A simple three-dimensional real-time reliable cellular array

Peter Gács*
Boston University

John Reif†
Harvard University, MIT

January 12, 1987

Abstract

We build a three-dimensional array of unreliable cellular automata that can simulate a universal Turing machine (more generally, a one-dimensional universal iterative array) reliably. This is the first reliable real-time simulation. The encoding is simple repetition, and no decoding is needed. The construction is based on Toom's work.

*Supported in part by NSF grant DCR-8405270. At the time of this research, the author was Visiting Scientist at the MIT.

†This work was supported in part by the Office of Naval Research contract N 00014-80-C-0647. Author's current address: Duke University, Durham, NC.

Preprint of paper appearing in *Journal of Computer and System Sciences*, Vol. 36, No. 2, April 1990, pp. 125-147.

1 Introduction

The theoretical problem of reliable computation is based on the assumption that computing devices of arbitrary size must be built from a few types of elementary components. Each component makes errors with some frequency independent of the size of the device to be built. What are the architectures enabling us to deal with all combinations of errors likely to arise for devices of a given size? Is it possible to achieve reliability without sacrificing the speed of computation?

We will consider the case when a failure does not incapacitate the component permanently, only causes it, in the step when it occurs, to violate its rule of operation. In the following steps, the component obeys its rule of operation again, until the next error. The case of permanent component failure may be of greater practical importance, but it has not been investigated in the same generality. (However, see [6] for some elegant geometrical results in a similar model.) There are reasons to believe that many of the techniques developed for the case of transient failure will be applicable for the case of permanent failure (see the Conclusions).

Let us remark that any model of reliable computation with unreliable components must use massive parallelism. Indeed, any information temporarily stored anywhere during computation is subject to decay and therefore must be actively maintained.

In 1953, von Neumann proposed a design for reliable Boolean circuits. In his model, each component had some constant probability of failure. For a circuit consisting of n perfect components, he built a circuit out of $O(n \log n)$ components, computing the same function. (For an efficient realization of his ideas, see [1].) In 1968, Taylor, using Gallager's low-density parity-check codes, constructed a Boolean circuit out of $O(K)$ unreliable components and memory elements, capable of holding K bits of information for a polynomial number of steps. This construction was improved by Kuznietsov, using an idea of Pinsker, increasing the storage time to an exponential function of K.

All the above constructions suffer from the same deficiency: the circuits use a rather intricate connection pattern which cannot be realized in three-dimensional space with wires of constant length. On the other hand, the natural assumption about a wire is that as its length grows, its probability of failure converges to 1.

A cellular space (medium) is a lattice of automata in, say, three-dimensional space where every automaton takes its input from a few of its closest neighbors. Such devices are now sometimes known as "systolic arrays". Typically, all automata are required to have the same transition function and are connected to the same relative neighbors, i.e. the device is translation-invariant. The spatial uniformity suggests the possibility of especially simple physical realization.

Cellular media are desirable computing devices, and it is easy to construct a one-dimensional cellular space that is a universal computer. (Take a one-tape Turing machine.) These devices are sometimes called iterative arrays. However, up to now there is no simple design for a reliable cellular medium made out of unreliable components. In 1974, solving a problem arising in mathematical physics, Toom constructed some two-dimensional infinite cellular spaces capable of holding a bit of information for any number of steps. In 1976, Tsirel'son constructed a one-dimensional medium holding a bit in n cells for a nearly exponential number of steps. However, Tsirel'son uses components of three different kinds, and the kind of the components changes in both space and time according to a grand plan not subject to errors.

In [2], using some of Kurdyumov's ideas, Gács constructed a one-dimensional cellular space M capable of reliable computation. Any computation using a one-dimensional array of K perfect automata and taking T steps can be modelled on M using $K \log^c(KT)$ cells and $T \log^c(KT)$ steps (c is a constant). However, the construction and the initial encoding of the inputs are quite complex.

In 1984, Reif noticed that a three-dimensional real-time reliable computing medium can be constructed using one of Toom's error-correcting rules in two dimensional slices and the rule of an arbitrary one-dimensional medium across the slices. The reliability of the infinite version of this construction follows from [9]. However, Toom's proof uses an elaborate topological argument, developed from the Peierls argument of statistical physics, that we could not adapt to an efficient finite version of the theorem. Gács used the technique of "k-sparse sets of errors" developed in [2] to give a more straightforward proof of Toom's theorem. Using this result, one can now do the following. Given K cells of a one-dimensional medium D working for T steps, one can build a three-dimensional medium M on the set $\{1, \ldots, K\} \times \mathbf{Z}_m^2$. Here $m = \log^{1+\epsilon}(KT)$ where $\epsilon \to 0$ as $KT \to \infty$, and $\mathbf{Z}_m$ is the group $\{0, \ldots, m-1\}$ of remainders modulo m (with $\mathbf{Z}_\infty = \mathbf{Z}$, the set of integers). When started with the appropriate input (each site in a torus-shaped slice $\{i\} \times \mathbf{Z}_m^2$ receives the same input symbol), the medium M will simulate D reliably step-for-step, without any time delay, for T steps. (This statement is made precise in the Theorem below.)

The strange topology of the torus $\mathbf{Z}_m^2$ is not necessary. A torus can be folded up into a square, by doubling up in each of the two dimensions simultaneously. However, the cells on the edges of the square will then obey a somewhat different transition function, i.e. we lose homogeneity.

This construction contrasts remarkably with the one given in [2]. A newer version of this latter construction works with a constant space- redundancy factor ($O(K)$ cells needed for K cells of input) and logarithmic time-delay. The present result has no time delay, i.e. it is *real-time*, and has a space factor $\log^{2+2\epsilon}(KT)$. Moreover, the encoding is just repetition, with no decoding.

The proof below gives the ridiculously small lower bound (10^{-25}) on the largest error probability permitting already reliable computation. However, the present paper was not written with the intent to find good constants, only to find out the asymptotic behavior of the redundancy. For better constants, probably very different techniques are needed. We made our estimates explicit only on the insistence of one of our referees. Bennett's experiments on Toffoli's Cellular Automata Machine give convincing empirical evidence that Toom's medium is nonergodic at error probabilities below 0.05. Let us note that for 2 dimensions and Toom's rule, our theoretical bound would also drop.

2 Statement of the result

As a notational convenience, we define, in analogy with Pascal and real analysis, the intervals

$$
\begin{aligned}
[a\,..\,b] &= \{x \in \mathbf{Z} : a \le x \le b\}, \\
[a\,..\,b) &= \{x \in \mathbf{Z} : a \le x < b\},
\end{aligned}
$$

etc. Let $G = [-1\,..\,1]$. For any function f and any subset H of its domain of definition, we will denote by $f(H)$ the restriction of f to H. We will apply this notion a little loosely, so if $x[1]\dots x[n]$ is a string and $1 < i < n$ then by $x[i+G]$ we mean the string $x[i-1]x[i]x[i+1]$. An r-dimensional *medium (cellular automaton, cellular space)* D is given by a finite set $S = S_D$ of states and a transition function D that assigns a value $D(x) \in S$ to all functions $x : G^r \to S$.

The set W of *cells*, or *sites* is a direct product of the form

$$\mathbf{Z}_{m_1} \times \cdots \times \mathbf{Z}_{m_r}.$$

For some positive integer time limit T, let $V = [0\,..\,T] \times W$. A function $x : V \to S_D$ is called an *evolution* of the medium D over W. An evolution x is called a *trajectory* if for all $t < T$ and $u \in W$ we have

$$x[t+1, u] = D(x[t, u + G^r]).$$

For $r = 1$ this is $D(x[t, u-1], x[t, u], x[u+1])$. We say that a *failure* occurred in x at $(t+1, u)$ (with respect to D) if (2) does not hold. If y is a trajectory then the event $x[v] \neq y[v]$ is a *deviation* (of the evolution x from y). Intuitively, a medium is reliable if it can keep the number of deviations small despite the occasional occurrence of failures.

Example: Toom's two-dimensional error-correcting medium R. This medium can be defined for any finite state-space. Let us define first the majority function $\mathrm{Maj}(x, y, z)$. If two of the three arguments coincide then their common value is the value of Maj, otherwise $\mathrm{Maj}(x, y, z) = x$. Let

$$H = \{(0,1), (-1,-1), (1,-1)\},\ R(x[G^2]) = \mathrm{Maj}(x[H]). \qquad (2.1)$$

In words, rule R says: "to obtain your next state, take the majority among your current state and the state of your northern, southwestern and southeastern neighbor". Any constant function is a trajectory of R. To fix the example, suppose $S_D = \{0, 1\}$, and y is the identically 0 trajectory. For some evolution x, we have a failure in $x[t, u]$ if it is not the value obtained by voting from the triple $x[t-1, u+H]$. We have a deviation if $x[t, u] = 1$.

Let $(\xi[t, u] \,:\, t \leq T, u \in W)$ be a system of random variables. We say that ξ is a ρ–perturbation of D if for each subset B of V the probability that for *all* $v \in B$ a failure occurs in ξ at v is at most $\rho^{|B|}$. (This condition is satisfied if the failures occur independently with probability ρ.) We will say that ξ is a ρ-perturbation of a trajectory y if it is a ρ-perturbation of D with $\xi[0,u] = y[0,u]$ for all u in W. Our goal is to find situations in which if the probability ρ of failure is small then the probability of deviations is also small: in other words, failures do not accumulate, there is a "healing effect".

Let us be given an arbitrary one-dimensional medium D, and the (possibly infinite) integers K, T. For any integer m, we define the sets $W = \mathbf{Z}_K \times \mathbf{Z}_m^2$ (our space) and $V = [0\,..\,T] \times W$ (our space-time). For any trajectory x of D on $[0\,..\,T] \times \mathbf{Z}_K$, we define the function $y : V \to S_D$ by

$$y[t,n,i,j] = x[t,n].$$

Thus, each cell of the one-dimensional medium D is repeated m^2 times, on a whole "plane" (a torus for finite m) in W.

THEOREM 2.1 *There is a three-dimensional medium M with S_D as its set of states, and function $\epsilon(n)$ such that $\lim_{n\to\infty} \epsilon(n) = 0$ and the following holds. For all K, T, there is an*

$$m = (\log(KT))^{1+\epsilon(KT)},$$

such that for any trajectory x of D over $[0\,..\,T] \times \mathbf{Z}_K$, if the trajectory y of M over V is defined by (2), then for any $\rho < 10^{-25}$, any ρ-perturbation ξ of y and for all v in V the probability of $\xi[v] \neq y[v]$ is less than $10^{25}\rho$.

The theorem says that in case of the medium M and the trajectories y, the probability of deviation can be uniformly bounded by $10^{25}\rho$. The trajectories y encode (by (3)) an arbitrary computation in an arbitrary cellular space D (e.g. a universal Turing machine), hence this theorem asserts the possibility of reliable computation in three-dimensional space. Moreover, the encoding used is a simple one: repetition $\log^{2+2\epsilon}(KT)$ times. The decoding is even simpler: if a plane of W represents a state s of a cell of D then each cell in this plane will be in state s with large probability.

We made some remarks on the unrealistically small upper bound 10^{-25} in the introduction.

3 The idea of the proof

We define $M(x[G^3]) = D(R(x[-1, G^2]), R(x[0, G^2]), R(x[1, G^2]))$ where the rule R is given in (2.1).

Majority voting over a centrally symmetrical neighborhood would not work. Charles Bennett's computer simulations convincingly show that such rules eliminate small islands by a phenomenon strikingly similar to surface pressure. Surface pressure is a function of curvature, hence a large island will grow if the failure probabilities favor its growth.

If $m = \infty$ then our theorem becomes a special case of Theorem 1 in [9]. Toom probably did not notice this consequence of his results since he gives only examples of media with a finite number of periodic stable trajectories, while our construction gives a continuum of stable trajectories (in three dimensions), and a countable number of periodic stable ones.

The proof in [9] is rather difficult, and what can be obtained from it immediately in the finite case is an m as large as $K + T$. Here we give a new proof of Toom's theorem (at least its special case for the rule M: some generalizations will be obvious), using the technique of k–sparse sets developed, following Kurdyumov's ideas, in [9]. This technique also led to a simpler proof of Tsirel'son's result. Our proof of Toom's theorem is not necessarily shorter or more elegant but its guiding idea is more elementary (even if some details are messy), therefore maybe more widely applicable. Also, it gives a sharper error bound leading to $m = \log^{1+\epsilon}(KT)$.

We suggest the reader look at the special case $D(x, y, z) = y$, i.e. the cells of D just want to keep their state forever, without looking at their neighbors. In this special case, the trajectories $x[t, u]$, $y[t, u]$ do not depend on the time t, and the planes $\{u\} \times \mathbf{Z}_m^2$ of the medium M work independently, each trying to keep the value $x[0, u]$. We can therefore ignore the parameter u, and we are left with the medium R of the Example. Without loss of generality, we can suppose then that $S_D = \{0, 1\}$ and that y is 0 everywhere. The essential part of the general proof given below concerns this special case. The proof is easier to read if we just forget of the extra dimension.

Let ξ be a ρ–perturbation of y. We want to prove that for appropriate values of m, for any t, u we will have $\xi[t, u] = 0$ with probability $1 - c\rho^{0.5}$. Thus whenever some island of cells in state 1 arises as a consequence of failures in ξ, it will be quickly erased, preventing the proliferation of cells in state 1 for a very long time.

Why is Toom's rule R likely to preserve a nearly constant initial configuration? Let us define the linear functions

$$l_1(x,y) = -2x + y,\ l_2(x,y) = 2x + y,\ l_3(x,y) = -y.$$

For an arbitrary subset F of $\mathbf{Z}^2$ we define $m_j(F) = \sup_{v \in F} l_j(v)$. We call $m_j(F)$ the *measurements* of F. For any real numbers a_1, a_2, a_3, let us define the triangle

$$I = L(a_1, a_2, a_3) = \{\, (x,y) \in \mathbf{Z}_m^2 \ :\ l_j(x,y) \le a_j \text{ for } j = 1,2,3 \,\}.$$

The numbers a_j are the measurements $m_j(I)$. Let us call $|I| = (a+b)/2 + c$ the *size* of triangle I. If the size is negative then the triangle is empty. For finite diameter m of the torus, the above definition does not work since inequality is not defined in $\mathbf{Z}_m$. But the definition can be easily extended as long as the size is less than m. Let us add this requirement to the definition of triangles.

For any sets H, I of the plane, let us define

$$d_j(H,I) = \inf\{\, l_j(x) \ :\ x \in I \,\} - m_j(H).$$

This quantity is called the *j-separation* of the sets H and I. It is easy to check the following propositions, which can be considered the "separating hyperplane theorem" for triangles.

LEMMA 3.1 *For triangles H and I, we have*

$$d_j(H,I) = m_j(I) - |I| - m_j(H).$$

Triangles H and I are disjoint if and only if there is a j such that their j-separation is positive.

The following assertion is easy to verify.

LEMMA 3.2 *Suppose that x is a trajectory of R in which at time t, the set of deviations from y is enclosed into the triangle $L(a,b,c)$. Then at time $t+1$, the same set is enclosed in $L(a-1, b-1, c-1)$.*

It is this speed of shrinking, *independent of size* of the set of deviations that distinguishes Toom's rule. It would suffice of $|I|/2$ error-free steps to eradicate the deviations entirely. Unfortunately, this simple argument will not suffice for our purpose because we will allow more frequent failures.

We will need more refined reasoning, based on a certain hierarchical characterization of failures and deviations. We will define the notions of a k-noise and the k-anarchy. The k-anarchy is derived from deviations by a process similar to the one used to derive the k-noise from failures. The main lemma will be analogous to Lemma 3.2. It says that outside the k-noise, the k-anarchy shrinks with constant speed. The k-noise and k-anarchy will be defined recursively, and the proof of the main lemma is inductive. Going up from level k to level $k+1$ can decrease the speed of shrinking of the covering triangles of the k-anarchy only by a certain amount which is the member of a convergent series.

The rule (among similar ones) for which Toom proved in [8] that it can keep a bit of information, was different from R. Instead of squeezing any kind of island in the same way, it squeezed an island of 0's from northeast and southwest, and an island of 1's from northwest and southeast. This proof can be relatively easily converted into a proof for the finite case (with $m = O(\log t)$ for keeping a bit of information in a torus of size m for t steps). However, as Leonid Levin noted, these rules do not seem to be applicable for reliable computation, since the rule D acting across planes can convert an island of 0's into an island of 1's, undoing the effect of squeezing. Another rule, used in earlier version of the present paper, is majority vote among a site and its northern and eastern neighbors. This rule works but is less convenien in proofs than the one adopted here, since it does not squeeze its triangles from all sides.

In the presence of failures, Toom was forced to abandon graphic reasoning, replacing it with an involved topological construction for the estimation of the probability of deviation.

4 Noise

Most failures are isolated. Sometimes a larger group of failures occurs but this is rare. This is the idea behind the notion of a k-noise. Let us be given a set E of failures in our space-time $\mathbf{Z} \times \mathbf{Z}_m^2$. The noise of order 0 is the set of failures. Noise of order $k+1$, (a $k+1$-noise) is the set of all points that have too much k-noise in their k-hypercube. In order to make this notion precise, we only have to say what is "k-hypercube" and what is "too much".

Let c_1, c_1' be integer parameters. We will choose

$$c_1 = 8052. \tag{4.1}$$

We define $r_1 = 104$, $r_k = 2^k$ for $k > 1$. This r_k is going to be the number of failure bursts tolerated within a k-hypercube. (The case $k = 1$ is treated specially in order to increase somewhat the upper bound on ρ in Lemma 4.1 below.) The size P_k of a k-hypercube is defined recursively as follows.

$$\begin{aligned} P_0 &= 1, \\ Q_k &= c_1 k^2 r_k, \\ P_k &= Q_k P_{k-1} \text{ for } k > 0. \end{aligned}$$

Let us denote $\mathsf{P}_k = [0\,..\,P_k)^3$,

$$\mathsf{P}_k[i, j_1, j_2] = (iP_k, j_1 P_k, j_2 P_k) + \mathsf{P}_k.$$

Sets of the form $u + \mathsf{P}_k[i, j_1, j_2]$ are called *k-colonies*. They are ***canonical*** if $u = 0$. Thus, colonies are cubes of width P_k. The notion of canonical colonies makes most sense if the measurements of our torus are divisible by P_k. To keep the Theorem simple, we do not make even this mild assumption. If P_k is relatively prime to the measurements of our torus then every colony is canonical. For a k-colony $E = \mathsf{P}_k[i, j_1, j_2]$, let us call its *neighborhood* the union of E and its 26 neighbor colonies $\mathsf{P}_k[i, j_1, j_2 - 1]$, etc. Each k-colony has a natural partitioning into $k-1$-colonies, which will be called its *subcolonies*.

Let us denote

$$\begin{aligned} C_k &= [0\,..\,P_k)^4 \\ C_k[h, i, j_1, j_2] &= (hP_k, iP_k, j_1 P_k, j_2 P_k) + C_k. \end{aligned}$$

Sets of the form $v + C_k[h, i, j_1, j_2]$ are called *k-hypercubes*. They are ***canonical*** if $v = 0$. We will call a k–hypercube a *k-event*. (Let us note that in a later paper on a two-dimensional error-correcting medium, the notions of a k-event and k-cube will be separated.) Each k-event has a natural partitioning into $(k-1)$- events that will be called its *sub-events*.

The *0-noise* of a set E of space-time is E itself. For $k > 0$, the k-hypercube C belongs to the k*-noise* of a set E of space-time if the number of its sub-events of order $(k-1)$ belonging to the $(k-1)$- noise is greater than r_k.

Example: a 1-event is in the 1-noise if it has at least two failures. In general, we can speak of the k-noise within any union of some disjoint canonical k-events.

LEMMA 4.1 *For $\rho < 10^{-24}$ the following holds. Let ξ be a ρ-perturbation of some medium. Let us denote by $\mathcal{E}$ the (random) set of failures in ξ. Let B be any k-event. For $k > 0$, the probability that B is in the k-noise of $\mathcal{E}$ is less than*

$$p_k = \rho^{5r_2 \cdots r_k}.$$

For $k = 0$ the same probability is ρ by definition. For $k =, 1, 2,$ the lemma gives the estimates ρ^5, ρ^{20}.

PROOF: By the definition of ρ-perturbation, for any set D, the probability of $D \subset \mathcal{E}$ is at most $\rho^{|D|}$. We can therefore increase the probabilities of all sets of failures by assuming that individual failures occur independently with probability ρ. Let us make this assumption. The following statement follows immediately from the definition.

LEMMA 4.2 *Let $B_0, B_1, \ldots$ be disjoint k-events. Let $\mathcal{Q}_i$ be the event that B_i is in the k-noise. The events $\mathcal{Q}_i$ are independent.*

We will prove Lemma 4.1 by induction. The probability of a 0-noise on a 0-cube (one space-time point) is at most ρ, hence the statement holds for $k = 0$.

Let us suppose now that the lemma holds for $k-1$. We prove it for k. To estimate the probability that B belongs to the k-noise we use the inductive assumption and Lemma 4.2. It gives that for each sequence of

$r_k + 1$ disjoint $(k-1)$-events in B the probability that they are all in the $(k-1)$-noise is less than

$$p_{k-1}^{r_k+1}.$$

The total number of possible such sequences is less than

$$\binom{Q_k^4}{r_k+1} < (c_1 k^2 r_k)^{4(r_k+1)}/(r_k+1)! < (ec_1^4 k^8 r_k^3)^{r_k+1}.$$

In the last inequality, we applied the relation $n! > n^n e^{1-n}$.

Let us distinguish now the cases $k = 1$ and $k > 1$. For $k = 1$, the probability that B is in the k-noise is less than $((104)^3 ec_1^4)^{105}\rho^{105}$. We have to show that this quantity is less than ρ^5, i.e. that $1/\rho$ is greater than $((104)^3 ec_1^4)^{1.05}$. Computation shows that the last expression is less than $2 \cdot 10^{23}$.

For $k > 1$ the probability that B is in the k-noise is less than

$$(ec_1^4 k^8 r_k^3)^{r_k+1} p_{k-1} p_k.$$

This is less than p_k if we have

$$(ec_1^4 k^8 r_k^3)^{r_k+1} < \rho^{-5 r_2 \cdots r_{k-1}},$$

i.e.

$$\frac{r_k+1}{5 r_2 \cdots r_{k-1}} \log(ec_1^4 k^8 r_k^3) < -\log \rho.$$

It is easy to see that the left-hand side is monotonically decreasing with k, therefore it is enough to check the inequality for $k = 2$. For $k = 2$ it turns into $ec_1^4 2^1 4 < 1/\rho$. The left-hand side is here less than $5 \cdot 10^{20}$, which completes the proof. ∎

5 Anarchy

5.1 Triangles

In the analog of Lemma 3.2 in the presence of failures, the statement "$x[t,w] = 0$ for all w outside $L(a,b,c)$" must be replaced with the statement "the k-anarchy is confined to $L(a,b,c)$". The notion of k-anarchy used here is the analog of the notion of k-noise introduced in Section 4.

For a set J of triangles we denote by $\bigcup J$ their union and by $|J|$ the sum $\sum_{J \in J} |J|$. We say that J *covers* a set if its union does so.

For a triangle $I = L(a,b,c)$ and positive number d, we define the new triangle

$$D(I,d) = L(a-d, b-d, c-d)$$

called the *deflation* of I by the amount d. For a set $\mathcal{I}$ of triangles, we define

$$D(\mathcal{I},d) = \{ D(I,d) \ : \ I \in \mathcal{I} \}. \tag{5.1}$$

For points $x = (x_1, x_2)$, $y = (y_1, y_2)$ in $\mathbf{Z}^2$ we measure their distance by

$$\max(|x_1 - y_1|, |x_2 - y_2|).$$

Distance is defined similarly in $\mathbf{Z}^3$. For a set E in $\mathbf{Z}^2$ and a positive number d we denote by $\Gamma(E,d)$ the set of points at a distance d or less from E. We call it the *d- blowup* of E. This notion is defined analogously in $\mathbf{Z}^3$. For a set E in $\mathbf{Z}^2$, let us denote by $\Delta(I,d)$ the smallest triangle I with the property that its deflation $D(I,d)$ contains E. These two operations are extended to sets similarly to (5.1). The relation between rectangular and triangular inflation is expressed by the following relation, which can be immediately verified.

$$\Gamma(E,d) \subset \Delta(I,2d). \tag{5.2}$$

Both the deflation of triangles and the blowup of sets are additive, in the sense of the following property:

$$\begin{aligned} D(D(I,c),d) &= D(I,c+d), \\ \Gamma(\Gamma(E,c),d) &= \Gamma(E,c+d). \end{aligned}$$

It is easy to verify the following: If the triangles I and J have nonempty intersection and $|I|+|J| < m$ (where m is the diameter of the torus) then the size of the smallest triangle containing their union is smaller than $|I| + |J|$. We can transform a finite set I of triangles with $|I| < m$ into a set I' of disjoint triangles in the following way: we successively replace any pair of intersecting triangles in the set with the smallest triangle containing their union (this process will be called *merging*), as long as we find intersecting triangles. We have

$$|I'| \leq |I|.$$

Let us define the projection proj of the three-dimensional space to the two-dimensional space of a slice by

$$\text{proj}(x, y, z) = (y, z).$$

There will be only two kinds of projection in the present paper: this projection, called simply *projection*, and the *space-projection* of a set in space-time.

5.2 Shrinking anarchy

A k-noise was defined as a function of a set in space-time that is interpreted as the set of failures in an evolution with respect to the medium M. Similarly, the k-anarchy is a function of a set in space that is interpreted as the set of deviations of an evolution from a certain trajectory of M. Let us thus be given an evolution $x[h, i, j_1, j_2]$ of medium M over a three-dimensional lattice W, and a trajectory $y[h, i, j_1, j_2]$ of M over the same space.

We define the notion of *health* at time t for colonies. The k- *anarchy* within a $(k + 1)$-colony C is the union of sick subcolonies of order k in C. The 0-anarchy at time t is the set of points (i, j_1, j_2) such that $x[t, i, j_1, j_2]$ differs from $y[t, i, j_1, j_2]$.

We will use the parameter

$$c_2 = 192. \tag{5.3}$$

For $k > 0$, a k-colony C is *healthy* at time t if there is a set I of disjoint triangles where $|I|$ is bounded by $c_2 r_k P_{k-1}$ such that the projection of the $(k - 1)$-anarchy in C is covered by $D(I, P_{k-1})$. In general, we can speak of k-anarchy within any union of disjoint canonical k-colonies.

The deflation by P_{k-1} in the definition of the k-anarchy makes sure that the triangles covering the projection of the $(k-1)$-anarchy are well separated from each other.

Let us introduce the "drag"

$$v_k = 0.5 \sum_{i=1}^{k} i^{-2} < 0.5\pi^2/6 < 0.8.$$

From now on, we will speak frequently about the k-noise for the set of failures (within some union of canonical k-events). Let us call it simply the *k-noise*. The k-noise at time t will mean the set of points v such that (t, v) belong to the k-noise.

The following lemma investigates the evolution x over a union C of k-colonies in the time interval $[t, t + sP_k]$ for $s \geq 0$. Let

$$C' = \Gamma(C, sP_k).$$

Let us denote by $W(n)$ the cube $[-n \ldots n)$. The following lemma confines several objects to a cube of the form $W(m/3)$ to make sure that triangles and colonies do not become degenerate by reaching around the torus.

LEMMA 5.1 (MAIN LEMMA) *Let C be the union of some disjoint canonical k-colonies such that C' is contained in $W(m/3)$. Let I be a set of disjoint triangles contained in $W(m/3)$. Suppose further that*

(5.4a) *The k-noise is empty in the set $[t \,..\, t + sP_k) \times C'$.*

(5.4b) *At time t, the projection of the k-anarchy of C' is covered by $D(I, P_k)$.*

Then at time $t + sP_k$, the projection of the k-anarchy of C is covered by

$$D(I, P_k + (s(1 - v_k) - 2)P_k).$$

5.3 The beginning of the proof

The case $k = 0$. Condition (5.4a) says that the 0-noise is empty, i.e. there are no failures. The projection of the 0-anarchy is enclosed into the 1-deflation of the disjoint triangles of I. An application of M means an application of Toom's rule to the 0-anarchy: it compresses the projection of the 0-anarchy into the 2-deflation of I. Similarly, each application of M results in a further compression.

In what follows we assume that the lemma is true for $k-1$, and prove it for k. For the sake of this proof, let us call a *small deflation* a deflation by the amount P_{k-1}, and a *big deflation* a deflation by the amount P_k. The k-anarchy, k-noise and k-events will be called *global* anarchy, noise and events, as opposed to the $(k-1)$-anarchy, noise and events, which will be called *local*. The notions "big", "small", "global", "local" are thus relative to the level k.

In our desire to achieve a somewhat less ridicuous bound on ρ in the Theorem, we will treat the case $k = 1$ a little differently from the other cases, since smaller constants are possible in this case, and these smaller constants at $k = 1$ become crucial in Lemma 4.1.

The local anarchy. Let us formulate an assertion that implies the Main Lemma.

LEMMA 5.2 *Under the assumptions of the Main Lemma, at time $t + sP_k$, for each k-colony E in C, there is a set $\mathcal{M}$ of triangles with $|\mathcal{M}| < c_2 r_k P_{k-1}$ such that the projection of the $(k-1)$-anarchy in E is covered by the small deflation of the system*

$$\mathcal{M} \cup D(I, P_k + s(1 - v_k)P_k).$$

PROOF OF THE MAIN LEMMA: Let E be a colony in C whose projection is not covered by $D(I, P_k + (s(1 - v_k) - 2)P_k)$. Then, as it is easy to see from (5.2), this colony has no intersection with the triangles in $D(I, P_k + s(1 - v_k)P_k)$. Application of Lemma 5.2 gives then that colony E is healthy. ∎

Lemma 5.2 can be proved using induction on s. For $s = 0$, it follows directly from the assumptions of the Main Lemma and the definition of health. The proof from s to $s + 1$ is the same for all s, so it is enough to prove the lemma for $s = 1$. Similarly, we can assume $t = 0$ without loss of generality. We will use the notation

$$\delta = P_{k-1}.$$

6 Shrinking local anarchy

We will proceed in time steps of size δ. Application of the Main Lemma for $k-1$ would eliminate the global anarchy if the set of failures did not have any local noise. The effect of the local noise will only be to slow the shrinking of the local anarchy. Indeed, in the time intervals when there is no local noise, the space-projection of earlier local noise can be added to the set I of triangles, and the lemma can be applied inductively. The local noise is not too large, since the global noise is empty. Therefore the slowdown will be small.

A refinement of this plan must consider that at the beginning, there is some local anarchy in C' outside I. We should worry that the local noise of failures might increase it, possibly turning thereby a big colony sick. We will find that this does not happen. Outside the local noise, the original local anarchy shrinks so fast that it will disappear near the beginning of the time period $[0\,..\,P_k)$. This is proven in Lemma 6.6 below. Hence all local anarchy outside the somewhat shrunken original triangles in I will be the result of the local noise of new failures: no accumulation takes place.

6.1 Clock cycles

We step through the period $[0\,..\,P_k)$ in steps of size δ. In an instant $q\delta$, the set

$$\Gamma_k(C,q) = \Gamma(C, P_k - q\delta)$$

contains the only cells at time $q\delta$ that can have any effects on the state of the cells in C at time P_k.

According to condition (5.4a) in the lemma, the global noise in the set C' during our time interval is empty. It follows that since C' is the union of 27 big colonies, the local noise on it consists of at most $27r_k$ small events.

Actually, we will step through the period $[0\,..\,P_k)$ in steps of size $\geq 3\delta$. The reason is that we want to use the fact that in the application of the Main Lemma for $k-1$, the expression

$$s(1 - v_{k-1}) - 2 \tag{6.1}$$

is positive if s is at least 3. Let us find therefore some integers

$$0 = q_0 < q_1 < \cdots < q_r = Q_k.$$

We will call these integers *boundaries*. The interval $[q_i \mathinner{..} q_{i+1})$ is called *singular* if the $(k-1)$-noise is nonempty during the time interval $[q_i\delta \mathinner{..} q_{i+1}\delta)$. Otherwise, this interval is called *regular*. We will enforce the following properties:

- Singular intervals have length 3.
- Every regular interval is surrounded by singular intervals.

Such boundaries q_i are easy to find, given that the number of singular integers is much smaller than Q_k. First we find a big interval $[a \mathinner{..} b)$ consisting of regular numbers. Let a' be the smallest number $\le a$ divisible by 3, and let b' be the largest number $\le b$ congruent to Q_k mod 3. Now we partition the interval $[0 \mathinner{..} a')$ into intervals of length 3 and keep the boundaries of those subintervals that contain singular numbers. We do the same with the interval $[b' \mathinner{..} Q_k)$. The two kinds of boundaries form the set $\{q_0, \dots, q_r\}$.

For each boundary q, we will define a colony I_q of disjoint triangles such that the following proposition holds.

LEMMA 6.1 *For all numbers q of the form q_i, at time $q\delta$, the small deflation of the system I_q covers the projection of the local anarchy in $\Gamma_k(C, q)$.*

Definition of I_q. The sets I_q will be defined inductively. Let us define I_0 first.

The assumptions of the Main Lemma together with the definition of healthy colonies imply that for each of the big colonies E_j $(i = 1, \dots, 27)$ in C', if the projection of E_j is not covered by the big deflation of I then there is a set of triangles with size less than $c_2 r_k \delta$ whose small deflation covers the projection of the local anarchy in E_j at time 0.

Let K be the union of these 27 sets of triangles. Now, we define

$$I_0 = (D(I, P_k - \delta) \cup K)'.$$

Thus, to find I_0 we first deflate I so that the projection of the global anarchy is still within a small deflation of the result. Then we add the triangles of K, containing the projection of the local anarchy in their small deflation. Finally, we make the resulting set of triangles disjoint by repeated merging.

We proceed to the definition of I_q for $q > 0$. Let us assume that q is a boundary number q_i and that I_q is defined.

Let us assume first that q_i begins a singular interval. Let $\mathcal{F}_q$ be the set of those small colonies (at most $27r_k$) that cover the space projection of the part of the local noise falling in the time interval $[q_i\delta \,..\, q_{i+1}\delta)$. Since we want to apply the Main Lemma for the case $k-1$, we have to cut out the small colonies that could be affected by $\mathcal{F}_q$. Therefore we define

$$C_q = \Gamma_k(C, q+3) \setminus \Gamma(\bigcup \mathcal{F}_q, 3\delta), \ C'_q = \Gamma(C_q, 3\delta).$$

The following lemma holds, by definition.

LEMMA 6.2 *During the time interval $[q_i\delta \,..\, q_{i+1}\delta)$ the local noise is empty over the set C'_q.*

In the definition of I_{q+1}, the set $\mathcal{F}_q$ will be taken into account via the following set of triangles:

$$\mathcal{L}_q = \Delta(\mathrm{proj}\mathcal{F}_q, 7\delta).$$

Here, the blowup by 7δ has two causes. A blowup by 6δ accounts, according to (5.2), for the spreading of the effect of the noise $\mathcal{F}_q$ during the interval $[q\delta \,..\, (q+3)\delta)$. The additional blowup by δ provides for the possibility of a later deflation by the same amount.

With these quantities, now we define $I_{q+3} = (I_q \cup \mathcal{L}_q)'$. Now assume that $[q \,..\, p)$ is a regular interval. We define

$$I_p = D(I_q, ((p-q)(1-v_{k-1}) - 2)\delta).$$

PROOF OF LEMMA 6.1. We use induction on q. It holds for $q = 0$ by the definition of I_0. Let us assume that it holds for q.

Suppose first that q is singular. By Lemma 6.2, the local noise is empty over C'_q in the period $[q\delta \,..\, (q+3)\delta)$. It follows from the application of Lemma 6.1 to q that at time $q\delta$, the projection of the local anarchy in C'_q is covered by a small deflation of I_q. Therefore we can apply the Main Lemma for $k-1$, with C_q in place of C, with I_q in place of I with $q\delta$ in place of t

and 3 in place of s. We obtain that at time $(q+1)\delta$, the projection of the local anarchy over C_q is covered by the small deflation of

$$D(\mathcal{I}_q, 3(1-v_{k-1})\delta - 2\delta).$$

Since $3(1-v_{k-1})-2$ is positive, the same set is therefore covered by the small deflation of $\mathcal{I}_q$. It follows from the definition of $\mathcal{L}_q$ and (5.2) that the projection of the difference between the sets C_q and $\Gamma_k(C, q+3)$ is covered by the small deflation of $\mathcal{L}_q$. This fact, together with the definition of $\mathcal{I}_{q+3}$, proves that the projection of the local anarchy at time $(q+3)\delta$ is covered by the small deflation of $\mathcal{I}_{q+3}$.

Suppose now that $[q\,..\,p)$ is a regular interval. We can apply the Main Lemma for $k-1$ again, with the same substitutions as above except for $s = p - q$ and $\Gamma_k(C,p)$ in place of C. We obtain the statement of the present lemma for p. ∎

6.2 Big and small triangles

The process of creating $\mathcal{I}_q$ involved three kinds of steps. Sometimes we added a new triangle. Sometimes we inflated or deflated an old triangle. And sometimes we merged two triangles into the smallest one containing both. For each element of $\mathcal{I}_q$ we can define the set of their *ancestors*. New triangles are their own ancestors. Inflating or deflating a triangle does not change its ancestors. Merging two triangles unites their ancestry. Those triangles that have some ancestry in $\mathcal{I}$ are called *big*, the rest are called *small*.

LEMMA 6.3 *Each big triangle has exactly one big ancestor.*

PROOF: In order to prove this statement, we have to show that in the process of constructing $\mathcal{I}_q$, two big triangles will never be merged. Let H and I be two such triangles. They are disjoint, so according to Lemma 3.1, there is a j such that their j- separation $d_j(H, I)$ is positive. The construction begins by deflating each big triangle by the amount $P_k - \delta$. This increases the j- separation by twice the same amount. From that time on, the big triangles suffer only merging with small triangles or deflation. Each deflation increases the separation. The merging decreases the separation at most by the size of the small triangle merged with the big one.

There are two kinds of original small triangle. The first kind are the elements of $\mathcal{K}$. These come from 27 different sets of size at most $c_2 r_k \delta$. It is easy to see that for any fixed j, the measurement m_j of a big triangle can be changed by merging with at most 18 from these 27 sets of triangles. This follows from the following simple geometrical fact.

LEMMA 6.4 *For any c, for $j = 1, 2, 3$, a strip of the form*

$$\{ u \ : \ c \le l_j(u) < c + c_2 r_k \}$$

can intersect the projection of at most 18 members of a set of 27 neighbor k-colonies.

To see this, it is enough to note that if we have an array of 9 squares of size 1 forming size forming a square of size 3 then a straight line of one of the mentioned three directions will intersect at most 6 of these squares.

The second kind of small triangles are all elements of $\mathcal{L}_q$ for some q. It is easy to see that the size of each such small triangle is at most 16δ. They come from 27 different sets with at most r_k elements.

This is also the maximum number of singular intervals. Therefore in the worst case, there will be $27 r_k$ cases when the separation decreases by the amount 16δ. Therefore the maximum total decrease is

$$(18c_2 + 16 \cdot 27) r_k \delta = (18c_2 + 432) r_k \delta.$$

It follows from (4.1) and (5.3) that we have $18c_2 + 432 < 2c_1 - 2$. Therefore the maximum total decrease is smaller than $(2c_1 - 2) r_k \delta < 2(P_k - \delta)$, and the j- separation of H and I will never be brought to 0. ∎

Shrinking big triangles. For each big triangle I in $\mathcal{I}$, and q of the form q_i, let us denote by $\Lambda_q(I)$ the big triangle in $\mathcal{I}_q$ whose ancestor it is. The following lemma estimates the total deflation suffered by a big triangle.

LEMMA 6.5 *Let I be an element of $\mathcal{I}$, and q a boundary number. Then we have*

$$\Lambda_q(I) \subset D(I, P_k + a\delta)$$

where $a = q(1 - v_{k-1}) - (18c_2 + 570) r_k$.

PROOF: Each big triangle is subjected, at the construction of I_0, to an initial deflation by P_k and inflation by δ.

Just as in the proof of Lemma 6.3 we can see that the total increase in a measurement during the singular intervals is at most $432r_k\delta$. Adding the original inflation by δ and the effect of merging with elements of $\mathcal{K}$ brings this to at most

$$(18c_2 + 433)r_k\delta. \tag{6.2}$$

To estimate the change caused by the regular intervals on $[0\,..\,q]$ let us note that there are at most $27r_k + 1$ such intervals. Their total length is at least $q - 3 \cdot 27r_k$. According to the definition of I_q, over such an interval $[q\,..\,p)$ the measurements of $\Lambda_q(I)$ decrease by the amount $((p-q)(1-v_{k-1})-2)\delta$. Adding all these decreases, we get at least

$$(q - 81r_k)(1 - v_{k-1})\delta - 2(27r_k + 1)\delta.$$

Taking into account the increase (6.2) and disregarding the factor δ, the total decrease is at least

$$\begin{aligned}(q - 81r_k)(1 - v_{k-1}) - 56r_k - (18c_2 + 433)r_k &> \\ q(1 - v_{k-1}) - (18c_2 + 570)r_k.\end{aligned}$$

∎

Vanishing small triangles. Let us denote by J_q the set of small triangles in I_q. This set is obtained by consecutive deflations and mergings between the elements of $\mathcal{K}$ and those of $\mathcal{L}_s$ for $s \le q$. The next lemma states that by the time $q = p$, the deflations eliminate the effect of $\mathcal{K}$.

LEMMA 6.6 *There is a boundary number q for which J_q is empty. Consequently, the triangles in J_{Q_k} are covered by the small deflation of the set*

$$\Big(\bigcup_s \mathcal{L}_s\Big)'.$$

The second statement follows since the first statement shows that by the time Q_k, all triangles with ancestors in $\mathcal{K}$ vanish.

PROOF: Suppose now that J_q never vanishes. Then in all regular intervals $[q\,..\,p)$, the *size* of J_q decreases by the amount $2((p-q)(1-v_k)-2)\delta$, since this is the decrease suffered by each triangle that does not disappear. The total decrease in size is thus at least (using again the fact that the number of regular intervals is at most $27r_k+1$), is the following.

$$\begin{aligned} 2(Q_k - 81r_k)(1-v_{k-1})\delta - 4(27r_k+1)\delta &> \\ r_k\delta(2c_1k^2(1-v_{k-1}) - 274) &\geq r_k\delta(0.4c_1 - 274). \end{aligned}$$

How large is the largest small triangle we can get? Using a reasoning similar to Lemma 6.4 we can see that the ancestors of any small triangle can come only from at most 12 of the 27 groups. Therefore the size of the largest small triangle is bounded by $12(c_2+16)r_k\delta = (12c_2+192)r_k\delta$. The small triangles will therefore disappear if we have

$$12c_2 + 192 \leq 0.4c_1 - 274.$$

Thus we need $12c_2 + 366 \leq 0.4c_1$. This follows from (4.1) and (5.3). ∎

PROOF OF LEMMA 5.2: Applying Lemma 6.1 for $q = Q_k$ we find that at time P_k, the system $D(I_q,\delta)$ covers the projection of the local anarchy in C. The system I_q consists of big and small triangles. According to the application of Lemma 6.5 for $q = Q_k$, for each big triangle I the triangle $\Lambda_q(I)$ is contained in $D(I, P_k + cP_k)$ where

$$c = 1 - v_{k-1} - (18c_2 + 570)/c_1k^2.$$

It follows from (4.1) and (5.3) that we have

$$18c_2 + 570 \leq 0.5c_1. \tag{6.3}$$

Therefore $\Lambda_q(I)$ is contained in $D(I, P_k(1+1-v_k))$. According to Lemma 6.6, the small triangles are covered by $\mathcal{M} = (\bigcup_s \mathcal{L}_s)'$. As estimated in the proof of Lemma 6.3, the size of $\mathcal{M}$ is at most

$$192r_k\delta = c_2r_k\delta. \tag{6.4}$$

∎

Let us note that the relation determining the definition of c_2 was (6.4). The relation determining the definition of c_1 was (6.3).

7 Proof of the theorem

Under the conditions of the Main Lemma, the k-anarchy is covered by a set of shrinking triangles. If those conditions hold long enough then therefore the k-anarchy will be eliminated. This is the statement of the next lemma.

LEMMA 7.1 *Let $k > 0$. Suppose that $P_k \leq m/3$. Suppose that the $(k-1)$-noise is empty in $W(P_k)$ during the interval $[t\,..\,t + P_k/2)$ and none of the canonical k-colonies in $W(P_k)$ belongs to the k-anarchy. Then the $(k-1)$-anarchy is empty in $W(P_k/2)$ at time $t + P_k/2$.*

PROOF: By the assumption of the present lemma there is a set $\mathcal{I}$ of triangles with size at most $8c_2 r_k P_{k-1}$ such that $D(\mathcal{I}, P_{k-1})$ covers the projection of the $(k-1)$-anarchy in $W(P_k)$. Therefore the Main Lemma is applicable with $k-1$ and $s = Q_k/2$. We obtain that the projection of the k-anarchy is covered by

$$D(\mathcal{I}, P_{k-1}(1 + Q_k(1 - v_k)/2 - 2)).$$

This set of triangles is empty. Indeed, otherwise its size, disregarding a factor P_{k-1}, is at most

$$8c_2 r_k - c_1 k^2 r_k (1 - v_k) + 2 \leq r_k(8c_2 - 0.2c_1) + 2.$$

It follows from (4.1) and (5.3) that this quantity is negative. ∎

LEMMA 7.2 *Suppose that no canonical k-event in the set $[0\,..\,T] \times W$ belongs to the k-noise. Then all k-colonies are healthy at all times in $[0\,..\,T]$.*

PROOF: Let us remember that we have not used anywhere the fact that the one-dimensional rule D used in the definition of M is homogenous in time (or space). Everything remains unchanged for a medium $D_t(x, y, z)$ dependent on time. Let $D_t(x, y, z) = D(x, y, z)$ with the old D for $t \geq 0$ and $D_t(x, y, z) = y$ for $t < 0$. We only consider ρ-perturbations $\xi[t, v]$ of the new medium M_t in which with probability 1 we have $\xi[t, v] = \xi[0, v]$ for all negative t. All earlier results remain in force for this ad hoc generalization.

By the above paragraph, the set of deviations is empty for any negative t. Therefore it is enough to prove that if all k-colonies are healthy at some time t then they are all healthy at time $t+P_k$. This follows by an application of the Main Lemma to any k-colony C with an empty set I of triangles. ∎

Let K and T be given, and let $L = \log(KT)$. The number m will be chosen to have the form $4P_l$ with

$$l = \lceil (2\log L)^{0.5} \rceil. \tag{7.1}$$

LEMMA 7.3 *For $L \geq 4$, the choice* (7.1) *satisfies*

$$16KTp_l \leq \rho, \tag{7.2}$$

and we have $m < L2^{2l(\log l + O(1))}$.

PROOF: By definition, we have $\log p_l = R(2.5 \cdot 2^{l(l+1)/2})$. Therefore (7.2) can be written as

$$2.5R2^{l(l+1)/2} \geq L + R + 4. \tag{7.3}$$

From (7.1) we have

$$2^{(l-1)^2/2} < L \leq 2^{l^2/2}, \tag{7.4}$$

hence $2^{l(l+1)/2} \geq L2^{l/2}$. Therefore the left-hand side of (7.3) is greater or equal than $2.5RL2^{l/2}$. It is therefore enough to check that this is greater than $L + R + 4$, i.e. that we have $1/L + 1/R + 4/LR \leq 2.5 \cdot 2^{l/2}$, which is obviously true. This proves (7.2).

Expanding the recursive definition of P_l and using (7.4) we have

$$\begin{aligned} m &= 4P_l = 4c_1^{2l}(l!)^2 \cdot 50 \cdot 2^{l(l+1)/2} \\ &< 2^{8+2l(\log l + \log c_1)+(l-1)^2/2+3l} < L2^{l(2\log l + 2\log c_1 + 3)+8}. \end{aligned}$$

This proves the second assertion of the lemma. ∎

PROOF OF THE THEOREM: We want to estimate the probability of having a deviation at a point (t, u) in space-time. Without loss of generality, we can assume $u = 0$.

We define

$$W_k = W(P_k),\ V_k = [0\,..\,P_k/2) \times W_k.$$

Let us define the sequence $t_0, \ldots, t_l$ of times and space-time blocks I_k as follows.

$$t_0 = t,\ t_{k+1} = t_k - P_{k+1}/2,\ I_k = (t_k, 0) + V_k.$$

If we imagine the direction of time pointing downwards then the sets I_k are 4-dimensional blocks forming a reversed tower on top of each other. The base of I_k is W_k.

Let E_0 be the property that for all k between 1 and l, the set of failures has no $(k-1)$-noise on I_k. Let E_1 be the property that no l-event belongs to the l-noise on V. By Lemma 4.1 the probability that a certain $(k-1)$-hypercube belongs to the $(k-1)$-anarchy is at most p_{k-1}. The set V_k contains $(2Q_k)^3 Q_k/2 = 4Q_k/2$ hypercubes of order $k-1$. Therefore the probability that E_0 does not hold on is at most

$$4 \sum_{k=1}^{\infty} Q_k^4 p_{k-1}.$$

Each term of this series is clearly at least 10 times smaller than the previous one. The first term is $4(c_1 r_1)^4 \rho$ which is less than $2 \cdot 10^{24} \rho$. By the same lemma, the probability that E_1 does not hold is at most $16KTp_l$ which is smaller than ρ, according to (7.2). Indeed, for each l-event, the probability that it belongs to the l-noise is estimated by p_l. Since $m = 4P_l$, the number of canonical l-events in $V = \mathbf{Z}_m^2 \times \mathbf{Z}_K \times [0\,..\,T]$ is at most $16KT$.

Let us now suppose that the properties E_0 and E_1 hold. The contrary happens only with probability less than $10^{25}\rho$. We verify that at time t_l the l-anarchy is empty in J_n. There are two cases. If t_l is negative then this follows from the assumption made in the proof of Lemma 7.2: the set of deviations is actually empty for negative t_n. Otherwise, this follows from E_1 and Lemma 7.2.

It follows from E_0 and Lemma 7.1 that the $(l-1)$-anarchy is empty in J_{l-1} at time t_{l-1}. Continuing this repeated application of E_0 and Lemma 7.1 gives that there is no deviation in $(t, 0)$.

Lemma 7.3 implies that m has the required bound on its rate of growth as a function of L. ∎

8 Conclusions

In this paper, a three-dimensional reliable cellular computer was given with an extremely simple construction. Any one-dimensional automaton consisting of K reliable components and working for T steps can be simulated in real-time, and essentially without encoding, but using repetition about $\log^2(KT)$. Let us discuss some related results and problems.

Better constants. The theoretical error-bound 10^{-25} is ridiculously small. Bringing it closer to the experimental bounds given by Bennett (0.05 for 2 dimensions, somewhat smaller for 3 dimensions) is a challenging task.

We are aware of some ways to shave off a few orders of magnitude from the constants, but we wanted to keep the exposition relatively transparent. (The case $k = 1$ needs separate treatment if reasonable constants are desired. For example, the definition of 1-anarchy must be changed to take advantage of the fact that the last term -2 can be avoided in (6.1) when $k = 1$.)

The need for lower-dimensional devices. There are some theoretical reasons to look for a reliable cellular array in 2 dimensions. It follows from general thermodynamical principles that in a "physical" error-correcting medium, each component turns free energy into heat at a constant rate. Therefore each component needs, essentially, a private line for feeding with free energy and for conducting away the heat. An extra dimension is taken up by these lines.

Difficulties in lower dimensions. In the work [3] (which is more recent than the present paper), a two-dimensional reliable cellular array is constructed. The error-correcting organization takes the form of a hierarchy of "colonies". During its work period, a colony of level k performs an additional layer of error-correction on the work of its constituent colonies of level $k - 1$. One of the main problems of an elaborate construction like this is that the organization itself is also bombarded by the failures.

It is an important technical point of both [2] and [3] that all essential structurel information about a given colony C can be expressed with the help of two variables τ, π called *phase variables.* Here τ counts the number work periods passed by C in the work period of the enclosing higher-order colony B, and π shows the position of the colony within B. The solution offered in [3] is based on the recognition that the phase-variables vary periodically therefore their maintenance is possible using Toom's Rule. Since there is no Toom Rule in 1 dimension, the maintenance of the phase variables must be achieved by other means. This accounts for the difference in complexity between the papers [2] and [3].

Redundancy. All results of the style discussed in the present paper indicate that if the size of the computation is N then the size of the simulating computation must be at least $N \log N$, *provided that some real computation is being performed.* The last qualification is necessary since information storage needs only a constant factor in redundancy. The factor $\log N$ can be considered the price we pay in redundancy for reliability. In the model of cellular arrays, it is possible to distinguish between the time and space requirements of the computation, and it is therefore possible to represent the redundancy as a product of the redundancies in space and time. Thus if t steps of computation of n cells of a deterministic medium are simulated reliably by t' steps of n' cells of a stochastic medium then t'/t is the time redundancy and n'/n is the space redundancy.

The results available to date on cellular arrays indicate that the product of the time and space redundancies must be logarithmic in the size of the computation. We can state this as a conjecture, but it seems a difficult one to prove, especially that the case of "no real computation" must be excepted.

The logarithmic redundancy can be shifted entirely to space, or almost entirely to time, as the following examples show.

- The 3-dimensional simulation of the present paper is real-time: there is no time redundancy, but there is space redundancy of size $\log^{2+2\epsilon} N$. The ideas of the present paper can be applied to [2] give a real-time 2-dimensional reliable simulation with space redundancy $\log^{1+\epsilon} N$. Indeed, one can use parallel lines (cycles) of the two-dimensional array to store one symbol each of the one-dimensional medium D to be simulated. Along each of these lines, the one-dimensional error-correcting medium works. Across the lines, we apply the rule D. The time-delay of the one-dimensional computation is irrelevant now since this computation is used only to preserve one symbol. The two-dimensional medium constructed this way is as complex as the one-dimensional medium it relies on.

- The 2-dimensional simulation described in [3] has constant space redundancy, and time redundancy $\log^{1+\epsilon} N$. The ideas of [3] combined with those in [2] givė the same parameters in 1 dimension.

There does not seem to be any reason for the exponent $1 + \epsilon$ in the theorem. If the computation is only the keeping of a bit for T steps then the construction and proof of [8] can be adapted to get $m = O(\log T)$. Some more topological insight will probably also adapt the proofs in [9] and [10].

Synchronization. Most works on the theory of interacting particle systems are concerned with the case of *continuous time*. The problem of dealing with *asynchrony* arises naturally in the contexy of reliable computation, too. There are simple ways to implement a synchronous error-free computation by an asynchronous but otherwise failure-free computation. Indeed, one can use the same topology for the asynchronous device, and a somewhat enlarged local alphabet. The state $x[t, n]$ is replaced by the triple

$$(x[t-1, n], x[t, n], t \bmod 3). \tag{8.1}$$

But no such simple construction is known to turn a reliable discrete-time cellular array into a reliable continuous-time cellular array. The hierarchical construction used for error-correction in [2] and [3] can probably also be used for synchronization. Each block of cells is supposed to be synchronized to a close tolerance and is periodically resynchronized. Higher order blocks are synchronized to progressively looser tolerances. Details must be worked out yet.

For the 3–dimensional case, Charles H. Bennett's physical analysis and simulation of the behavior of transformations like (8.1) holds some promise.

Permanent failures. If we desire to extend the present investigation to permanent failures then some other assumptions must also be changed. If namely permanent failures appear with some constant frequency then in a constant number of steps, almost all components will be out of service. There are several ways to introduce restrictions preventing this. One is to declare that all permanent failures are due to manufacturing defects, and their number does not increase. This model is almost like our original one, except that at time 0 some randomly chosen cells are permanently damaged: nothing can be assumed about their transitions. It is our conviction that the hierarchical reasoning applied in the present paper will be successfully applicable to this model.

References

[1] R.L.Dobrushin and S.I.Ortyukov: Upper bound on the redundancy of self-correcting arrangements of unreliable elements, *Problems of Information Transmission* **13** No.3 (1977), 201-208.

[2] P.Gács: Reliable Computation with Cellular Automata, *Journal of Computer and System Sciences* **32** No.1 (1986) 15-78.

[3] P.Gács: Error-correction in two dimensions: a trade-off in complexity, *Technical Report*, Boston University.

[4] G.L.Kurdyumov: An example of a nonergodic homogenous one-dimensional random medium with positive transition probabilities, *Soviet Mathematical Doklady* **19** No.1 (1978), 211-214.

[5] A.V.Kuznietsov: Information storage in a memory assembled from unreliable components, *Problems of Information Transmission* **9** No.3 (1973), 254-264.

[6] F.T.Leighton, C.E.Leiserson: Wafer-scale integration of systolic arrays (extended abstract), *Proc. of the 23-rd Annual Symposium on the Foundations of Computer Science*, Chicago 1982, 297-311

[7] M.C.Taylor: Reliable information storage in memories designed from unreliable components, *Bell System Tech. J.* **47** No.10 (1968), 2299-2337.

[8] A.L.Toom: Nonergodic multidimensional systems of automata, *Problems of Information Transm.* **10** 239-246

[9] A.L.Toom: Stable and attractive trajectories in multicomponent systems, *Advances in Probability* **6** (Multicomponent systems, ed. R.L.Dobrushin) Dekker, NY 1980, 549-575

[10] A.L.Toom: Estimates for the measures describing the behavior of stochastic systems with local interaction, (in Russian), in: *Interactive Markov Processes and their Application to the Mathematical Modelling of Biological Systems*, ed. Dobrushin, Kryukov, Toom. Acad. Sci. USSR, Pushchino 1982.

[11] B.S.Tsirel'son: Reliable information storage in a system of locally interacting unreliable elements, *in* "Interacting Markov Processes in Biology", Lecture Notes in Math., Vol. 653, Springer-Verlag, New York/Berlin, 1978.

[12] J. von Neumann: Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components. *Automata Studies* (Shannon, McCarthy, eds.) Princeton Univ.Press NJ 1956

Acknowledgement: Charles H. Bennett's experiments with Toom's media and his discussions of the related physics were a source of inspiration. Conversations with Leonid Levin were helpful. K.Ganesan's and W.Wang's careful reading decreased the number of errors.