A NEW VERSION OF TOOM'S PROOF

PETER GÁCS
BOSTON UNIVERSITY

ABSTRACT. There are several proofs now for the stability of Toom's example of a two-dimensional stable cellular automaton and its application to fault-tolerant computation. Simon and Berman simplified and strengthened Toom's original proof: the present report is simplified exposition of their proof.

1. INTRODUCTION

Let us define cellular automata.

Definition 1.1. For a finite m, let $\mathbb{Z}_m$ be the set of integers modulo m; we will also write $\mathbb{Z}_\infty = \mathbb{Z}$ for the set of integers. A set $\mathbb{C}$ will be called a *one-dimensional set of sites*, or *cells*, if it has the form $\mathbb{C} = \mathbb{Z}_m$ for a finite or infinite m. For finite m, and $x \in \mathbb{C}$, the values $x+1$ $x-1$ are always understood modulo m. Similarly, it will be called a two- or three-dimensional set of sites if it has the form $\mathbb{C} = \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2}$ or $\mathbb{C} = \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \times \mathbb{Z}_{m_3}$ for finite or infinite m_i. One- and three-dimensional sets of sites are defined similarly.

For a given set $\mathbb{C}$ of sites and a finite set $\mathbb{S}$ of states, we call every function $\xi : \mathbb{C} \to \mathbb{S}$ a *configuration*. Configuration ξ assigns state $\xi(x)$ to site x. For some interval $I \subset (0, \infty]$, a function $\eta : \mathbb{C} \times I \to \mathbb{S}$ will be called a *space-time configuration*. It assigns value $\eta(x,t)$ to cell x at time t.

In a space-time vector $\langle x, t \rangle$, we will always write the space coordinate first. ⌟

Definition 1.2. Let us be given a function function $\mathrm{Trans} : \mathbb{S}^3 \to \mathbb{S}$ and a one-dimensional set of sites $\mathbb{C}$. We say that a space-time configuration η in one dimension is a *trajectory* of the *one-dimensional (deterministic) cellular automaton* $\mathrm{CA}(\mathrm{Trans})$

$$\eta(x,t) = \mathrm{Trans}(\eta(x-B, t-T), \eta(x, t-T), \eta(x+B, t-T))$$

holds for all x, t. Deterministic cellular automata in several dimensions are defined similarly. ⌟

Since we want to analyze the effect of noise, we will be interested in random space-time configurations.

Partially supported by NSF grant CCR-9204284.

Definition 1.3. For a given set **C** of sites and time interval I, consider a probability distribution **P** over all space-time configurations $\eta : \mathbb{C} \times I \to \mathbb{S}$. Once such a distribution is given, we will talk about a *random space-time configuration* (having this distribution). We will say that the distribution **P** defines a *trajectory* of the *ε-perturbation*

$$\mathrm{CA}_{\varepsilon}(\mathrm{Trans})$$

if the following holds. For all $x \in \mathbb{C}, t \in I$, $r_{-1}, r_0, r_1 \in \mathbb{S}$, let E_0 be an event that $\eta(x + j, t - 1) = r_j$ $(j = -1, 0, 1)$ and $\eta(x', t')$ is otherwise fixed in some arbitrary way for all $t' < t$ and for all $x' \neq x$, $t' = t$. Then we have

$$\mathbf{P}[\, \eta(x,t) = \mathrm{Trans}(r_{-1}, r_0, r_1) \mid E_0 \,] \leqslant \varepsilon.$$

⌟

A simple stable two-dimensional deterministic cellular automaton given by Toom in [3] can be defined as follows.

Definition 1.4 (Toom rule)**.** First we define the neighborhood

$$H = \{\langle 0,0\rangle, \langle 0,1\rangle, \langle 1,0\rangle\}.$$

The transition function is, for each cell x, a majority vote over the three values $x + g_i$ where $g_i \in H$. ⌟

As in [2], let us be given an arbitrary one-dimensional transition function Trans and the integers N, T.

Definition 1.5. We define the three-dimensional transition function Trans′ as follows. The interaction neighborhood is $H \times \{-1, 0, 1\}$ with the neighborhood H defined above. The rule Trans′ says: in order to obtain your state at time $t + 1$, first apply majority voting among self and the northern and eastern neighbors in each plane defined by fixing the third coordinate. Then, apply rule Trans on each line obtained by fixing the first and second coordinates.

For a finite or infinite m, let $\mathbb{C}$ be our 3-dimensional space that is the product of $\mathbb{Z}_m^2$ and a 1-dimensional (finite or infinite) space **A** with $N = |\mathbf{A}|$. For a trajectory ζ of Trans on **A**, we define the trajectory ζ' of Trans′ on $\mathbb{C}$ by $\zeta'(i, j, n, t) = \zeta(n, t)$. ⌟

Let ζ' be a trajectory of Trans′ and η a trajectory of $\mathrm{CA}_{\varepsilon}(\mathrm{Trans}')$ such that $\eta(0, w) = \zeta'(0, w)$.

Theorem 1. *Suppose* $\varepsilon < \frac{1}{32 \cdot 12^8}$. *If* $m = \infty$ *then we have*

$$\mathbf{P}[\, \eta(w,t) \neq \zeta'(w,t) \,] \leqslant 24\varepsilon.$$

If m is finite then we have

$$\mathbf{P}[\, \eta(w,t) \neq \zeta'(w,t) \,] \leqslant 24tm^2N(2 \cdot (12)^2 \varepsilon^{1/12})^m + 24\varepsilon.$$

The proof we give here is a further simplification of the simplified proof of [1].

Definition 1.6. Let Noise be the set of space-time points v where η does not obey the transition rule Trans$'$. Let us define a new process ξ such that $\xi(w,t) = 0$ if $\eta(w,t) = \zeta'(w,t)$, and 1 otherwise. Let

$$\mathrm{Corr}(a,b,u,t) = \mathrm{Maj}(\xi(a,b,u,t), \xi(a+1,b,u,t), \xi(a,b+1,u,t)).$$

⌟

For all points $(a,b,u,t+1) \notin \mathrm{Noise}(\eta)$, we have

$$\xi(a,b,u,t+1) \leqslant \max(\mathrm{Corr}(a,b,u-1,t), \mathrm{Corr}(a,b,u,t), \mathrm{Corr}(a,b,u+1,t)).$$

Now, Theorem 1 can be restated as follows:

Suppose $\varepsilon < \frac{1}{32 \cdot 12^8}$. If $m = \infty$ then

$$\mathbf{P}[\,\xi(w,t) = 1\,] \leqslant 24\varepsilon.$$

If m is finite then

$$\mathbf{P}[\,\xi(w,t) = 1\,] \leqslant 24tm^2N(2 \cdot (12)^2\varepsilon^{1/12})^m + 24\varepsilon.$$

2. Proof using small explanation trees

Definition 2.1 (Covering process)**.** If $m < \infty$ let $\mathbb{C}' = \mathbb{Z}^3$ be our *covering space*, and $\mathbf{V}' = \mathbb{C}' \times \mathbb{Z}$ our covering space-time. There is a projection $\mathrm{proj}(u)$ from $\mathbb{C}'$ to $\mathbb{C}$ defined by

$$\mathrm{proj}(u)_i = u_i \bmod m \qquad (i = 1,2).$$

This rule can be extended to $\mathbb{C}'$ identically. We define a random process ξ' over $\mathbb{C}'$ by

$$\xi'(w,t) = \xi(\mathrm{proj}(w), t).$$

The set Noise is extended similarly to Noise$'$. Now, if $\mathrm{proj}(w_1) = \mathrm{proj}(w_2)$ then $\xi'(w_1,t) = \xi'(w_2,t)$ and therefore the failures at time t in w_1 and w_2 are not independent. ⌟

Definition 2.2 (Arrows, forks)**.** In figures, we generally draw space-time with the time direction going down. Therefore, for two neighbor points u, u' of the space $\mathbb{Z}$ and integers a, b, t, we will call *arrows*, or *vertical edges* the following kinds of (undirected) edges:

$$\{\langle a,b,u,t\rangle, \langle a,b,u',t-1\rangle\}, \{\langle a,b,u,t\rangle, \langle a+1,b,u',t-1\rangle\},$$
$$\{\langle a,b,u,t\rangle, \langle a,b+1,u',t-1\rangle\}.$$

We will call *forks*, or *horizontal edges* the following kinds of edges:

$$\{\langle a,b,u,t\rangle, \langle a+1,b,u,t\rangle\}, \{\langle a,b,u,t\rangle, \langle a,b+1,u,t\rangle\},$$
$$\{\langle a+1,b,u,t\rangle, \langle a,b+1,u,t\rangle\}.$$

We define the graph $\mathbf{G}$ by introducing all possible arrows and forks. Thus, a point is adjacent to 6 possible forks and 6 possible arrows: the degree of $\mathbf{G}$ is at most

$$r = 12.$$

(If the space is $d+2$-dimensional instead of 3, then $r = 6(d+1)$.) We use the notation $\mathrm{Time}(\langle w, t\rangle) = t$. ⌟

The following lemma is key to the proof, since it will allow us to estimate the probability of each deviation from the correct space-time configuration. It assigns to each deviation a certain tree called its “explanation”. Larger explanations contain more noise and have a correspondingly smaller probability. For some constants c_1, c_2, there will be $\leqslant 2^{c_1 L}$ explanations of size L and each such explanation will have probability upper bound $\varepsilon^{c_2 L}$.

Lemma 2.3 (Explanation Tree)**.** *Let u be a point outside the set Noise′ with $\xi'(u) = 1$. Then there is a tree $\mathrm{Expl}(u, \xi')$ consisting of u and points v of $\mathbf{G}$ with $\mathrm{Time}(v) < \mathrm{Time}(u)$ and connected with arrows and forks called an explanation of u. It has the property that if n nodes of* Expl *belong to* Noise′ *then the number of edges of* Expl *is at most $4(n-1)$.*

This lemma will be proved in the next section. To use it in the proof of the main theorem, we need some easy lemmas.

Definition 2.4. A *weighted tree* is a tree whose nodes have weights 0 or 1, with the root having weight 0. The *redundancy* of such a tree is the ratio of its number of edges to its weight. The set of nodes of weight 1 of a tree T will be denoted by $F(T)$.

A subtree of a tree is a subgraph that is a tree. ⌟

Lemma 2.5. *Let T be a weighted tree of total weight $w > 3$ and redundancy λ. It has a subtree of total weight w_1 with $w/3 < w_1 \leqslant 2w/3$, and redundancy $\leqslant \lambda$.*

Proof. Let us order T from the root r down. Let T_1 be a minimal subtree below r with weigth $> w/3$. Then the subtrees immediately below T_1 all weigh $\leqslant w/3$. Let us delete as many of these as possible while keeping T_1 weigh $> w/3$. At this point, the weight w_1 of T_1 is $> w/3$ but $\leqslant 2w/3$ since we could subtract a number $\leqslant w/3$ from it so that w_1 would become $\leqslant w/3$ (note that since $w > 3$) the tree T_1 is not a single node.

Now T has been separated by a node into T_1 and T_2, with weights $w_1, w_2 > w/3$. Since the root of a tree has weight 0, by definition the possible weight of the root of T_1 stays in T_2 and we have $w_1 + w_2 = w$. The redundancy of T is then a weighted average of the redundancies of T_1 and T_2, and we can choose the one of the two with the smaller redundancy: its redundancy is smaller than that of T. □

Theorem 2 (Tree Separator)**.** *Let T be a weighted tree with weight w and redundancy λ, and let $k < w$. Then T has a subtree with weight w' such that $k/3 < w' \leqslant k$ and redundancy $\leqslant \lambda$.*

Proof. Let us perform the operation of Lemma 2.5 repeatedly, until we get weight $\leqslant k$. Then the weight w' of the resulting tree is $> k/3$. □

Lemma 2.6 (Tree Counting)**.** *In a graph of maximum node degree r the number of weighted subtrees rooted at a given node and having k edges is at most $2r \cdot (2r^2)^k$.*

Proof. Let us number the nodes of the graph arbitrarily. Each tree of k edges can now be traversed in a breadth-first manner. At each non-root node of the tree of degree i from which we continue, we make a choice out of r for i and then a choice out of $r-1$ for each of the $i-1$ outgoing edges. This is r^i possibilities at most. At the root, the number of outgoing edges is equal to i, so this is r^{i+1}. The total number of possibilities is then at most r^{2k+1} since the sum of the degrees is $2k$. Each point of the tree can have weight 0 or 1, which multiplies the expression by 2^{k+1}. □

Proof of Theorem 1. Let us consider each explanation tree a weighted tree in which the weight is 1 in a node exactly if the node is in Noise′. For each n, let $\mathcal{E}_n$ be the set of possible explanation trees Expl for u with weight $|F(\mathrm{Expl})| = n$. First we prove the theorem for $m = \infty$, that is Noise′ = Noise. If we fix an explanation tree Expl then all the events $w \in \mathrm{Noise}'$ for all $w \in F = F(\mathrm{Expl})$ are independent from each other. It follows that the probability of the event $F \subset \mathrm{Noise}'$ is at most ε^n. Therefore we have

$$\mathbf{P}[\,\xi(u) = 1\,] \leqslant \sum_{n=1}^{\infty} |\mathcal{E}_n| \varepsilon^n.$$

By the Explanation Tree Lemma, each tree in $\mathcal{E}_n$ has at most $k = 4(n-1)$ edges. By the Tree Counting Lemma, we have

$$|\mathcal{E}_n| \leqslant 2r \cdot (2r^2)^{4(n-1)},$$

Hence

$$\mathbf{P}[\,\xi(u) = 1\,] \leqslant \frac{2r}{\varepsilon} \sum_{n=0}^{\infty} (16r^{16}\varepsilon)^n = \frac{2r}{\varepsilon}(1 - 16r^{16}\varepsilon).$$

In the case $\mathbb{C} \neq \mathbb{C}'$ this estimate bounds only the probability of $\xi'(u) = 1$, $|\mathrm{Expl}(u, \xi')| \leqslant m$, since otherwise the events $w \in \mathrm{Noise}'$ are not necessarily independent for $w \in F$. Let us estimate the probability that an explanation $\mathrm{Expl}(u, \xi')$ has m or more nodes. It follows from the Tree Separator Theorem that Expl has a subtree T with weight n' where $m/12 \leqslant n' \leqslant m/4$, and at most m nodes. Since T is connected, no two of its nodes can have the same projection. Therefore for a fixed tree of this kind, for each node of weight 1 the events that they belong to Noise′ are independent. Hence for each tree T of these sizes, the probability that T is such a subtree of Expl is at most $\varepsilon^{m/12}$. To get the probability that there is such a subtree we multiply by the number of such subtrees. An upper bound on the number of places for the root is tm^2N. An upper bound on the number of trees from a given root is obtained from the Tree Counting Lemma. Hence

$$\mathbf{P}[\,|\mathrm{Expl}(u, \xi')| > m\,] \leqslant 2rtm^2N \cdot (2r^2\varepsilon^{1/12})^m.$$

□

3. THE EXISTENCE OF SMALL EXPLANATION TREES

3.1. Some geometrical facts.

Let us introduce some geometrical concepts.

Definition 3.1. Three linear functionals are defined as follows for $v = \langle x, y, z, t \rangle$.

$$L_1(v) = -x, \qquad L_2(v) = -y, \qquad L_3(v) = x + y.$$

⌟

Notice $L_1(v) + L_2(v) + L_3(v) = 0$.

Definition 3.2. For a set S, we write

$$\mathrm{Size}(S) = \sum_{i=1}^{3} \max_{v \in S} L_i(v).$$

⌟

Notice that for a point v we have $\mathrm{Size}(\{v\}) = 0$.

Definition 3.3. A set $\mathcal{S} = \{S_1, \ldots, S_n\}$ of sets is *connected by intersection* if the graph $G(\mathcal{S})$ is connected which we obtain by introducing an edge between S_i and S_j whenever $S_i \cap S_j \neq \emptyset$. ⌟

Definition 3.4. A *spanned set* is an object $\mathbb{P} = \langle P, v_1, v_2, v_3 \rangle$ where P is a space-time set and $v_i \in P$. The points v_i are the *poles* of $\mathbb{P}$, and P is its *base set*. We define $\mathrm{Span}(\mathbb{P})$ as $\sum_{i=1}^{3} L_i(v_i)$. ⌟

Lemma 3.5 (Spanned Set Creation). *If P is a set then there is a spanned set $\langle P, v_1, v_2, v_3 \rangle$ on P with* $\mathrm{Span}(\mathbb{P}) = \mathrm{Size}(P)$.

Proof. Assign v_i to a point of the set P in which L_i is maximal. □

The following lemma is our main tool.

Lemma 3.6 (Spanning). *Let $\mathbb{L} = \langle L, u_1, u_2, u_3 \rangle$ be a spanned set and $\mathcal{M}$ be a set of subsets of L connected by intersection, whose union covers the poles of $\mathbb{L}$. Then there is a set $\{\mathbb{M}_1, \ldots, \mathbb{M}_n\}$ of spanned sets whose base sets M_i are elements of $\mathcal{M}$, such that the following holds. Let M'_i be the set of poles of $\mathbb{M}_i$.*

(a) $\mathrm{Span}(\mathbb{L}) = \sum_i \mathrm{Span}(\mathbb{M}_i)$.
(b) *The union of the sets M'_j covers the set of poles of $\mathbb{L}$.*
(c) *The system $\{M'_1, \ldots, M'_n\}$ is a minimal system connected by intersection (that is none of them can be deleted) that connects the poles of $\mathbb{L}$.*

Proof. Let $M_{i_j} \in \mathcal{M}$ be a set containing the point u_j. Let us choose u_j as the j-th pole of M_{i_j}. Now leave only those sets of $\mathcal{M}$ that are needed for a minimum spanning tree $\mathcal{T}$ of the graph $G(\mathcal{M})$ connecting $M_{i_1}, M_{i_2}, M_{i_3}$. Keep deleting points from each set (except u_j from M_{i_j}) until every remaining point is necessary for a connection among u_j. There will only be two-and three-element sets, and any two of them intersect in at most one element. Let us draw an edge between each pair of points if they belong to a common set M'_i. This turns the union

$$V = \bigcup_i M'_i$$

into a graph. (Actually, this graph can have only two simple forms: a point connected via disjoint paths to the poles u_i or a triangle connected via disjoint paths to these poles.) For each i and j, there is a shortest path between M'_i and u_j. The point of M'_i where this path leaves M'_i will be made the j-th pole u_{ij} of M_i. For $j \in \{1,2,3\}$ we have $u_{i_j j} = u_j$ by definition. This rule creates three poles in each M_i and each point of M'_i is a pole.

Let us show $\sum_i \mathrm{Span}(\mathbb{M}_i) = \mathrm{Span}(\mathbb{L})$. We can write

$$\sum_i \mathrm{Span}(\mathbb{M}_i) = \sum_{v \in V} \sum_{i,j : v = u_{ij}} L_j(v). \tag{1}$$

For a point $v \in V$, let

$$I(v) = \{\, i : v \in M'_i \,\}.$$

For $i \in I(v)$ let $E_i(v)$ be the set of those $j \in \{1,2,3\}$ for which either $i = i_j$ or $v \neq u_{ij}$. Because graph $\mathcal{T}$ is a tree, for each fixed v the sets $E_i(v)$ are disjoint. Because of connectedness, they form a partition of the set $\{1,2,3\}$. Let $e_i(j,v) = 1$ if $j \in E_i(v)$ and 0 otherwise, then we have $\sum_i e_i(j,v) = 1$ for each j.

We can now rewrite the sum (1) as

$$\sum_{j=1}^{3} \sum_{v \in V} L_j(v)(e_{i_j}(j,v) + \sum_{v \in V} \sum_{i \in I(v) \setminus \{i_j\}} (1 - e_i(j,v))).$$

If $i = i_j \in I(v)$ then by definition we have $1 - e_i(j,v) = 0$, therefore we can simplify the sum as

$$\sum_{j=1}^{3} \sum_{v \in V} L_j(v) e_{i_j}(j,v) + \sum_{i \in I(v)} \sum_{j=1}^{3} L_j(v)(1 - e_i(j,v)).$$

The first term is equal to $\mathrm{Span}(\mathbb{L})$; we show that the last term is 0. Moreover, we show $0 = \sum_{j=1}^{3} L_j(v) \sum_{i \in I(v)} (1 - e_i(j,v))$ for each v. Indeed, $\sum_{i \in I(v)} (1 - e_i(j,v))$ is independent of j since it is $|I(v)| - \sum_i e_i(j,v) = |I(v)| - 1$. On the other hand, $\sum_j^3 L_j(v) = 0$ as always. □

3.2. **Building an explanation tree.** Let us define the excuse of a space-time point.

Definition 3.7. Let $v = \langle a,b,u,t+1 \rangle$ with $\xi'(v) = 1$. If $v \notin \mathrm{Noise}'$ then there is a u' such that $\xi'(w) = 1$ for at least two members w of the set

$$\{\, \langle a,b,u',t \rangle, \langle a+1,b,u',t \rangle, \langle a,b+1,u',t \rangle \,\}.$$

We define the set $\mathrm{Excuse}(v)$ as such a pair of elements w, and as the empty set in all other cases. By Lemma 3.5, we can turn $\mathrm{Excuse}(v)$ into a spanned set, $\langle \mathrm{Excuse}(v), w_1, w_2, w_3 \rangle$ with span 1. Denote

$$\mathrm{Excuse}_i(v) = w_i.$$

⌟

The following lemma utilizes the fact that Toom's rule "makes triangles shrink".

Lemma 3.8 (Excuse size)**.** *If $\mathbb{V} = \langle V, v_1, v_2, v_3 \rangle$ is a spanned set and v_i are not in* Noise′ *then we have*

$$\sum_{j=1}^{3} L_j(\mathrm{Excuse}_j(v_j)) = \mathrm{Span}(\mathbb{V}) + 1.$$

Proof. Let T be the triangle

$$\{\, u : L_1(u) \leqslant 0,\ L_2(u) \leqslant 0,\ L_3(u) \leqslant 1 \,\}.$$

We have $\mathrm{Size}(T) = 1$, and $\mathrm{Excuse}(v) \subset v + T$. Since the chosen poles turn $\mathrm{Excuse}(v)$ into a spanned set of size 1, the function L_j achieves its maximum in $v + T$ on $\mathrm{Excuse}_j(v)$. We have

$$L_j(\mathrm{Excuse}_j(v)) = \max_{u \in v+T} L_j(u) = L_j(v) + \max_{u \in T} L_j(u).$$

Hence we have

$$\begin{aligned}\sum_j L_j(\mathrm{Excuse}_j(v_j)) &= \sum_j \max_{u \in T} L_j(u) + \sum_j L_j(v_j)\\ &= \mathrm{Size}(T) + \mathrm{Span}(\mathbb{V}) = 1 + \mathrm{Span}(\mathbb{V}).\end{aligned}$$

□

Definition 3.9 (Clusters)**.** In a subgraph of the graph $\mathbf{G}$, let us call two nodes u, v of the above graph with $\mathrm{Time}(u) = \mathrm{Time}(v) = t$ *equivalent* if there is a path between them made of arrows, using only points x with $\mathrm{Time}(x) \leqslant t$. An equivalence class will be called a *cluster*. For a cluster K we will denote by $\mathrm{Time}(K)$ the time of its points. We will say that a fork or arrow *connects* two clusters if it connects some of their nodes. ⌟

If a cluster contains a point in Noise′ then clearly it contains no other points.

Definition 3.10 (Cause graph)**.** For a cluster K we define the *cause graph* $G_K = \langle V_K, E_K \rangle$ as follows. The elements of G_K are those clusters R with $\mathrm{Time}(R) = \mathrm{Time}(K) - 1$ which are reachable by an arrow from K. For $R, S \in V_K$ we have $\{R, S\} \in E_K$ iff for some $v \in R$ and $w \in S$ we have $\mathrm{Time}(v) = \mathrm{Time}(w) = \mathrm{Time}(K) - 1$ and $\{v, w\} \in \mathrm{Forks}$. ⌟

Lemma 3.11. *The cause graph G_K is connected.*

Proof. The points of K are connected via arrows using points x with $\mathrm{Time}(x) \leqslant \mathrm{Time}(K)$. The clusters in G_K are therefore connected with each other only through pairs of arrows going trough K. The tails of each such pair of arrows in $\mathrm{Time}(K) - 1$ are connected by a fork. □

Definition 3.12. A *spanned cluster* is a spanned set that is a cluster. ⌟

The explanation tree will be built from an intermediate object defined below. Let us fix a point u_0: from now on we will work in the subgraph of the graph $\mathbf{G}$ reachable from u_0 by arrows pointing backward in time. Clusters are defined in this graph.

Definition 3.13. A *partial explanation tree* is an object of the form $\langle C_0, C_1, E\rangle$. Elements of C_0 are spanned clusters called *unprocessed nodes*, elements of C_1 are *processed nodes*, these are nodes of $\mathbf{G}$. The set E is a set of arrows or forks $\{u, v\}$ between nodes and poles of the spanned clusters. From this structure a graph is formed if we identify each pole of a spanned cluster $\mathbb{K}$ with $\mathbb{K}$ itself. This graph is required to be a tree.

The *span* of such a tree will be the sum of the spans of its unprocessed clusters and the number of its forks. ⌟

The explanation tree will be built by applying repeatedly a "refinement" operation to partial explanation trees.

Definition 3.14 (Refinement)**.** Let T be a partial explanation tree, and let the spanned cluster $\mathbb{K} = \langle K, v_1, v_2, v_3\rangle$ be one of its unprocessed nodes, with v_i not in Noise'. We apply an operation whose result will be a new tree T'.

Consider the cause graph $G_K = \langle V_K, E_K\rangle$ defined above. Let $\mathcal{M} = V_K \cup E_K$, that is the family of all clusters in V_K (sets of points) and all edges in G_K connecting them, (two-element sets). Let L be the union of these sets, and $\mathbb{L} = \langle L, u_1, u_2, u_3\rangle$ a spanned set where $u_i = \mathrm{Excuse}_i(v_i)$. Lemma 3.11 implies that the set $\mathcal{M}$ is connected by intersection. Applying the Spanning Lemma 3.6 to $\mathbb{L}$ and $\mathcal{M}$, we find a family $\mathbb{M}_1, \dots, \mathbb{M}_n$ of spanned sets with

$$\sum_i \mathrm{Span}(\mathbb{M}_i) = \mathrm{Span}(\mathbb{L}) = \sum_i L_i(u_i).$$

It follows from Lemma 3.8 that the latter sum is $\mathrm{Span}(\mathbb{K}) + 1$, and that u_i are among the poles of these sets. Some of these sets are spanned clusters, others are forks connecting them, adjacent to their poles. Consider these forks again as edges and the spanned clusters as nodes. By the minimality property of Lemma 3.6, they form a tree $U(\mathbb{K})$ that connect the three poles of $\mathbb{K}$.

The refinement operation takes an unprocessed node $\mathbb{K} = \langle K, v_1, v_2, v_3\rangle$ in the tree T. This node is connected to other parts of the tree by some of its poles v_j.

The operation deletes cluster K, and keeps those poles v_j that were needed to keep connect $\mathbb{K}$ to other clusters and nodes in T. It turns these into processed nodes, and adds the tree $U(\mathbb{K})$ just built, declaring each of

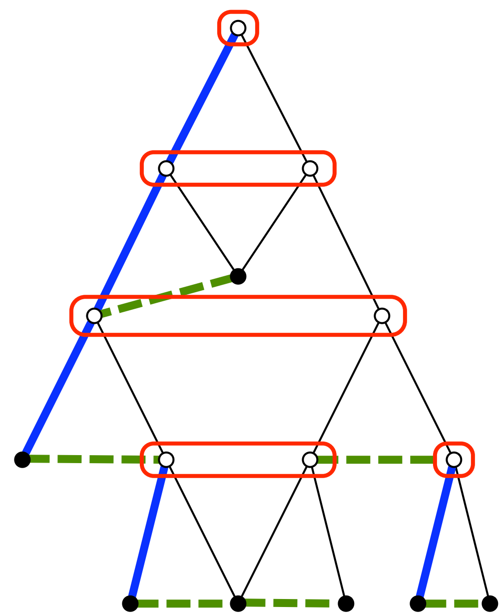

FIGURE 1. An explanation tree. The black points are noise. The squares are other points of the explanation tree. Thin lines are arrows not in the explanation tree. Framed sets are clusters to which the refinement operation was applied. Thick solid lines are arrows, thick broken lines are forks of the explanation tree.

its spanned clusters unprocessed nodes. Then it adds the arrow from these v_j to $\mathrm{Excuse}_j(v_j)$. Even if none of these nodes were needed for connection, it keeps v_1 and adds the arrow from v_1 to $\mathrm{Excuse}_1(v_1)$. ⌟

The refinement operation increases both the span and the number of arrows by 1.

Let us build now the explanation tree. We start with a node $u_0 \notin \mathrm{Noise}'$ with $\xi'(u_0) = 1$ and from now on work in the subgraph of the graph $\mathbf{G}$ of points reachable from u_0 by arrows backward in time. Then $\langle \{u_0\}, u_0, u_0, u_0 \rangle$ is a spanned cluster, forming a one-node partial explanation tree if we declare it an unprocessed node. We apply the refinement operation to this partial explanation tree, as long as we can. When it cannot be applied any longer then all nodes are either processed or one-point spanned clusters belonging to Noise'. See an example in Figure 1.

Proof of Lemma 2.3. What is left to prove is the estimate on the number of edges. Let us contract each arrow $\langle u, v \rangle$ of the explanation tree one-by-one into its bottom point v. The edges of the resulting tree are the forks. All the processed nodes will be contracted into the remaining one-node clusters that are elements of Noise$'$. If n is the number of these nodes then there are $n-1$ forks in this remaining tree. The span of the explanation tree just constructed is the sum of sizes of the forks, that is $n-1$.

The number of arrows in the tree is at most $3(n-1)$. Indeed, each introduction of at most 3 arrows by the refinement operation was accompanied by an increase of the span by 1. The total number of edges of the explanation tree is thus at most $4(n-1)$. □

References

[1] Piotr Berman and Janos Simon, *Investigations of fault-tolerant networks of computers*, Proc. of the 20-th Annual ACM Symp. on the Theory of Computing, 1988, pp. 66–77. 1

[2] Peter Gács and John Reif, *A simple three-dimensional real-time reliable cellular array*, Journal of Computer and System Sciences **36** (1988), no. 2, 125–147. 1

[3] Andrei L. Toom, *Stable and attractive trajectories in multicomponent systems*, Multicomponent Systems (New York) (R. L. Dobrushin, ed.), Advances in Probability, vol. 6, Dekker, New York, 1980, Translation from Russian, pp. 549–575. 1