is not to be confused with some physical time, which, for its part, is also considered to be an emergent coarse-grained quantity. The well known problem of time is, for the time being, not treated in detail in the following, as it is a big problem of its own, needing a careful and separate analysis of its own (see, however, [25, 26]). That is, at the moment the above clock time is neither considered to be dynamical nor observer dependent. There is, however, a discussion concerning the presumed emergence of a new primordial time scale which sets the scale for the regime where quantum fluctuations hold sway in [9].

A simple example of such a local dynamical law as we have in mind is given in the following definition.

**Definition 2.1 (Example of a local law).** At each clock time step a certain quantum \( q \) is exchanged between, say, the nodes \( n_i, n_k \), connected by the bond \( b_{ik} \) such that

\[
s_i(t + \tau) - s_i(t) = q \sum_k J_{ki}(t)
\]

(i.e. if \( J_{ki} = \pm 1 \) a quantum \( q \) flows from \( n_k \) to \( n_i \), etc).

The second part of the law describes the back reaction on the bonds (and is, typically, more subtle). This is the place where the so-called ‘hysteresis interval’ enters the stage. We assume the existence of two ‘critical parameters’ \( 0 \leq \lambda_1 \leq \lambda_2 \) with

\[
J_{ik}(t + \tau) = 0 \quad \text{if} \quad |s_i(t) - s_k(t)| =: |s_{ik}(t)| > \lambda_2
\]

\[
J_{ik}(t + \tau) = \pm 1 \quad \text{if} \quad 0 < \pm s_{ik}(t) < \lambda_1
\]

with the special proviso that

\[
J_{ik}(t + \tau) = J_{ik}(t) \quad \text{if} \quad s_{ik}(t) = 0.
\]

On the other hand,

\[
J_{ik}(t + \tau) = \begin{cases} 
\pm 1 & J_{ik}(t) \neq 0 \\
0 & J_{ik}(t) = 0
\end{cases} \quad \text{if} \quad \lambda_1 \leq \pm s_{ik}(t) \leq \lambda_2.
\]

In other words, bonds are switched off if local spatial charge fluctuations are too large, switched on again if they are too small, their orientation following the sign of local charge differences, or remain inactive.

Another interesting law arises if one exchanges the role of \( \lambda_1 \) and \( \lambda_2 \) in the above law, that is, bonds are switched off if the local node fluctuations are too small and are switched on again if they exceed \( \lambda_2 \). We emulated all of these laws on a computer and studied a lot of network properties. The latter law has the peculiar feature that it turned out to have very short transients in the simulations, i.e. it reaches an attractor in a very short clock time. Furthermore, these attractors or state-cycles turned out to be very regular, that is, they had a very short period of
typically six, i.e., the whole network returned to a previous state after only six clock time steps, which is quite remarkable, given the seeming complexity of the evolution and the huge phase space [27].

Remarks.

(a) It is important that, generically, such laws, as introduced above, do not lead to a reversible time evolution, i.e. there will typically exist attractors or state-cycles in total phase space (the overall configuration space of the node and bond states). On the other hand, there exist strategies (in the context of cellular automata!) to design particular reversible network laws (cf, e.g., [29]) which are, however, typically of second order. Usually the existence of attractors is considered to be important for pattern formation. On the other hand, it may suffice that the phase space, occupied by the system, shrinks during the course of evolution, that is, that one has a flow into smaller subvolumes of phase space.

(b) In the above class of laws a direct bond–bond interaction has not yet been implemented. We are prepared to incorporate such a (possibly important) contribution in a next step if it turns out to be necessary. In any case there are not so many ways to do this in a sensible way. Stated differently, the class of possible physically sensible interactions is perhaps not so numerous.

(c) As in the definition of evolution laws of spin networks by, for example, Markopoulou et al (see [30, 31]), there are in our case more or less two possibilities: treating evolution laws within an integrated spacetime formalism or regarding the network as representing space alone with the time evolution being implanted via some extra principle (which is the way we have chosen above). The inter-relation of these various approaches and frameworks, while being very interesting, is, however, far from obvious at the moment and needs a separate detailed investigation.

Observation 2.2 (Gauge invariance). The above dynamical law depends nowhere on the absolute values of the node charges but only on their relative differences. By the same token, charge is nowhere created or destroyed. We have

$$\Delta \left( \sum_{Q \in X} s(n) \right) = 0.$$  \hspace{1cm} (7)

To avoid artificial ambiguities we can, for example, choose a fixed reference level and take as an initial condition, respectively, the constraint

$$\sum_{Q \in X} s(n) = 0.$$  \hspace{1cm} (8)

There are many different aspects of our class of cellular networks one can study in this context. One can, for example, regard them as complex dynamical systems, or one can undertake to develop a statistical or stochastic framework, etc. In a purely geometric sense, however, they are evolving graphs. As, in this paper, we are primarily concerned with the analysis of the microstructure of (quantum) spacetime, it seems to be a sensible strategy to suppress, at least in a first step, all the other features such as, for example, the details of the internal-state spaces of nodes and bonds and concentrate instead on their pure wiring diagram and its reduced (graph) dynamics. This is already an interesting characteristic of the network (perhaps somewhat reminiscent of the Poincaré map in the theory of chaotic systems) as bonds can be switched on and off during the course of clock time so that already the wiring diagram
will constantly change. Furthermore, as we will see, it encodes the complete near- and far-order structure of the network, that is, it tells us which regions are experienced as nearby or far away (in a variety of possible physical ways such as the strength of correlations or with respect to some other physically meaningful metric such as, for example, statistical distance, etc). Evidently, this is one of the crucial features we expect from something like physical spacetime. In the above simple scenario with $J_{ik} = \pm 1$ or 0 one can, for example, draw a directed bond, $d_{ik}$, if $J_{ik} = +1$, with $J_{ik} = -J_{ki}$ implied, and delete the bond if $J_{ik} = 0$. This leads to a (clock) time-dependent graph, $G(t)$, or wiring diagram. In other words, we will deal in the following mainly with the evolution and structure of large dynamical graphs.

We close this section with a brief résumé of the characteristics an interesting network dynamics should encode (in our view).

Résumé 2.3. Irrespective of the technical details of the dynamical evolution law under discussion it should emulate the following, in our view crucial, principles, in order to match certain fundamental requirements concerning the capability of emergent and complex behaviour:

(a) As is the case with, say, gauge theory or general relativity, our evolution law on the surmised primordial level should implement the mutual interaction of two fundamental substructures, put sloppily: ‘geometry’ acting on ‘matter’ and vice versa, where in our context ‘geometry’ is assumed to correspond in a loose sense with the local and/or global bond states and ‘matter’ with the structure of the node states.

(b) By the same token the alluded self-referential dynamical circuitry of mutual interactions is expected to favour a kind of undulating behaviour or self-excitation over a return to some uninteresting ‘equilibrium state’ as is frequently the case in systems consisting of a single component which acts directly back on itself. This propensity for the ‘autonomous’ generation of undulation patterns is in our view an essential prerequisite for some form of ‘protoquantum behaviour’ we hope to recover on some coarse-grained and less primordial level of the network dynamics.

(c) In the same sense we expect the overall pattern of switched-on and switched-off bonds to generate a kind of ‘protogravity’.

3. The cellular network as a (random) graph

We start with the introduction of some graph-theoretical definitions (see, e.g., [28]).

Definition 3.1 (Simple locally finite (un)directed graph)).

(a) We write the simple labelled graph as $G := (V, E)$ where $V$ is the countable set of nodes $\{n_i\}$ (or vertices) and $E$ is the set of bonds (edges). The graph is called simple if there do not exist elementary loops and multiple edges; in other words, each existing bond connects two different nodes and there exists at most one bond between two nodes. (We could of course also discuss more general graphs.) Furthermore, for simplicity, we assume the graph to be connected, i.e. two arbitrary nodes can be connected by a sequence of consecutive bonds called an edge sequence or walk. A minimal edge sequence, that is one with each intermediate node occurring only once, is called a path (note that these definitions may change from author to author).

(b) We assume the graph to be locally finite (but this is not always really necessary), that is, each node is incident with only a finite number of bonds. Sometimes it is useful to make the stronger assumption that this vertex degree, $v_i$ (the number of bonds being incident with $n_i$), is globally bounded away from $\infty$. 

Remark. The notions node, vertex or bond, edge are used synonymously in the following (vertex and edge are perhaps more common in pure graph theory). The same applies to notions like vertex degree and valency.

Observation/definition 3.2. Among the paths connecting two arbitrary nodes there exists at least one with minimal length. This length we denote by \( d(n_i, n_k) \). This \( d \) has the properties of a metric, i.e.

\[
\begin{align*}
  d(n_i, n_i) &= 0 & (9) \\
  d(n_i, n_k) &= d(n_k, n_i) & (10) \\
  d(n_i, n_l) &\leq d(n_i, n_k) + d(n_k, n_l). & (11)
\end{align*}
\]

(The proof is more or less evident.)

Corollary 3.3. With the help of the metric one obtains a natural neighbourhood structure around any given node, where

\( U_m(n_i) \) comprises all the nodes, \( n_k \), with \( d(n_i, n_k) \leq m \), \( \partial U_m(n_i) \) the nodes with \( d(n_i, n_k) = m \).

Remark. The restriction to connected graphs is, for the time being, only made for convenience. If one wants to study geometric phase transitions of a more fragmented type, it is easy to include these more general types of graphs. In the context of random graphs (which we will introduce below) one can even derive probabilistic criteria concerning geometric properties such as connectedness, etc.

We mentioned above that, for the time being, we want to concentrate our investigation on the geometric, i.e. graph, properties of our cellular network which are themselves (clock) time dependent. On the other hand, the graphs or networks we are actually interested in are expected to be extremely large (number of nodes \( \geq 10^{100} \)). According to our philosophy they are to emulate the full physical vacuum together with all its more or less macroscopic excitations, or, in other words, the entire evolving universe. Furthermore, the assumed clock-time interval \( \tau \) is extremely short (in fact of Planck-time order). On the other hand, it is part of our working philosophy that the phenomena we are observing in, for example, present-day high-energy physics and, a fortiori, macroscopic physics, are of the nature of collective (frequently large-scale) excitations of this medium both with respect to space and time (in Planck units). In other words, each of these patterns is expected to contain, typically, a huge number of nodes and bonds and to stretch over a large number of clock-time intervals. This then suggests the following approach which has been fruitful again and again in modern physics.

The statistical hypothesis 3.4. Following the above arguments, it makes sense to study so-called graph properties within a certain statistical framework to be further explained below as long as one is interested in patterns which are quasi-macroscopic compared to the Planck scale.

Remark. Similar strategies have been pursued in the investigation of cellular automata (see, e.g., [32, 33])

One strategy, i.e. erecting a probability space of graphs over the given set of nodes, will be introduced now.
The random graph idea 3.5. Take all possible labelled graphs over \( n \) nodes as probability space \( \mathcal{G} \) (i.e. each graph represents an elementary event). The maximal possible number of bonds is \( N := \binom{n}{2} \), which corresponds to the unique simplex graph (denoted usually by \( K_n \)). Give each bond the independent probability \( 0 \leq p \leq 1 \) (more precisely, \( p = \text{the probability that there is a bond between the two nodes under discussion} \)). Let \( G_m \) be a graph over the above vertex set, \( V \), having \( m \) bonds. Its probability is then

\[
pr(G_m) = p^m q^{N-m}
\]

where \( q := 1 - p \). There exist \( \binom{N}{m} \) different labelled graphs \( G_m \), having \( m \) bonds, and the above probability is correctly normalized, i.e.

\[
pr(\mathcal{G}) = \sum_{m=0}^{N} \binom{N}{m} p^m q^{N-m} = (p + q)^N = 1.
\]

This probability space is sometimes called the space of binomially random graphs and is denoted by \( \mathcal{G}(n, p) \). Note that the number of edges is binomially distributed, i.e.

\[
pr(m) = \binom{N}{m} p^m q^{N-m}
\]

and

\[
\langle m \rangle = \sum m pr(m) = Np.
\]

Proof of the latter statement.

\[
\langle m \rangle = \frac{d}{d\lambda} \bigg|_{\lambda=1} \left( \sum \binom{N}{m} \lambda^m p^m q^{N-m} \right) = \frac{d}{d\lambda} \bigg|_{\lambda=1} (\lambda p + q)^N = Np
\]

or, where \( e_i \) are the independent Bernoulli \((0, 1)\)-variables (as to this notion cf, e.g., [34]) belonging to the bonds (i.e. \( e_i \) is the random variable which takes the two values \((1, 0)\) with the respective probabilities \((p, q = 1 - p)\)):

\[
\langle e_i \rangle = p \quad \text{hence} \quad \langle m \rangle = \sum_{i=1}^{N} \langle e_i \rangle = Np.
\]

(The systematic use of the above Bernoulli variables also leads to conceptual clarifications in other calculations.)

A variant approach 3.6. A slightly different probability space can be constructed by considering only graphs with a fixed number, \( m \), of bonds and give each the probability \( \binom{N}{m}^{-1} \) as there are exactly \( \binom{N}{m} \) of them. The corresponding probability space, \( \mathcal{G}(n, m) \), is called the space of uniform random graphs.

The latter version is perhaps a little bit more common in pure mathematics as this concept was introduced mainly for purely combinatorial reasons which have nothing to do with our own strand of ideas. The whole theory was developed by Erdős and Rényi in the late 1950s and early 1960s to cope with certain notorious (existence) problems in graph theory (for more information see, e.g., [35, 38], brief but concise accounts can also be found in chapter VII of [28] or in [39]).
Observation 3.7. The two random graph models behave similarly if $m \approx pN$. Note, however, that there exists a subtle difference between the two models anyway. In the former model all elementary bond random variables are independent, while in the latter case they are (typically weakly) dependent.

(While being plausible this statement needs a proof which can be found in, for example, [35].)

The really fundamental observation made already by Erdős and Rényi (a rigorous proof of this deep result can, for example, be found in [38]) is that there are what physicists would call phase transitions in these random graphs. To go a little bit more into the details we have to introduce some more graph concepts.

Definition 3.8 (Graph properties). Graph properties are certain particular random variables (indicator functions of so-called events) on the above probability space $\mathcal{G}$; i.e. a graph property, $Q$, is represented by the subset of graphs of the sample space having the property under discussion. To give some examples: (a) connectedness of the graph; (b) existence and number of certain particular subgraphs (such as subsimplices, etc); (c) other geometric or topological graph properties, etc.

Remark. In addition to these variables there are other more general random variables (graph characteristics) describing the fine structure of graphs, some of which we will introduce below.

In this context Erdős and Rényi made the following important observation.

Observation 3.9 (Threshold function). A large class of graph properties (e.g. the monotonically increasing ones, cf the above-cited literature) have a so-called threshold function, $m^*(n)$, so that for $n \to \infty$ the graphs under discussion have property $Q$ almost certainly for $m(n) > m^*(n)$ and almost certainly not for $m(n) < m^*(n)$ or vice versa (more precisely, for $m(n)/m^*(n) \to \infty$ or 0; for the details see the above literature). The above version applies to the second kind of graph probability space, $\mathcal{G}(n,m)$. A corresponding result holds for $\mathcal{G}(n,p)$ with $p(n)$ replacing $m(n)$. That is, by turning on the probability $p$, one can drive the graph one is interested in beyond the phase transition threshold belonging to the graph property under study. Note that, by definition, threshold functions are only unique up to ‘factorization’, i.e. $m^*_2(n) = O(m^*_1(n))$ is also a threshold function.

Example 3.10 (Connectedness). The threshold function for the graph property connectedness is

$$m^*(n) = \frac{1}{2} n \log(n) \text{ respectively } p^*(n) = \frac{\log(n)}{n}. \quad (18)$$

Note that with the help of the above observation, i.e. for $m \approx p(n^2)$, we have for $n$ large: 
$
\binom{n}{2} \approx n^2/2$ and hence $pn^2/2 \approx n \log(n)$, i.e. $p \approx \log(n)/n$

In the following our main thrust will be towards the development of the concept of protospacetime as an order parameter manifold or superstructure floating in our network QX and the concept of physical points. We therefore illustrate the method of random graphs, graph properties and graph characteristics by applying it to a particular feature which is of importance in the following.

Definition 3.11 (Subsimplices and cliques). With $G$ a given fixed graph and $V_i$ a subset of its vertex set $V$, the corresponding induced subgraph over $V_i$ is called a subsimplex (ss), $S_i$, or complete subgraph, if all its nodes are connected by a bond. In this class there exist certain maximal subsimplices (mss), that is, every addition of another node destroys this property. These mss are called cliques in combinatorics and are the candidates for our physical points carrying a presumably rich internal (dynamical) structure.
We consider all possible graphs, \( G \), over the fixed vertex set \( V \) of \( n \) nodes. For each subset \( V_i \subset V \) of order \( r \) (i.e. the number of elements) we define the following random variable:

\[
X_i(G) := \begin{cases} 
1 & \text{if } G_i \text{ is an } r\text{-simplex}, \\
0 & \text{otherwise}
\end{cases}
\]  

where \( G_i \) is the corresponding induced subgraph over \( V_i \) with respect to \( G \in \mathcal{G} \) (the probability space). Another random variable is then the \textit{number of } r\text{-simplices in } G, \text{ denoted by } Y_r(G), \text{ and we have}

\[
Y_r = \sum_{i=1}^{\binom{n}{r}} X_i
\]

with \( \binom{n}{r} \) the number of \( r \)-subsets \( V_i \subset V \). With respect to the probability measure introduced above we then have for the \textit{expectation values}:

\[
\langle Y_r \rangle = \sum_i \langle X_i \rangle
\]

and

\[
\langle X_i \rangle = \sum_{G \in \mathcal{G}} X_i(G) \text{pr}(G_i = r\text{-simplex in } G).
\]

With the sum running over all \( G \in \mathcal{G} \) and where \( X_i \) is one or zero we obtain

\[
\langle X_i \rangle = \text{pr}(G_i \text{ an } r\text{-simplex, } G - E_i \text{ an arbitrary graph})
\]

where \( G - E_i \) is the remaining graph after all the edges belonging to \( G_i \) have been eliminated. This yields

\[
\langle X_i \rangle = p \binom{r}{\ell} \sum_{G' \in \mathcal{G}'} \text{pr}(G')
\]

where \( \mathcal{G}' \) is the probability space of graphs over \( V \) with all the bonds \( E_i \) being omitted. The maximal possible number of bonds belonging to \( \mathcal{G}' \) is

\[
|E'| = \binom{n}{2} - \binom{r}{2}.
\]

Each of these bonds can be on or off with probability \( p \) or \( (1 - p) \). To each graph of \( \mathcal{G}' \) belongs a unique labelled sequence of \( p \)'s and \( q \)'s and every such sequence does occur (i.e. with either \( p \) or \( q \) at label \( i \)). Hence we have

\[
\sum_{G} \text{pr}(G') = (p + q)^{|E'|} = 1^{|E'|} = 1
\]

and we obtain

\[
\langle X_i \rangle = p \binom{r}{\ell}.
\]

The case that such an \( r \)-simplex is already maximal, i.e. is actually a clique, can be calculated as follows. The addition of a single further vertex would destroy the property of being an \( ss \). In other words, each of the vertices in the graph \( G \), not lying in the node-set \( V_i \), is connected with the above vertex set, \( V_r \), by fewer than \( r \) bonds. This can now be encoded quantitatively as follows. The probability that the induced subgraph \( G_r \) over \( r \) arbitrarily chosen vertices is already an \( mss \) is the product of the \textit{independent} probabilities that it is an \( ss \) and that \textit{each} of the
remaining \((n - r)\) vertices is connected with \(V_r\) by fewer than \(r\) bonds. The latter probability is \((1 - p^r)^{n-r}\), hence

\[
\Pr(G_r \text{ is a clique}) = (1 - p^r)^{n-r} p^\binom{r}{2}.
\]  
(28)

There are now exactly \(\binom{n}{r}\) possible \(r\)-subsimplices over the node set \(V\). We therefore arrive at the following important conclusion.

**Conclusion 3.12 (Distribution of subsimplices and cliques).** The expectation value of the random variable number of \(r\)-subsimplices is

\[
\langle Y_r \rangle = \binom{n}{r} p^\binom{r}{2}.
\]  
(29)

For \(Z_r\), the number of \(r\)-cliques (i.e. maximal! \(r\)-simplices) in the random graph, we have then the following relation:

\[
\langle Z_r \rangle = \binom{n}{r} (1 - p^r)^{n-r} p^\binom{r}{2}.
\]  
(30)

It is remarkable, both physically and combinatorially, that these quantities, as functions of \(r\) (the order of subsimplices), have quite a peculiar numerical behaviour. We are, among other things, interested in the typical order of cliques (where typical is understood in a probabilistic sense and order means the number of nodes in the clique).

**Observation/definition 3.13 (Clique number).** The maximal order of cliques (i.e. the greatest order of occurring cliques) contained in \(G\) is called its clique number, \(\text{cl}(G)\). It is another random variable on the probability space \(\mathcal{G}(n, p)\).

An analysis of the above expression \(\langle Y_r \rangle = \binom{n}{r} p^\binom{r}{2}\) as a function of \(r\) shows that it is typically very large (for \(n\) sufficiently large) for all \(r\) lying in a certain interval below some critical value, \(r_0\), and drops rapidly to zero for \(r > r_0\).

**Conclusion 3.14.** From the above one can infer that this value \(r_0\) is a good approximation for the above-defined clique number of a typical random graph, depending only on \(p\) and \(n\). In other words, it approximates the order of the largest occurring cliques in a typical random graph. An estimate for \(r_0\) is

\[
r_0 \approx 2 \log(n) / \log(p^{-1}) + O(\log \log(n))
\]  
(31)

(cf chapter XI.1 of [35]).

**Remark.** A numerical analysis of the geometric web of cliques and various of its characteristics is given in section 4.

**Conclusion 3.15.** By estimating the variance of \(Z_r\) we can conclude that the typical orders of occurring cliques are expected to lie between \(r_0(n)/2\) and \(r_0(n)\).

**Remark 3.16.** To make the above reasoning perhaps more transparent, it is again helpful to exploit the properties of the elementary \((0, 1)\)-edge-variables \(e_i\). The probability that \(r\) arbitrarily selected bonds exist in the random graph is \(\Pr(e_1 = \cdots = e_r = 1) = p^r\), the complimentary possibility (i.e. that some of these bonds are missing), hence has the probability \((1 - p^r)\).
Résumé 3.17 (Random graph approach). There is of course no absolute guarantee that our network, following a deterministic evolution law and typically reaching after a certain transient one of possibly several attractors, can in every respect be regarded as the evolution of true random graphs. In other words, its behaviour cannot be entirely random already by definition. Quite the contrary, we expect a shrinking of the huge accessible phase space during its evolution which manifests itself on a more macroscopic level as pattern creation.

The underlying strategy is rather the following. At each clock time step, \( G(t) \) is a graph having a definite number of active bonds, say \( m \). Surmising that \( G(t) \) is sufficiently generic or typical we may be allowed to regard it as a typical member of the corresponding family \( G(n, m) \) or \( G(n, p) \) (at least as far as certain gross features are concerned). Via this line of inference we are quite confident of being able to obtain some clues concerning the qualitative behaviour of our network, the microscopic time evolution of which is too erratic to follow in detail. As to this working philosophy, the situation is not so different from the state of affairs in many areas of, say, ordinary statistical physics or complex dynamical systems. However, nevertheless, a more detailed analysis (underpinned by concrete numerical simulations) of the validity or limits of such an ansatz would be desirable but has to be postponed in order not to expand this paper too much (cf, for example, the completely similar situation in the context of cellular automata, being expounded to some extent in the above-mentioned literature).

Some steps in this direction were taken by our former student Th Nowotny as part of his diploma thesis. Other perhaps characteristic numerical deviations between theoretical results based on certain \( \text{a priori} \) statistical assumptions and computer simulations of concrete models were also found in [36] by Antonsen, who presented an approach which is different from ours in several respects but belongs to the same general context.

4. The emergence of (proto)spacetime

In this core section of our investigation we are going to describe how the presumed underlying (discrete) fine structure of our continuous spacetime may look. We want, however, to emphasize at this point that, while most of the details and observations we will present are rigorously proved, the overall picture is still only hypothetical. That is, we are able to describe a (we think) fairly interesting scenario in quite some detail but are not yet able to show convincingly that our network actually evolves into exactly such a phase as we are going to expound in the following under one of the dynamical laws we described above. We will, in fact, show that the unfolding of something like an extended, macroscopically stable spacetime structure having, \( \text{a fortiori} \), an integer macroscopic dimension, is quite a subtle phenomenon from a purely probabilistic point of view and that only the understanding of the nature of these subtleties will lead us to the correct class of dynamical laws.

The central theme of this paper is the description and analysis of a certain superstructure, \( ST \), emerging within our network \( QX \) as a consequence of a process which can be interpreted as a geometrical phase transition. In this picture, \( ST \), which we experience as spacetime on a more macroscopic scale, plays the role of an \( \text{order parameter manifold} \). Its emergence signals the transition from a \( \text{disordered} \) and \( \text{chaotic} \) initial phase to a phase developing a \( \text{near-/far-order} \), i.e. a \( \text{causal structure} \), and stable \( \text{physical points} \) or \( \text{lumps} \) (Menger).

Our qualitative picture concerning the initial scenario is the following (more details can be found in, for example, section 4 of [24]). The network, \( QX \), started from a presumably densely entangled \( \text{initial phase} \), \( QX_0 \), in which on average every pair of nodes was connected.
by an active bond with high probability (i.e. it was almost a simplex). This chaotic initial phase was characterized by extremely large fluctuations of local node/bond charges and very short-lived/short-ranged correlations. This is a result of the dense entanglement and the kind of network laws we described in section 2 (they typically favour to some extent an overshooting). There are, in particular, no stable cliques or lumps and no traces of a kind of protospace.

We conjecture now that, under certain favourable conditions, that is, an appropriately chosen dynamical law together with the occurrence of a suitably large collective spontaneous deviation of the network state in some region, the network is capable of leaving this chaotic initial phase, starting from such a nucleation centre, and unfolds towards a new more ordered attractor, the above-mentioned phase \( QX/ST \). This geometric phase transition is triggered, respectively, accompanied by an annihilation of a certain fraction of active bonds (in the language of our network laws: \( J_{ik} = 0 \)) due to too large differences in neighbouring node charges. In this process the individual and incoherent elementary fluctuations are expected to be reorganized in a more macroscopic pattern (in the language of synergetics they become slaved; see, e.g., [37]). We note in passing that what we indicated above may be described in a more conventional context as the big-bang scenario and it remains an ambitious task to fill in all the intermediate steps.

We now describe in broad outline our idea of the underlying discrete substratum of spacetime and relate it, in a next step, to the more rigorously defined mathematical correlatives within the random graph framework. We want to emphasize that, for the time being, the only structural or quantitative input we are employing to encode the effects of the presumed geometric phase transition is the decrease in the number of occupied bonds in the wiring diagram of the network. That is, we will mainly be occupied with the description and analysis of the structural and geometric changes, occurring in the underlying graph, when the bond probability, \( p(t) \), decreases with \( t \) the clock time.

The qualitative picture 4.1 (Physical points).

(a) Physical points have a (presumably rich) internal structure, i.e. they consist of a (presumably) large number of nodes and bonds. In the words of Menger they are lumps.
(b) We suppose that what we are used to describing as fields at a spacetime point (in fact, rather distributions in, for example, quantum field theory), are really internal excitations of these lumps.
(c) In order to have a qualitative measure to tell the physical points apart, that is, to discern what happens within a certain point or between different points, we identify the physical points with particularly densely connected subgraphs of our network or graph. This then motivates our interest in maximal subsimplices or cliques as natural candidates.
(d) Typically (i.e. if a certain fraction of bonds has been eliminated), some of these lumps overlap with each other in a stronger or weaker sense, forming so-called local groups, while others will cease to overlap. This will then establish a kind of protocausality or nearfar-order in our protospace and will be one of the central topics in our analysis.

To obtain an idea of how the order of the typical cliques depends on the bond probability, \( p \), we apply the formula for the clique number, provided in section 3, to a graph having, say, \( 10^{100} \) nodes with \( p \) varying (we omitted the \( \log(\log(n)) \)-term). The results are shown in table 1. We will proceed by compiling a couple of simple observations concerning (m)ss, proofs being partly omitted (if they are obvious).
Table 1.

| $p$  | 0.9 | 0.8 | 0.7 | 0.6 | 0.5 | 0.4 | 0.3 | 0.2 | 0.1 |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| $r_0$ | 4444 | 2083 | 1333 | 909 | 666 | 500 | 400 | 285 | 200 |

**Observation 4.2.**

(a) If the node degree, $v_i$, of $n_i$ is smaller than $\infty$ then $n_i$ can lie in at most a finite set of different simplices, an upper bound being provided by the number of different subsets of bonds emerging from $n_i$, that is $2^v_i$.

(b) The set of subsimplices is evidently partially ordered by inclusion.

(c) Furthermore, if $S$ is a simplex, each of its subsets is again a simplex (called a face).

(d) It follows that each of the chains of linearly ordered simplices (containing a certain fixed node) is finite. The corresponding length can be calculated in a similar way as in (a) by selecting chains of sets of bonds, ordered by inclusion. In other words, each chain has a maximal element. By the same token each node lies in at least one (but generically several) mss.

(e) An mss where $n_i$ is a member can comprise at most $(v_i + 1)$ nodes, in other words, its order is bounded by the minimum of these numbers when $n_i$ varies over the mss.

**Proof of (a).** Assume that $S_k, S_l$ are two different simplices containing $n_i$. By definition $n_i$ is linked with all the other nodes in $S_k$ or $S_l$. As these sets are different by assumption, the corresponding subsets of bonds emerging from $n_i$ are different. On the other hand, not every subset of such bonds corresponds to a simplex (there respective endpoints need not form a simplex), which proves the upper bound stated above. $\square$

**Observation 4.3.** The class of simplices, in particular the mss, containing a certain fixed node, $n_i$, can be generated in a completely algorithmic way, starting from $n_i$. The first level consists of the bonds with $n_i$, an end node, the second level comprises the triples of nodes (triangles), $(n_{i0}n_{i1}n_{i2})$, with the nodes linked with each other and so forth. Each level set can be constructed from the preceding one and the process stops when an mss is reached.

**Remark.** Note that at each intermediate step, i.e. having already constructed a certain particular subsimplex, one has in general several possibilities as to how to proceed. On the other hand, a chain of such choices may differ at certain places from another one but may lead in the end to the same final simplex (in other words, being simply a permutation of the nodes of the former simplex).

Denoting the (m)ss under discussion by capital $S$ with certain indices or labels attached to it, this process can be pictorially abbreviated as follows:

$$S(n_0 \rightarrow \cdots \rightarrow n_k). \quad (32)$$

With $S(n_0 \rightarrow \cdots \rightarrow n_k)$ given, each permutation will yield the same mss, i.e.

$$S(n_0 \rightarrow \cdots \rightarrow n_k) = S(n_{\pi(0)} \rightarrow \cdots \rightarrow n_{\pi(k)}). \quad (33)$$

Furthermore, each mss can be constructed in this way, starting from one of its nodes. Evidently, this could be done for each node and for all possible alternatives as to the choice of the next node in the above sequence.
Definition 4.4. Let $G_v$ be a class of subgraphs of $G$.

(a) $\cap G_v$ is the graph with $n \in V_{\cap G_v}$ if $n \in$ every $V_{G_v}$, $b_{ik} \in E_{\cap G_v}$ if $b_{ik} \in$ every $E_{G_v}$.

(b) $\cup G_v$ is the graph with $n \in V_{\cup G_v}$ if $n \in V_{G_v}$ for at least one $v$, $b_{ik} \in E_{\cup G_v}$ if $b_{ik} \in E_{G_v}$ for at least one $v$.

As every node or bond belongs to at least one mss (as can be easily inferred from the above algorithmic construction), we have

Corollary 4.5.

$$\cup S_v = G. \quad (34)$$

After these preliminary remarks we now turn to our main task, that is, the analysis of the web of these mss as the elementary building blocks of the next higher level of organization.

4.1. The embryonic epoch

In its surmised transition from the almost maximally connected and chaotic initial phase to the fully developed phase, $QX/ST$ (i.e. $QX$ plus superstructure $ST$), the underlying graph passes through several clearly distinguishable epochs. In the following we study two main phases of the network. The first one, which we call embryonic, is characterized by a still pronounced common overlap of the emerging lumps or physical points. That is, they can still interact directly with each other which implies that there exists no true far-order or larger distances on the network of lumps. This epoch should be typical (in our picture) for the infinitesimal time interval just after the big bang.

We begin with the epoch where only a small fraction of the bonds is shut off. Let us, for example, assume that $\alpha$ bonds with

$$1 \ll \alpha < n/2 \ll n(n-1)/2 = N \quad \text{for } n \text{ large} \quad (35)$$

are temporarily dead with $n$ the order of the graph or network (number of nodes) and $N$ the maximal possible number of bonds. In other words, the network is supposed to be still near the initial phase. We observe that $\alpha$ arbitrarily selected bonds can connect at most $k \leq 2\alpha$ different nodes, hence there still exist at least $(n-k)$ nodes which are maximally connected, namely they are spanning a still huge subsimplex $S' \subset G$. On the other hand, there are at most $k \leq 2\alpha$ nodes with one or more incident bonds missing in the corresponding induced subgraph.

$V_G$ can hence be split in the following way:

$$V_G = V_S \cup V_N \quad (36)$$

with $V_N$ the unique set of those nodes so that for each node $n_i$ in $V_N$ there exist other nodes (also lying in $V_N$) with the respective bonds to $n_i$ missing. $V_S$, on the other hand, is the set of remaining nodes which are maximally connected (by construction); i.e. they form an ss,

$$|V_N| = k \leq 2\alpha. \quad (37)$$

Definition 4.6. $[\cup G_i]$ is the induced subgraph spanned by the nodes occurring in $\cup V_i$. Note that in general $[\cup G_i] \supset \cup G_i$, that is, it may rather be called its ‘closure’.
Observation 4.7.

(a) The simplex $S'$ is contained in each of the occurring mss, $S_\nu$, i.e.

$$S' \subset \cap S_\nu$$

and it holds a fortiori $S' = \cap S_\nu$.  \hfill (38)

Hence we have

$$|V_{S'}| = |V_{\cap S_\nu}| \geq n - 2\alpha.$$  \hfill (39)

(b) Note that $S'$ itself is never maximal since $[S' \cup n_1]$ is always a larger simplex with $n_1 \in N$ and $[S' \cup n_1]$ is the induced subgraph spanned by $V_{S'}$ and $n_1$.

(c) To each maximal simplex $S_\nu \subset G$ belongs a unique maximal subsimplex $N_\nu \subset N$ with

$$S_\nu = [S' \cup N_\nu].$$  \hfill (40)

(d) It is important for what follows that $S'$ can be uniquely characterized, without actually knowing the $S_\nu$, by the following two properties:

1. $S'$ is an ss so that all bonds connecting nodes from $V_{S'}$ with $V - V_{S'}$ are ‘on’.
2. $S'$ is maximal in this class of ss, that is, each node in $V - V_{S'}$ has at least one bond missing with respect to the other nodes in $V - V_{S'}$. An induced subgraph in $G$, having these properties, is automatically the uniquely given $S'$!

Corollary 4.8. From the maximality of the $N_\nu$ follows a general structure relation for the $\{S_\nu\}$ and $\{N_\nu\}$:

$$\nu \neq \mu \Rightarrow S_\nu \neq S_\mu \Rightarrow N_\nu \neq N_\mu$$  \hfill (41)

and neither

$$N_\nu \subset N_\mu \quad \text{nor} \quad N_\mu \subset N_\nu$$  \hfill (42)

namely there always exists at least one $n_\nu \in V_{N_\nu}$ s.t. $n_\nu \notin V_{N_\mu}$ and vice versa.

Proof of the above observation.

(a) Starting from an arbitrary node $n \in G$, it is by definition connected with all the nodes in $S'$, since if say $n, n'$ are not connected they both belong to $N$ (by definition); i.e. irrespective of how we proceed in the construction of some $S_\nu$, $S'$ can always be added at any intermediate step, hence $S' \subset \cap S_\nu$. On the other hand, assume that $n \notin \cap S_\nu$. This implies that $n$ is connected with each node in $\cup S_\nu$. We showed above that $\cup S_\nu = G$, hence $n$ is connected with all the other nodes, i.e. it is not in $N$, that is, $n \in S'$, which proves the statement.

(b) As $n \in N$ is connected with each $n' \in S'$ (by definition of $N$ and $S'$), the subgraph $[S' \cup n]$ is again a (larger) simplex.

(c) We have $S' \subset S_\nu$ for all $\nu$, hence

$$S_\nu \neq S_\mu \quad \text{implies} \quad N_\nu \neq N_\mu$$  \hfill (43)

with $N_{\nu,\mu}$ the corresponding subgraphs in $N$.

With $S_\nu$ being a simplex, $N_\nu$ is again a subsimplex which is maximal in $N$. Otherwise $S_\nu$ would not be maximal in $G$.

On the other hand, each $S_\nu = [S' \cup N_\nu]$ is uniquely given by a maximal $N_\nu$ in $N$ as each node in $N$ is connected with all the nodes in $S'$. □
We see from the above that as long as \( \alpha \), the number of dead (missing) bonds, is sufficiently small, i.e. \( 2\alpha < n \), there necessarily exists a non-void overlap \( S' \), among the class of \( \text{mss} \), \( S_\nu \). This overlap will become smaller as \( \alpha \) increases. By the same token the number of \( \text{mss} \) will increase for a certain range of the parameter \( \alpha \), while the respective size of the \( \text{mss} \) will shrink. The above results hold for each given graph. On the other hand, the above unique characterization of \( S' \) in (d) makes it possible to attack the problem of the order of \( S' \) within the framework of random graphs in a more quantitative manner. Given a member \( G \) of \( G(n, p) \), \( S' \) is fixed by (d) of the above observation. We are interested in the probability of \( S' \) having, say, \( r \) nodes.

The strategy is, as usual, to try to express the probability of such a configuration within \( G \) as the product of certain more elementary and (if possible) independent probabilities. Unfortunately, this turns out to be relatively intricate in the above case and we are, at the moment, only able to provide certain upper and lower bounds for the probability under discussion. As this example shows that such questions may not always have simple and straightforward answers, it is perhaps worthwhile to dwell a little bit on this point.

The typical difficulties one usually encounters in this context are the following. The structure of the set of graphs in \( G(n, p) \) having a prescribed property may be rather complicated, so that it is difficult to avoid multiple counting of members when trying to calculate the order of such a set. A frequent reason for this is the intricate entanglement of the various pieces of a complicated graph, a case in point being the above description of \( S' \) in \( G \). In our case the peculiar entanglement can be seen as follows.

Selecting \( r \) arbitrary vertices, the probability that the corresponding induced subgraph forms an \( \text{ss} \) is \( p^{(\text{ss})} \) (see section 3). If this subgraph is to qualify as \( S' \), i.e. \( S' = \cap S_{\nu} \cup S_{\nu} = G \), each of the nodes in \( N \) is connected with every node in \( S' \). The probability for this property is \( p^{(n-r)} \). The difficult part of the reasoning concerns the subgraph \( N \). We call the probability that \( N \) has just the structure being described above, \( pr(N) \).

The following observation is helpful. As \( S' \) is unique in \( G \), i.e. occurs only once, the corresponding random variable, \( X_{S'} \), that we hit at such an \( S' \) of order \( r \), when browsing through the set of induced \( r \)-subgraphs, is zero with at most one possible exception, that is, if \( S' \) has just the order \( r \). Therefore, the corresponding expectation value of \( X_{S'} \) is, by the same token, also the probability of the property \( (S'_r) \).

Conclusion 4.9. The probability that a random graph contains such an \( S' \) of order \( r \) is

\[
pr(S'_r) = \binom{n}{r} p^{(\text{ss})} p^{(n-r)} pr(N). \tag{44}
\]

As far as we can see, it is not easy to disentangle \( pr(N) \) into more elementary independent probabilities and master the complex combinatorics. Therefore, we will, at the moment, only give (possibly crude) upper and lower bounds.

\( N \), having \( (n-r) \) nodes, is characterized as follows. Labelling the nodes from (1) to \( (n-r) \), none of them are allowed to have the maximal possible degree (with respect to \( N \)), i.e. \( (n-r-1) \). The first step is simple. Starting with, say, node (1), the probability that at least one bond is missing is the complement of the probability that all possible bonds are present, i.e. \( (1 - p^{(n-r-1)}) \). The following steps will, however, become more and more cumbersome. Take, for example, node (2). In the above probability both the probability that either the bond \( b_{12} \) is missing or that it is not, are already contained. If \( b_{12} \) is not missing then some other bond \( b_{1j} \), \( j > 2 \) are present is admissible, in the latter case this possibility is forbidden. Depending on which choice we make at each step the algorithmic construction bifurcates in a somewhat
involved manner. Evidently, this first step yields a crude upper bound on \( pr(N) \). Making at each step (i) the particular choice that there are always missing bonds among the bonds pointing to nodes (j) with \( j > i \) provides a lower bound. Hence we have

**Conclusion 4.10.**

\[
(1 - p^{n-r-i}) \geq pr(N) \geq \prod_{j=1}^{n-r-i} (1 - p^{n-r-j}) = \prod_{j=1}^{n-r-i} (1 - p^j)
\]

(45)

and for \( pr(S' = \emptyset) \):

\[
(1 - p^{n-i}) \geq pr(S' = \emptyset) \geq \prod_{j=1}^{n-i} (1 - p^j).
\]

(46)

**Observation 4.11.** The lower bound is interesting! Perhaps surprisingly, the occurring product is an important number-theoretic function belonging to the field of partitions of natural numbers (see any good textbook about combinatorics or the famous book of Hardy and Wright [40], the standard source being [41]). Our above random graph approach offers the opportunity to (re)derive and prove this number-theoretic formula by purely probabilistic means, i.e. give it an underpinning which seems to be, at first glance, quite foreign. We will come back to this interesting point elsewhere.

It is important to have effective estimates for the regime of probabilities, \( p(n) \), so that \( S' \) is empty with a high probability. According to our philosophy this signals the end of the embryonic epoch, where all the supposed protopoints still overlap (and hence are capable of direct interaction) and the beginning of the unfolding of a new phase with, as we hope, a more pronounced near- and far-order among the physical points.

Such an estimate can, in fact, be provided with the help of the above inequality. We have

\[
pr(S' = \emptyset) > \prod_{i=1}^{n-i} (1 - p^i) > \prod_{i=1}^{\infty} (1 - p^i) = \sum_{k=0}^{\infty} a_k p^k
\]

(47)

for \( 0 < p < 1 \). The following (highly non-trivial) observation is due to Euler (cf the above-mentioned literature for more recent proofs):

**Theorem 4.12.**

\[
\prod_{i=1}^{\infty} (1 - p^i) = \sum_{k=0}^{\infty} (-1)^k \left( p^{k(3k+1)} + p^{k(3k-1)} \right) = 1 - p - p^2 + p^5 + p^7 \ldots
\]

\[
\approx 1 - p - p^2
\]

(48)

for \( p \) small.

**Conclusion 4.13.** For \( p \) near zero, \( S' \) is empty with arbitrarily large probability \( \approx 1 \).

This shows that there exists, in fact, a regime of small \( p \)-values where the embryonic epoch no longer prevails. This holds the more so for an \( n \)-dependent \( p \) (which is very natural) and \( p(n) \nwarrow 0 \). On the other hand, note that there exists a possibly substantial class of bond configurations which have, up to now, been excluded in the above estimate, the inclusion of which would increase the relevant probability further.

To get a feeling of how good the above exact estimate may be, we construct an (we think, not untypical) example. The construction goes as follows. Take \( 2k \) nodes, choose a subset \( G_1 \) consisting of exactly \( k \) nodes \((n_1, \ldots, n_k)\), make \( G_1 \) a simplex. With the remaining \( k \) nodes
(\(n_1', \ldots, n_k'\)) we proceed in the same way, i.e. we now have two subsimplices \(G_1, G_1'\). We now choose a one-to-one map from \((n_1, \ldots, n_k)\) to \((n_1', \ldots, n_k')\), say \(n_i \rightarrow n_i'\). We now connect all the \(n_i\) with the \(n_j'\) except for the \(k\) pairs \((n_i, n_i')\). The graph \(G\) so constructed has

\[ |E_G| = 2k(2k - 1)/2 - k = 2k(2k - 2)/2. \]  

We see from this that, as in our above network scenario, the number of missing bonds is a relatively small fraction. We can now make the following sequence of (easily proved) observations.

**Observation 4.14.** \(G_1 = S_1\) is already an mss as each \(n_i' \in G_1'\) has one bond missing with respect to \(G_1\). One obtains a new mss by exchanging exactly one \(n_i\) with its partner \(n_i'\), pictorially:

\[ [S_1 - n_i + n_i'] \]  

yielding \(k\) further mss. One can proceed by constructing another class of mss, now deleting \((n_i, n_j)\) and adding their respective partners, i.e.

\[ [S_1 - n_i - n_j + n_i' + n_j'] \]  

This can be done until we end up with the mss

\[ [S_1 - n_1 - \cdots - n_k + n_1' + \cdots + n_k'] = S_1'. \]  

The combinatorics goes as follows:

\[ |\{\text{mss}\}| = \sum_{v=0}^{k} \binom{k}{v} = (1 + 1)^k = 2^k \]  

i.e. our \(2k\)-node graph (with \(k\) bonds missing) contains exactly \(2^k\) mss of order \(k\). Evidently \(S' = \emptyset\).

We showed above that \(S'\) is non-empty as long as \(\alpha\), the number of missing bonds, is smaller than \(n/2\), \(n\) the order of the graph \(G\), since \(\alpha\) bonds can at most connect \(2\alpha\) different nodes. On the other hand, \(\alpha = n/2\) implies

\[ |E_G| = \binom{n}{2} - n/2 = n(n - 2)/2 \]  

or an average vertex degree

\[ \langle v(n) \rangle_s = n - 2 \]  

which is still very large. The example constructed in the preceding observation has

\[ n = 2k, \quad \alpha = k = n/2, \quad S' = \emptyset. \]  

In other words, the parameter \(\alpha\) is just the critical one given in observation 4.7. One may hence surmise that \(\alpha = n/2\) is perhaps the threshold for \(S' = \emptyset\) in the sense that, say,

\[ pr(S' = \emptyset) = O(1) \]  

for \(\alpha \gtrsim n/2\).
On the other hand, within the framework of random graphs, we have obtained the rigorous but presumably not optimal estimate

$$p r(S' = \emptyset) > \prod_1^{\infty} (1 - p^j) \approx 1 - p - p^2$$

(59)

for \(p\) small. For a large graph \(p = \frac{1}{2}\) implies, however,

$$\alpha \approx \frac{1}{2} \left( \frac{n}{2} \right) \gg n/2.$$  

(60)

That is, there is still a wide gap between these two values, a point which needs further clarification.

4.2. The unfolded epoch

For \(p\)-values away from \(p \approx 1\) we expect the clique-graph, \(C\), to consist of a huge number of cliques, \(S_i\), each being surrounded by its local group, that is cliques having an overlap with \(S_i\), while there is no longer an overlap with the majority of cliques in the graph. In other words, in this scenario the clique graph is much more unfolded and has at least the potential to figure as a kind of protospace\(\text{t}ime\). This epoch shall now be analysed in more detail.

From the general analysis we learned that the majority of cliques have an order lying between \(r_0/2\) and \(r_0\), \(r_0 \approx \log(n)/\log(1/p)\). The expected number of \(r\)-cliques is \(\langle Z_r \rangle = {n \choose r} (1 - p^r)^n p^C\). We want the average order of cliques to be much greater than one, so that the internal structure of the lumps or physical points is still sufficiently complex but, on the other hand, the order should be infinitesimal compared to the number of nodes in the graph, that is

$$1 \ll r_0 \ll n.$$

(61)

In a first step we will now estimate the number of relevant cliques in the network or graph for given bond-probability \(p\), which implies an \(r_0\) of the above type (the numerical calculation follows below). We approximate factorials by the following version of Stirling’s formula:

$$x! \approx \sqrt{2\pi x} x^x e^{-x}.$$  

(62)

This yields

$$\left( \frac{n}{r} \right) \approx \frac{n^r}{\sqrt{2\pi r r^r (n - r)^{n-r}}}$$

(63)

(the exponentials drop out).

As we have to deal with extremely large numbers it is useful to take logarithms:

$$\log \left( \frac{n}{r} \right) \approx n \log(n) - (n - r) \log(n - r) \approx r \log(n)$$

(64)

where we have left out marginal contributions and approximated \(\log(n - r)\) by \(\log(n)\). We hence have for the assumed range of \((n, r)\)-values

$$\left( \frac{n}{r} \right) \approx n^r.$$  

(65)
With this estimate we can now estimate $\langle Z_r \rangle$, the expected number of $r$-cliques.

$$\log(\langle Z_r \rangle) \approx r \log(n) + n \log(1 - p') + \frac{1}{2}r^2 \log(p).$$  \quad (66)$$

For the range of parameters we are interested in we can approximate $\log(1 - p')$ by $-\log(e)p'$. Evaluating the latter expression for the values of the list in table 1, for example, we see that irrespective of the huge prefactor $n$ the second term in the above equation becomes negligibly small (for example, for $p = 0.7$, $r = 1000$ and $n = 10^{100}$ we have $np' \approx 10^{-54}$. Henceforth we will therefore work with the formula

$$\log(\langle Z_r \rangle) \approx r \log(n) + \frac{1}{2}r^2 \log(p).$$  \quad (67)$$

To get an idea of the size of the numbers involved we provide, in table 2, a list for the parameters $p = 0.7$, $r_0 = 1333$, taken from table 1.

### Table 2.

| $r$  | $400$ | $500$ | $650$ | $1200$ | $1300$ | $1400$ |
|------|-------|-------|-------|--------|--------|--------|
| $\log(\langle Z_r \rangle)$ | $2.9 \times 10^4$ | $3.2 \times 10^4$ | $3.3 \times 10^4$ | $1.2 \times 10^4$ | $3 \times 10^3$ | $5 \times 10^2$ |

Observation 4.15. We see that over a large scale of $r$-values, ranging from, say, $100 \leq r \leq 1200$, we have a $\langle Z_r \rangle$ of $O(10^{104})$. On the other hand, we noted above that from general estimates [35] one can infer that the typical cliques are supposed to have $r$-values of between $r_0/2$ and $r_0$. The above list shows that, at least for our possibly extreme values, this number is still appreciable below $r_0/2 \approx 650$. On the other hand, the upper critical value, i.e. $r_0 \approx 1300$ seems to be more reliable.

In a next step we want to estimate the average number of cliques which overlap with a given fixed clique, $S_0$ of a generic order between, say, $r_0/2$ and $r_0$. This set of cliques we have dubbed the local group or infinitesimal neighbourhood of the respective lump or physical point, $S_0$. This problem winds up as a calculation of certain conditional probabilities. As a subset of graphs we take those containing a fixed $r$-clique, $S_0$. The respective relative probability (which is the same as the measure of the set of graphs under discussion) is $(1 - p^r)^{n - r}p^r$. In this subclass of graphs we now consider another subset, consisting of the graphs containing yet another $r'$-clique, having an overlap of order $l$ with the fixed given $S_0$.

The (conditional) probability that an $r$-set and an $r'$-set of nodes span an $r$-clique, $S_0$, $r'$-clique, $S_1$, respectively, with common overlap of order $l$ is

$$P_{r,l} \frac{p^{(r)}}{(1 - p^r)^{n - r}p^{r}}$$

with $P_{r,l}$ to be calculated in the following way. We have to apply the same principle as in the calculation of individual cliques, i.e. incorporate the fact that they are maximal members in the class of $(r, r')$-simplices. On the other hand, they now overlap and we have to avoid a double counting of elementary events (i.e. configurations). Maximality plus overlap is then encoded in the following way:

$$P_{r,l} = (1 - p^{r + r' - l})(1 - p^{r - l})(1 - p^{r' - l})(1 - p^{r' - l})$$

(69)
with \( P_{r,l} \) describing the probability that there is no other node in the graph under consideration which has all its links to nodes in \( S_0, S_1 \) being occupied.

In this new probability space of graphs, containing a fixed \( r \)-clique, \( S_0 \), there exist \( \binom{r}{m-r} \) possibilities to place an \( r' \)-clique with overlap \( l \). Hence we have the result

**Conclusion 4.16.** The expected number of \( r' \)-cliques, overlapping with a fixed given \( r \)-clique, \( S_0 \), is

\[
\sum_{l=1}^{r-1} \langle N(S_0; r', l) \rangle
\]

with

\[
\langle N(S_0; r', l) \rangle = \frac{\binom{r}{n-r} p^{n-r} (1-p)^r}{(1-p')^n p^r} P_{r,l}. \tag{71}
\]

To estimate the cardinality, \( \langle N_{loc,gr} \rangle \), of the local group of \( S_0 \), one has first to sum over all \( r' \)-values between, say, \( r_0/2 \) and \( r_0 \). We showed in table 2 that the numbers are rather uniform over a wide range of \( r \)-values so that it is sufficient to estimate the cardinality for some generic value, \( r' \approx r \), say, and simply multiply this result with the width of the interval (e.g. \( 10^3 \) in the above table). In a first step we have to convince ourselves that, as in the above numerical calculation, we can again neglect the \( P_{r,l} \)-factors if we are only interested in estimates of the kind order of. Taking logarithms we see that all terms arising from \( \log(P_{r,l}) \) are either very small or at most of \( O(1) \). The same holds for the term \( (n-r) \log(1-p') \). They can be neglected compared with the other terms and we hence end up with

\[
\sum_{l=1}^{r-1} \langle N(S_0; r', l) \rangle \approx \sum_{l=1}^{r-1} \left( \binom{r}{n-r} p^{n-r} (1-p)^r \right) \frac{1}{(1-p')^n p^r} P_{r,l}. \tag{72}
\]

On the other hand, the expected number of \( r' \)-cliques not overlapping with \( S_0 \) is

\[
\langle N(S_0; r', l = 0) \rangle = \binom{n-r}{r'} (1-p')^{n-r} p^{r'} \approx \frac{n-r}{r'} p^{r'} \tag{73}
\]

which for \( r' \approx r \) is of the same order as \( \langle Z_r \rangle \) itself (with \( n \approx (n-r) \) for \( n \gg r \)). In other words, we already observe a certain degree of unfolding on the level of the clique graph as compared to the underlying graph, \( G \). However, nevertheless, as all the estimates are only of order of type, we cannot use the latter estimate as a substitute for (72) which can also be very large in principle.

**Remark 4.17.** Previously we approximated \( \log\binom{n}{r} \) in expressions like \( \log\binom{n}{r} P^{r(n)} \) by \( r \log(n) \), neglecting the term \(-r \log(r)\). This suffices as long as \( n \gg r \gg 1 \). However, if \( l \approx r' \) as in the above expression, we have to take a further term into account in the approximation of the Stirling formula (cf. equation (64)).

We have

\[
\log\langle N(S_0; r', l) \rangle \approx l (r \log(r) - log(l)) + (r' - l) (\log(n) - \log(r' - l))
\]

\[
+ \frac{1}{2} (r')^2 - l^2 \log(p) \tag{74}
\]

For \( l \ll r' \) we can still neglect the correction term and obtain

\[
\log\langle N(S_0; r', l) \rangle \approx (r' - l) (\log(n) + \frac{1}{2} (r')^2 \log(p) \approx \log(Z_{r'}) - l \log(n). \tag{75}
\]
For $l \ll r'$ (and $r, r'$ both generic, i.e. of roughly the same order) we have on the other hand
\[
\log\langle N(S_0; r', l) \rangle \approx l(\log(r) - \log(l)) + (r' - l)(\log(n) - r'\log(p))
\]
which is much smaller than the preceding expression.

**Conclusion 4.18 (Typical size of local group in clique graph).** From the above we infer that the dominant contribution comes from cliques having a relatively small overlap with the given lump, $S_0$. Neglecting small factors and correction terms we have approximately for generic $r, r'$:
\[
|\text{local group}| \approx \sum_{l \ll r'} |\text{number of cliques in } G| / n^l
\]
with the main contribution coming from $l = 1$.

That is, the size of the typical local group is roughly the $n$th part of the total number of cliques.

**Observation 4.19.** From tables 1 and 2 we see that the local groups in the clique graph stemming from a typical random graph are still very large, but much smaller than the original neighbourhood of a node in the underlying graph, $G$.

**Remark 4.20.** If we assume that members of the local group should have an appreciable overlap (i.e. $l \gg 1$), the size would shrink drastically as can be seen from the above general formula.

4.3. Some desirable properties of (proto)spacetime

What we have achieved so far is the following. We did not tackle the full problem head-on, i.e. to show that something like an unfolded (proto)spacetime emerges under the evolution of a certain critical dynamical law from a chaotic initial phase. Instead of that we replaced this (presumably) very complicated problem by a slightly different one. Assuming a suitable statistical generic behaviour, we described the emergence of a two-storey structure within the framework of random graphs. More precisely, we showed that the wiring diagram of the network, i.e. the underlying (clock-time-dependent) graph, passes generically through various epochs with decreasing bond probability, $p$, starting from an embryonic epoch with almost all cliques overlapping for $p$ in the infinitesimal neighbourhood of one, and unfolding, for $p$ away from one, into a web of cliques (the clique graph), displaying a certain degree of near-/far-order. In this epoch or phase each lump has a certain local neighbourhood of other lumps overlapping with it, these local groups now being significantly smaller than the whole web.

We want, however, to stress the following point. The emergence of something highly organized (from a dynamical and macroscopical point of view) such as macroscopic spacetime cannot expected to be a mere probabilistic effect coming about as a byproduct without the injection of a (possibly very peculiarly chosen) class of dynamical laws. To learn more about the (un)naturalness of such a dynamical process, we want to show in the remaining space what kind of results can already be achieved in the pure random graph approach and compare them with what is needed in order to have a fully developed spacetime picture.

As to the underlying graph, $G$, we worked with numerical data like
\[
n = 10^{100} \quad p = 0.7 \quad r_0 = 1333.
\]
The number of generic cliques was then approximately
\[
\langle Z_r \rangle \approx 10^{100}.
\]
The size of the typical local group of a clique was roughly
\[ \langle N_{\text{loc gr}} \rangle \approx \langle Z \rangle / n \] (80)
which is, by the same token, the average vertex degree, \( \langle v_{\text{cl}} \rangle \), in the clique graph, i.e. the average number of bonds ending on a vertex.

From this, one can estimate the bond probability, \( p_{\text{cl}} \), in the clique graph, with an existing bond meaning overlap.
\[ p_{\text{cl}} = (n - 1)/\langle v_{\text{cl}} \rangle \approx \langle Z \rangle / \langle N_{\text{loc gr}} \rangle \approx n^{-1}. \] (81)

**Observation 4.21.** With the numerical data we used above we have an edge probability in the clique graph
\[ p_{\text{cl}} \approx n^{-1} = 10^{-100} \] (82)
which is very small (compared with, for example, the \( p = 0.7 \) of the underlying graph).

Another important question is whether the clique graph is generically connected. The threshold we presented in example 3.10 is
\[ p^\star(n) = \log(n)/n \approx 10^4/10^{10^4} = 10^{-(10^4-4)} \ll p_{\text{cl}}. \] (83)

**Conclusion 4.22.** For the numerical data we employed the web of lumps is almost certainly connected.

Take, on the other hand, a typical class of graphs which are very regular, admitting large or infinite distances between nodes and have a pronounced near-far-order, i.e. lattices. Usually the vertex degree, \( v \), is both low and constant over the entire lattice. For node number, \( n \), to infinity we hence have
\[ v \ll p^\star(n)(n - 1) = v^* \approx \log(n) \] (84)
that is, a typical lattice graph is a connected member of the regime of random graphs which are generically disconnected. This shows that graphs which may be of particular interest for physics and numerical approximation are, on the other hand, not generic in a probabilistic sense. In the following we show that, in fact, a too large \( \langle v \rangle \) or \( p \) usually implies a very low diameter of the graph which is typically equal to two.

**Definition 4.23 (Diameter).** The diameter, \( \text{diam}(G) \), of a graph is defined as the greatest distance between two arbitrary nodes of the graph, i.e.
\[ \text{diam}(G) = \max_{i,j} d(n_i, n_j). \] (85)

Surprisingly, this diameter behaves rather uniformly over a wide range of \( p \)-values in a random graph. In order to derive a quantitative estimate we will proceed as follows. In a first step we calculate the probability that an arbitrary pair of nodes, \( (n_i, n_j) \), is neither directly connected nor connected with each other via an intermediate node, \( x \), by a pair of bonds. For a fixed pair \( n_i, n_j \) and \( x \) running through the set of remaining nodes this probability is (following the same line of reasoning as above) \((1 - p)(1 - p^2)^{(n-2)}\). There are \( \binom{n}{2} \) such pairs. We denote by \( A \) the random variable (number of such pairs) and have
\[ \langle A \rangle = \sum_{k \geq 1} k \ pr(A = k) = \binom{n}{2} (1 - p)(1 - p^2)^{(n-2)}. \] (86)
On the other hand,
\[ (A) \geq \sum_{k \geq 1} pr(A = k) = pr(\exists \text{ such a pair}) =: pr(A) = 1 - pr(A = 0) \]  
(87)

with
\[ pr(A = 0) = pr(\text{diam}(G) \leq 2). \]  
(88)

Hence we have

**Observation 4.24.** The probability that \( \text{diam}(G) \leq 2 \) is \( 1 - pr(A) \geq 1 - (A) \), which is of course only non-trivial for \( (A) < 1 \).

**Remark.** Note that there is exactly one graph having a diameter equal to one, i.e. the simplex over the \( n \) vertices. In the large-\( n \) limit this contribution becomes marginal.

For \( n \) large, more precisely, \( n \to \infty \), one can now calculate the \( n \)-dependent \( p^*(n) \)-threshold so that \( \text{diam}(G) = 2 \) holds almost certainly for \( p(n) > p^*(n) \) (cf [35]). We content ourselves with a simpler but nevertheless impressive result. With \( p > 0 \) fixed and \( n \to \infty \) we infer from the above estimate
\[ \log((A)) \approx 2 \log(n) - \log 2 + \log(1 - p) + (n - 2) \log(1 - p^2) \]  
(89)

which is for fixed \( p \) and \( n \to \infty \) dominated by the last term which goes to \( -\infty \).

**Conclusion 4.25.** With \( p > 0 \) fixed and \( n \to \infty \), \( G(n, p) \) has diameter two almost certainly. Now inserting the numerical values for our clique graph \( C \) in the above formula, we see that the very small \( p_{cl} \) is overcompensated by the huge \( n_{cl} \) with the result that not only the underlying graph \( G \) but also the derived clique graph \( C \) still has a diameter equal to two almost certainly.

The above result is, in fact, not so bad, whereas it exhibits the limits of the pure random graph approach. It rather shows that the unfolded web of lumps we expect to emerge as a kind of (proto)spacetime is presumably the result of some dynamical fine tuning which, among other things, accomplishes the possibility of larger distances between the lumps. This fine tuning is expected to shrink the available phase space so that the set of graphs we have to take into account is in effect much smaller and more special. Note that on a continuum scale the driving force which fuels the unfolding towards larger scales is a combined effect of Einstein gravity and the (not completely understood) behaviour of the quantum vacuum. On our much more primordial scale there will be a certain counterpart in the form of a peculiar class of dynamical network laws, which may then lead, for example, on the scale of the web of lumps or a coarser scale, to an effective theory which serves as a proto form of macroscopic gravity (in the mentioned companion paper we attempt to show this for quantum theory).

We want to remark that we recently came across a (we think) deep speculative remark by Penrose [42] in which he ventured the idea that, perhaps, some sort of (quantum) non-locality is already necessary on ordinary scales in order to understand the formation of crystals and the like. In the case of quantum theory we will argue that it just turns out to be such an effective theory, encoding non-local effects in a seemingly local framework. We expect the same kind of non-locality to be at work in the organization of our web of lumps.

A network or graph characteristic which has the potential to describe the large-scale organization of our web of lumps is a concept we introduced and analysed in [1], i.e. the notion of (intrinsic) graph dimension or scaling dimension. We do not want to go into the details here as the topic poses difficult questions in their own right. We only briefly discuss
this concept and relate it to the questions we are investigating in this paper. We begin by providing two definitions of a possible graph dimension, which turn out to be the same on generic configurations but may be different on exceptional ones (cf [1]).

**Definition 4.26.** With $U_k(x)$ describing the $k$-neighbourhood of the node $x$, i.e. the nodes, $n_i$, with distance $d(x, n_i) \leq k$, $\partial U_k(x)$ its boundary, i.e. the nodes having $d(x, n_i) = k$, we define

$$D_1^k(x) := \ln |U_k(x)| / \ln(k)$$  
$$D_2^k(x) := \ln |\partial U_k(x)| / \ln(k) + 1$$  

and call

$$\overline{D} := \limsup D^{(1,2)}_k, \quad \underline{D} := \liminf D^{(1,2)}_k$$

the upper, lower dimension (for more details see [1]). If they coincide we call them the respective dimensions of the graph or network.

We showed in [1] that for, say, regular lattices, embedded in a background spacelike $\mathbb{R}^n$, the above intrinsic dimensions coincide with the usual Euclidean dimensions of the embedding space (a property well known to hold for fractal-like dimensions such as ours). On the other hand, we constructed examples (i.e. graphs) for practically every (non-integer) value of $D^{(1,2)}_k$ in [1], showing that it is, in fact, a useful generalization of the ordinary continuous dimension concept. Furthermore, we could show, among other things, that $D$, provided it exists at all, is practically independent of the special node $x$ but is rather a property of the whole graph.

From a physical, more specifically, dynamical point of view the above dimensions exactly encode just the characteristics of the system which really matter physically (e.g. in the context of phase transitions). They count basically the number of new partners (nodes) which can be reached or influenced within clock-time $k, k \to \infty$. As the continuum space dimension of our universe seems to be rather special (presumably being even fine-tuned), it is of central importance to understand how and why some definite (and integer) dimension such as, for example, three may emerge in our network (cf also the discussion in [19]). Our approach makes it possible to analyse this great and long-standing question in a dynamical and microscopic way. As this very intricate subject poses a variety of subtle problems in their own right, we refrain from treating it here (due to the limits of space).

We want to close the paper by emphasizing an aspect of our two-storey structure of the network which will be of crucial importance for fundamental physics in general and for understanding quantum theory in particular.

**Observation 4.27 (The two-storey concept).**

(a) Given a network or graph, $G$, of the above kind, we can construct its associated clique graph, $C_G$, and thus establish the two-storey concept, mentioned already in the introduction. We hence have two kinds of distances and metric (causal) relations in the network: one defined by the original node distance in $G$, the other defined by the distance between lumps (defined by overlap) in $C_G$.

(b) It is important that two lumps, $S_1, S_2$, which are some distance apart in $C_G$, may nevertheless be connected by a certain (possibly substantial) number of interbonds or short paths, extending from nodes in $S_1$ to nodes in $S_2$ (see the construction of the cliques described in the preceding sections).
(c) That is, there may exist two types of information transport or correlation being exchanged in the network. A relatively coherent (and possibly classical) one, exchanged among the lumps, and a more stochastic and less coherent one (of possibly quantum nature) between individual groups of nodes lying in the respective lumps and which can be almost instantaneous.

Remark. These phenomena will be employed and analysed in the forthcoming paper [9] on quantum theory. For some technical results concerning degrees of interdependence of sets of nodes see the final section of [24].

References

[1] Nowotny T and Requardt M 1998 Dimension theory on graphs and networks J. Phys. A: Math. Gen. 31 2447 (Nowotny T and Requardt M 1997 Preprint hep-th/9707082)
[2] Nowotny T and Requardt M 1999 Pregeometric concepts on graphs and cellular networks Chaos Solitons Fractals 10 469 (invited paper)
[3] Requardt M 1998 Cellular networks as models for Planck-scale physics J. Phys. A: Math. Gen. 31 7997 (Requardt M 1998 Preprint hep-th/9806135)
[4] Requardt M 2000 Spectral analysis and operator theory on (infinite) graphs… Preprint math-ph/0001026 (submitted to J. Phys. A: Math. Gen.)
[5] Penrose R 1971 Quantum Theory and Beyond ed T Bastin (Cambridge: Cambridge University Press)
Penrose R 1972 Magic without Magic, John Archibald Wheeler ed I R Klauder (San Francisco, CA: Freeman)
[6] Smolin L 1991 Space and time in the quantum universe Conceptual Studies of Quantum Gravity (Einstein Studies vol 2) ed A Ashtekar and J Stachel (Boston, MA: Birkhäuser)
[7] ‘t Hooft G 1988 Deterministic and quantum mechanical systems J. Stat. Phys. 53 323
‘t Hooft G 1990 Quantization of discrete deterministic theories Nucl. Phys. B 342 471
[8] Isham C J 1995 Structural issues in quantum gravity Lecture given at the GR14 A New Approach to Functional Analysis on Graphs Conf. (Florence, 1995) Preprint gr-qc/9510063
[9] Requardt M 1998 Preprint forthcoming
[10] Menger K 1970 Albert Einstein Physicist Scientist 3rd edn, ed P A Schilpp (London: Cambridge University Press)
[11] Stachel J 1991 Einstein and quantum mechanics Conceptual Problems of Quantum Gravity (Einstein Studies vol 2) ed A Ashtekar and J Stachel (Boston, MA: Birkhäuser)
[12] 1996 Chaos Solitons Fractals 7 (special issue)
El Nasch M S 1996 Time symmetry breaking and Cantorian space-time Chaos Solitons Fractals 7 499
El Nasch M S 1998 Von Neumann geometry and E∞∞ quantum spacetime Chaos Solitons Fractals 9 2023
[13] Neumann J V 1962 Collected Works vol IV, ed Taub A H (New York: Pergamon)
[14] Waits D J and Strogatz S H 1998 Collective dynamics of small-world networks Nature 393 441
[15] Rosen N 1947 Statistical geometry and fundamental particles Phys. Rev. 72 298
[16] Roy S 1998 Statistical Geometry and Applications to Microphysics and Cosmology (Dordrecht: Kluwer)
[17] Schweizer B and Sklar A 1983 Probabilistic Metric Spaces (New York: North-Holland)
[18] Nagels G 1985 Space as a bucket of dust Gen. Rel. Grav. 17 545
[19] Bombelli L, Lee J, Meyer D and Sorkin R 1987 Space-time as a causal set Phys. Rev. Lett. 59 521
[20] Kauffman L H 1998 Noncommutativity and discrete physics Physica D 120 125
[21] Alexandroff P and Hopf H 1974 Topologie (New York: Springer) (reprint)
[22] Frohlich J and Nielsen H B 1991 Origin of Symmetries (Singapore: World Scientific)
[23] Requardt M 1999 Space-time as an orderparameter manifold in random networks and the emergence of physical points Preprint gr-qc/9902031
[24] Isham C J 1992 Canonical quantum gravity and the problem of time Lectures Presented at the NATO Advanced Study Institute on Recent Problems in Mathematical Physics (Salamanca, June 1992) Preprint gr-qc/9210011
[25] Butterfield J and Isham C J 1999 On the emergence of time in quantum gravity The Argument of Time ed J Butterfield (Oxford: Oxford University Press) to appear
(Butterfield J and Isham C J 1999 Preprint gr-qc/9901024)
[26] Nowotny Th Diploma Thesis a compilation of some of the extensive computer simulations can be found on the web site, http://www.physik.uni-leipzig.de/~nowotny/research.html (commentary in German)
Spacetime as a statistical geometry of lumps in random networks

[28] Bollobás B 1998 Modern Graph Theory (Graduate Texts in Mathematics) (New York: Springer)
[29] Toffoli T and Margolus N 1987 Cellular Automaton Machines (Cambridge, MA: MIT Press)
[30] Markopoulou F and Smolin L 1997 Causal evolution of spin networks Nucl. Phys. B 508 409
   (Markopoulou F and Smolin L 1997 Preprint gr-qc/9702025)
   Markopoulou F 1997 Dual formulation of spin network evolution Preprint gr-qc/9704013
[31] Borissov R 1997 Graphical evolution of spin network states Phys. Rev. D 55 6099
   (Borissov R 1996 Preprint gr-qc/9606013)
[32] Wolfram S 1984 Universality and complexity in cellular automata Physica D 10 1
   Wolfram S 1983 Statistical mechanics of cellular automata Rev. Mod. Phys. 55 601
[33] Farmer D 1982 Information dimension and the probabilistic structure of chaos Z. Naturf. 37a 1304
[34] Feller W 1968 An Introduction to Probability Theory vol I (New York: Wiley)
[35] Bollobas B 1985 Random Graphs (London: Academic)
[36] Antonsen F 1994 Random graphs as a model for pregeometry Int. J. Theor. Phys. 33 1189
[37] Haken H 1978 Synergetics (New York: Springer)
[38] Bollobas B 1986 Combinatorics (London: Cambridge University Press)
[39] Karoński M 1995 Random graphs Handbook of Combinatorics vol I, ed R Graham, M Grötschel and L Lovász
   (Amsterdam: Elsevier) ch 6
[40] Hardy G H and Wright E M 1960 An Introduction to the Theory of Numbers (Oxford: Oxford University Press)
[41] Andrews G E 1998 The Theory of Partitions (Cambridge: Cambridge University Press)
[42] Penrose R 1998 The Geometric Universe ed S A Huggett et al (Oxford: Oxford University Press) (afterword)