



HAL
open science

The self-assembly of paths and squares at temperature 1

Pierre-Etienne Meunier

► **To cite this version:**

Pierre-Etienne Meunier. The self-assembly of paths and squares at temperature 1. 2013. hal-00912937v1

HAL Id: hal-00912937

<https://hal.science/hal-00912937v1>

Preprint submitted on 2 Dec 2013 (v1), last revised 9 Dec 2013 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The self-assembly of paths and squares at temperature 1

Pierre-Étienne Meunier *

Abstract

We prove that the number of tile types required to build squares of size $n \times n$, in Winfree’s abstract Tile Assembly Model, when restricted to using only non-cooperative tile bindings, is at least $2n - 1$, which is also the best known upper bound. Non-cooperative self-assembly, also known as “temperature 1”, is where tiles bind to each other if they match on one or more sides, whereas in cooperative binding, some tiles can bind only if they match on multiple sides.

Our proof introduces a new programming technique for temperature 1, that disproves the very intuitive and commonly held belief that, in the same model, assembling *paths* between two points A and B cannot be done with less tile types than the Manhattan distance $\|\vec{AB}\|_1$ between them. Then, we prove a necessary condition for these “efficient paths” to be assembled, and show that this necessary condition cannot hold in completely filled squares.

This result proves the oldest conjecture in algorithmic self-assembly, published by Rothemund and Winfree in STOC 2000, in the case where growth starts from a corner of the square. As a corollary, we establish n as a lower bound on the tile complexity of the general case. The problem of determining the minimal number of tile types to self-assemble a shape is known to be Σ_2^P -complete.

1 Introduction

Self-assembly is the process through which unorganized, simple, components automatically coalesce according to simple local rules to form some kind of target structure. It sounds simple, but the end result can be extraordinary. For example, researchers have been able to self-assemble a wide variety of structures experimentally at the nanoscale, such as regular arrays [28], fractal structures [10,20], smiling faces [18,26], DNA tweezers [30], logic circuits [16,21], neural networks [17], and molecular robots [12]. These examples are fundamental because they demonstrate that self-assembly can, in principle, be used to manufacture specialized geometrical, mechanical and computational objects at the nanoscale. Potential future applications of nanoscale self-assembly include the production of smaller, more efficient microprocessors and medical technologies that are capable of diagnosing and even treating disease at the cellular level.

Controlling nanoscale self-assembly for the purposes of manufacturing atomically precise components will require a bottom-up, hands-off strategy. In other words, the self-assembling units themselves will have to be “programmed” to direct themselves to assemble efficiently and correctly. Molecular self-assembly is rapidly becoming a ubiquitous engineering paradigm, and we need to develop a theory to inform us of its algorithmic capabilities and ultimate limitations.

In 1998, Erik Winfree [27] introduced the abstract Tile Assembly Model (aTAM), a simplified discrete mathematical model of algorithmic DNA nanoscale self-assembly pioneered by Seeman [22]. The aTAM is an asynchronous nondeterministic cellular automaton that models crystal growth

*Computer Science, California Institute of Technology. pmeunier@caltech.edu. Supported in part by National Science Foundation Grant CCF-1219274.

processes. Put another way, the aTAM essentially augments classical Wang tiling [25] with a mechanism for sequential growth of a tiling. This contrasts with Wang tiling in which only the existence of a valid mismatch-free tiling is considered, and *not* the order of tile placement. In the aTAM, the fundamental components are translatable but un-rotatable square or cube *tiles* whose sides are labeled with colored *glues* colors, each with an integer *strength*. Two tiles that are placed next to each other *interact* if the glue colors on their abutting sides match, and they *bind* if the strengths on their abutting sides match and sum to at least a certain (integer) *temperature*. Self-assembly starts from a *seed* tile type and proceeds nondeterministically and asynchronously as tiles bind to the seed-containing-assembly. Despite its deliberate simplification, the aTAM is a computationally expressive model. For example, by using cooperative binding (that is, by having some of the tiles bind on two or more sides), Winfree [27] proved that it is Turing universal, which implies that self-assembly can be directed by a computer program. Here, we study noncooperative binding.

Tile self-assembly in which tiles can be placed only in a *noncooperative* fashion is colloquially referred to as “temperature-1 self-assembly”. Despite the esoteric name, this is a fundamental and ubiquitous form of growth: it refers to growth from *growing and branching tips* where each new tile is added if it can match on at least *one side*.

It has been known for some time that a more general form of growth where some of the tiles must match on two or more sides, i.e. *cooperative* growth, leads to highly non-trivial behavior: arbitrary Turing machine simulation [11, 19], efficient production of $n \times n$ squares and other simple shapes using $\Theta(\log n / \log \log n)$ tile types [1], efficient production of arbitrary finite connected shapes using a number of tile types that is within a log factor of the Kolmogorov complexity of the shape [24], and even intrinsic universality: the existence of a single tile set that simulates arbitrary tile assembly systems [7].

However, the capabilities of two-dimensional noncooperative self-assembly remain largely unknown: several generalizations and restrictions of this model have been studied, that conjectured in all cases that noncooperative binding could not be as powerful as cooperative binding [4, 5, 9, 13, 15, 19]. The first fully general separation results were only proven recently [14], in the context of intrinsic universality [6–8]. However, the computational capabilities, in the Turing sense, of this model, remain largely unknown.

The conjecture that we prove in this paper was first stated by Rothmund and Winfree [19]¹: the minimal number of tile types to assemble $n \times n$ squares is $2n - 1$. A restriction of this result, where it is required that all the tiles be assembled with all their neighbors, appeared in the same paper. Moreover, computing the minimal number of tile types required to *deterministically* assemble a shape from a seed of size one is known to be NP-complete [2], and Σ_2^p -complete in the non-deterministic case [3].

1.1 Main results

Although a number of terms have not been formally defined, we give an overview of our two main results now. See the definitions in section 2. Our ultimate goal is to prove Rothmund and Winfree’s conjecture [19] that a tile assembly system assembling only squares of size $n \times n$, from a seed of size 1, has at least $2n - 1$ tile types. Our first result disproves a statement stronger than this conjecture; namely, that the tile complexity of a square is the same as the tile complexity of its diagonal. Although widely believed, this statement is false:

Theorem 3.1. Let n be an integer. There is a tile assembly system $\mathcal{T}_n = (T_n, \sigma_n, 1)$, and two

¹In Figure 2 of that paper

points $A, B \in \mathbb{Z}^2$, such that $\|\overrightarrow{AB}\|_1 = n$, the terminal assemblies of \mathcal{T}_n are all finite, they all include a path from A to B , and $|T_n| = 4n/5 + O(1)$.

The fact that the constructions of this theorem are possible, even though they are quite elementary, is not obvious at all; indeed, the intuition from words and automata theory is that any attempt to “reuse” tile types will enable us to “pump” the path, as in the pumping lemma of finite automata [23], and thus any tile assembly system that can produce paths repeating tile types will also be able to produce ultimately periodic, infinite paths. This intuition is valid in a restricted setting where two adjacent tiles always agree on their abutting sides [9, 13].

Then, we will prove the following theorem, which gives the optimal lower bound on the tile complexity of squares, when growth starts from a corner:

Theorem 4.2. Let $\mathcal{T} = (T, \sigma, 1)$ be a temperature 1 tile assembly system, with σ a single tile at $(0, 0)$, and n an integer. If all terminal assemblies producible by \mathcal{T} are of domain $\{0, \dots, n-1\}^2$, then $|T| \geq 2n - 1$.

2 Definition and preliminaries

We begin by defining the two-dimensional abstract tile assembly model. A *tile type* is a unit square with four sides, with the top side called N, the bottom side S, the right side E, and the left side W, each consisting of a glue *label* and a nonnegative integer *strength*. We assume a finite set T of tile types, but an infinite supply of copies of each type. An *assembly* is a positioning of the tiles on the discrete plane \mathbb{Z}^2 , that is, a partial function $\alpha : \mathbb{Z}^2 \dashrightarrow T$.

We say that two tiles in an assembly *interact*, or are *stably attached*, if the glue labels on their abutting side are equal, and have positive strength. An assembly α induces a weighted *binding graph* $G = (V, E)$, where $V = \text{dom}(\alpha)$, and there is an edge $(a, b) \in E$ if and only if a and b interact, and this edge is weighted by the glue strength of that interaction. The assembly is said to be τ -stable if any cut of G has weight at least τ .

A *tile assembly system* is a triple $\mathcal{T} = (T, \sigma, \tau)$, where T is a finite tile set, σ is called the *seed*, and τ is the *temperature*. Throughout this paper, we will always have $|\text{dom}(\sigma)| = 1$ and $\tau = 1$. Therefore, we can make the simplifying assumption that all glues have strength one without changing the behavior of the model.

Given two assemblies α and β , we write $\alpha \rightarrow_1^{\mathcal{T}} \beta$ if α and β are both τ -stable, $\text{dom}(\alpha) \subseteq \text{dom}(\beta)$, and $|\text{dom}(\beta) \setminus \text{dom}(\alpha)| = 1$. We say that γ is *producible* from α , and write $\alpha \rightarrow^{\mathcal{T}} \gamma$ if there is a (possibly empty) sequence $\alpha = \alpha_1, \dots, \alpha_n = \beta$ such that $\alpha_1 \rightarrow_1^{\mathcal{T}} \dots \rightarrow_1^{\mathcal{T}} \alpha_n$.

The *productions* of a tile assembly system $\mathcal{T} = (T, \sigma, \tau)$, written $\mathcal{A}[\mathcal{T}]$, is the set of all assemblies producible from σ . An assembly α is called *terminal* if there is no β such that $\alpha \rightarrow_1^{\mathcal{T}} \beta$. The set of terminal assemblies is written $\mathcal{A}_{\square}[\mathcal{T}]$.

An important fact about temperature 1 tile assembly, that we will use heavily, is that any path of the binding graph can grow immediately from the seed, independently from anything else. For any path $P = P_1, \dots, P_n$ on the lattice grid of \mathbb{Z}^2 , and any integer $1 \leq i \leq n$, we say that the *output side* of P_i is the side adjacent to P_{i+1} , if $i < n$, and its *input side* is the side adjacent to P_{i-1} , if $i > 1$. The first tile of a path does not have an input side, and the last one does not have an output side.

Finally, the *Manhattan distance* between two points $A = (x_A, y_A)$ and $B = (x_B, y_B)$ of \mathbb{Z}^2 , written $\|\overrightarrow{AB}\|_1$, is $|x_A - x_B| + |y_A - y_B|$. Moreover, for any side name $s \in \{N, S, E, W\}$, we write \bar{s} its opposite side (e.g. $\bar{N} = S$).

2.1 The known upper bound

The only known way to assemble squares of size $n \times n$ at temperature 1 with $2n - 1$ tile types, is by using the “comb” design of Figure 1, already described in Rothemund and Winfree’s paper [19].

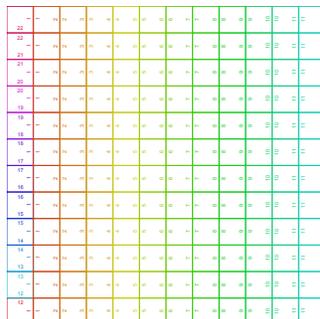


Figure 1: The comb design for size 10×10

3 Building efficient paths

A major obstacle in proving the claimed lower bound for squares, is that building only a square’s diagonal can require way less tiles than building the whole square. In this section, we show the following result:

Theorem 3.1. *There is a tile assembly system $\mathcal{T} = (T, \sigma, 1)$, such that for all terminal assembly $a \in \mathcal{A}_{\square}[\mathcal{T}]$, a is finite and of width $\frac{5(|T|+2)}{4} - 23$.*

Proof. Since a path of height n that is monotonic in the y-dimension, and has less than n tile types with input side S, can be “pumped”, our path will need to have “caves”, or non-monotonic subpaths, and reuse them several times. Moreover, in order for all the assemblies of \mathcal{T} to be finite, the caves must be exited by a different path every time; however, since we want the caves to “save” tile types, these “exit paths” cannot all be new tile types. Therefore, one possible way to solve these constraints is to grow a regular monotonic path P_0 first, then build a cave C , and reuse a part of P_0 as its exit path. The next time we want to reuse C , we can use another part of P_0 as its exit path.

If the exit paths used in previous instance of a cave are all blocked in new instances, we will get only finite assemblies. Figure 2 is an example of such a path.

This figure, however, is not a terminal assembly; since our caves, and their exit paths, are used several times, the same assemblies can grow from all their repetitions. Fortunately, we can arrange the shape of our main path so that no collision ever happens between these repetitions: see Figure 3. Now, this figure has 38 tile types, and is of width 27; it does not yet save tile types. But by inserting:

- n new tile types in place of glue 6,
- n new tile types in place of glue 14,
- n new tile types in place of glue 24, and
- n new tile types in place of glue 26

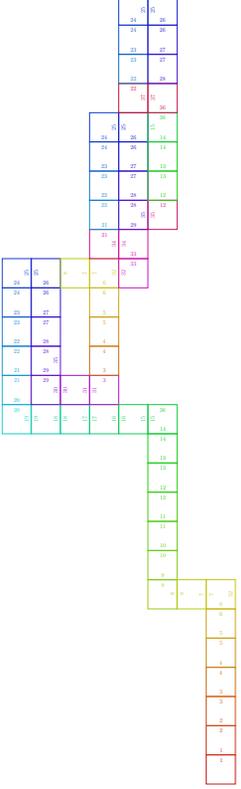


Figure 2: Reusing caves and exit paths

We add only $4n$ tile types, but the assembly is now $5n$ wider, and the result follows. An example of path that actually saves tile types is given on Figure 4.

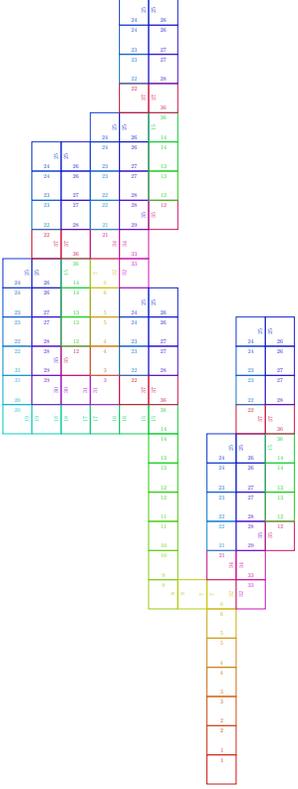


Figure 3: Early branches, not blocking the main assembly

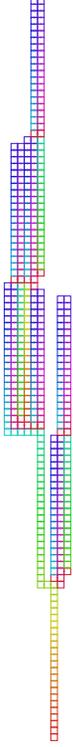


Figure 4: A path of width 112 with 106 tile types

□

In the rest of this paper, we will say that a path from A to B on some temperature 1 tile assembly system \mathcal{T} is *efficient* if it can be assembled using strictly less than $\|\overrightarrow{AB}\|_1 - 1$ tile types, and all of \mathcal{T} 's productions are finite.

4 Building filled squares requires $2n - 1$ tile types

In this section, we prove that if a tileset has less than $2n - 1$ tile types and has squares in its terminal assemblies, then it also has terminal assemblies that are not squares. Our technique to prove this will be the following: assuming we are given such a tileset, we will first choose, using Algorithm 4.1, the assembly sequences that “lose” as much information as possible about its past,

so as to “confuse” the tileset; then, if these assembly sequences can still build efficient paths, we will prove a necessary condition on these, (this is Lemma 4.4), that cannot hold in completely filled squares (Theorem 4.5).

4.1 A path-building algorithm

We first define the Algorithm we use to choose assembly sequences. There are two possible “priority modes” for this algorithm, namely *right-priority* and *left-priority*:

Algorithm 4.1. Let $\mathcal{T} = (T, \sigma, 1)$ be a tile assembly system, \preceq be an ordering on T , and S be a non-empty set of paths producible by \mathcal{T} . Let P_0 be the initial path, with just σ . For any $i \geq 0$, if several tiles can bind to P_i , let P_{i+1} be the one such that:

1. P_{i+1} binds to the first side of P_i from its input side, in clockwise order if we are building a right-priority path, and in counterclockwise order if we are building a left-priority one.
2. There is at least one path in S , of which $P_0 \dots P_{i+1}$ is a prefix.
3. If there are several such choices, we choose the smallest tile with respect to \preceq .

4.2 Building filled squares from a seed in a corner

Proving the claimed lower bound is significantly simpler if we assume that the seed is in a corner. We begin by showing that if paths built using Algorithm 4.1 can still be efficient, then we can *contract* them:

Definition 4.2. A *left-tentacular* (respectively *right-tentacular*) path is a path P such that at least one tile is repeated twice along P , say at indices i and j , and the suffix of P after j can be completely grown from i .

Formally, this last condition can be written:

$$\{P_k | 0 \leq k < |P|\} \cap \{P_k + \overrightarrow{P_j P_i} | j \leq k < |P|\} = \emptyset$$

Moreover, we say that a path is *two-way tentacular* if it is both left- and right-tentacular.

Definition 4.3. Let $\mathcal{T} = (T, \sigma, 1)$ be a tile assembly system. A path $P \in \mathcal{A}[\mathcal{T}]$ is said to be *fragile* if a tile type is repeated along P , say at positions i and j ($i < j$), and P cannot grow from $P_{0\dots i} \cup (P_{j\dots |P|-1} + \overrightarrow{P_j P_i})$.

In the rest of the proof, this choice of assembly sequence will be referred to as “breaking P ”.

Lemma 4.4. Let A be a point of \mathbb{Z}^2 , and $\mathcal{T} = (T, \sigma, 1)$ be a tile assembly system such that:

- $\text{dom}(\sigma) = \{(0, 0)\}$.
- $\forall \alpha \in \mathcal{A}[\mathcal{T}], \text{dom}(\alpha) \subseteq \{0, \dots, n-1\}^2$.
- There is a path $P \in \mathcal{A}[\mathcal{T}]$ from O to A , with less than $\|\overrightarrow{OA}\|_1$ tile types.

Then P is either fragile or two-way tentacular.

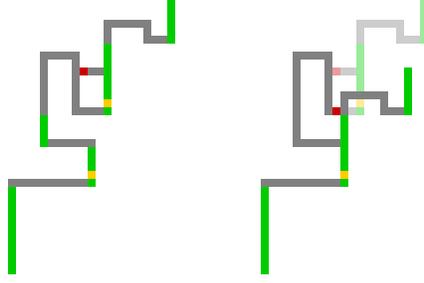


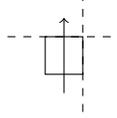
Figure 5: Glues visible from the east are in green or yellow; yellow tiles are of the same type, and red tiles are crashes.

Proof. Let P be a path from O to A , built using Algorithm 4.1. Without loss of generality, we assume that $y_A \geq 0$ and $x_A \geq 0$.

We first introduce the idea of *visible glue*: we say that a glue between two tiles of P is visible from the east (respectively from the north) if P_i interacts with P_{i+1} on its north face (respectively on its east face), and the glue between them is the rightmost (respectively topmost) one, on an infinite horizontal (respectively vertical) line between P_i and P_{i+1} .

Let us call P_E the tiles of P that have their north glue visible from the east, and P_S the tiles of P that have their east glue visible from the south.

There are clearly n tiles in P_E , and n tiles in P_S . Moreover, P_E and P_S are disjoint: indeed, if P_i is a tile whose north glue is visible from the east, then P still has to cross the vertical line on the right of P_i , before reaching B :



And the horizontal line, along with P and the borders of the square, define a closed curve: by Jordan's curve theorem, this curve partitions the plane into two connected components. Since P starts in one and ends in the other, this crossing can only occur above P_i . Indeed, the only way for P to go back into the seed's component is to cross the horizontal line above P_i , but then P_i would not be in P_E . Therefore, P_i cannot be in P_S .

Now, because P was grown using Algorithm 4.1, if a tile type is repeated in P_E , say at positions P_i and P_j , then a translation of $P_{i\dots}$ (by $\overrightarrow{P_i P_j}$ started growing after P_j , ultimately crashing into something (possibly after several repetitions). A new branch was started from this crashed path and reached B (in the case where this new path reached B , we consider the empty branch from B to itself as a valid branch). Therefore, since the two north glues are still visible after this, the early restart of this branch can grow from $P_{i\dots j}$; If it can grow completely from P , it means P is right-tentacular; else, this new branch crashes into something.

In this last case, grow P until P_i . Then, grow a translation by $\overrightarrow{P_j P_i}$ of the path that grew after P_j . Whether or not this translated subpath can grow completely, the original P cannot grow anymore from this new assembly: P is fragile.

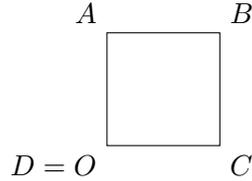
The same argument on the tiles of P_W and P_N proves that P is also left-tentacular or fragile. Remark that $P_W \cup P_N$ is not necessarily disjoint from $P_E \cup P_S$ (in the case where P has strictly more than $2n$ tile types, these two sets may even be equal). \square

We can now prove the claimed result:

Theorem 4.5. *Let $\mathcal{T} = (T, \sigma, 1)$ be a tile assembly system whose terminal assemblies' domains are all $\{0, \dots, n-1\}^2$, and such that $\text{dom}(\sigma) = \{(0, 0)\}$. Then:*

$$|T| \geq 2n - 1$$

Proof. Let A, B, C and D be the following points of a square that \mathcal{T} can assemble:



Since \mathcal{T} can fill this square, it must contain in particular a path from D , to A . Let P_0 be the path from O to A built right-priority using Algorithm 4.1.

We now define a sequence $(S_i)_i$ of assemblies, S_0 being the assembly where only P_0 has grown. For any $i \geq 0$, let A_i be the leftmost point, that is to the right of S_i , but not in it. If there are several such points, let A_i be the highest one. Now, let P_{i+1} be a right-priority path from O to A_i , built using Algorithm 4.1. Moreover, let Q_{i+1} be the prefix of P_{i+1} that goes to the first occurrence of P_{i+1} 's rightmost or topmost point (whichever comes last on P_{i+1}).

There are three main cases:

1. If Q_{i+1} is fragile, we “break it”. Let r be the index of the last tile of Q_{i+1} that can grow with the assembly sequence described in the proof of Lemma 4.4. Moreover, let U_{i+1} be the assembly grown with that assembly sequence, and V_{i+1} the part of Q_{i+1} that cannot grow anymore, that is, $V_{i+1} = (Q_{i+1})_{r+1} \dots$

In the case where Q_{i+1} is a suffix of some path of S_i , it is possible that other parts of S_i had been assembled from V_{i+1} . Let S'_i be the subassembly of S_i , with exactly the parts of S_i that were not assembled from V_{i+1} .

Finally, let $S_{i+1} = S'_i \cup U_{i+1}$.

2. If Q_{i+1} is right-tentacular, there are two subcases:

- (a) Its first contraction vector $\overrightarrow{v_{i+1}}$ is to the left, i.e. $x_{\overrightarrow{v_{i+1}}} \leq 0$, and $Q_{i+1} = P_{i+1}$, that is, the leftmost point to the right of S_i is not on Q_{i+1} itself. In this case, we can do the same as in case 1 above. Indeed, since the last point of Q_{i+1} is next to the leftmost point to the right of S_i , contracting it to the left will necessarily end in a collision with S_i , before the end of P_{i+1} .

- (b) Its first contraction vector is strictly to the right. Let a and b be the indices of the two contracted tiles on P_{i+1} , with $a < b$, and $R_{i+1} = (P_{i+1})_{0\dots a} \cup ((P_{i+1})_{b+1\dots} + \overrightarrow{(P_{i+1})_b(P_{i+1})_a})$

It is not possible that $B \in \text{dom}(P_{i+1})$, for else R_{i+1} would get out of $\{0, \dots, n-1\}^2$, contrarily to our hypothesis that all the assemblies stay within $\{0, \dots, n-1\}^2$. Therefore, we let this early restart grow as long as it can, until it reaches its highest point or else crashes into something.

3. Else, by Lemma 4.4, Q_{i+1} has the right number of tile types. There are two subcases:

- (a) if $A_i \neq B$, we let $S_{i+1} = S_i \cup P_{i+1}$, and resume the construction.
- (b) if $A_i = B$, the construction is over. Indeed, we have found a path, $P_{i+1} = Q_{i+1}$, from O to B , with $\|\overrightarrow{OB}\|_1$ distinct tile types.

Finally, we need to argue that this construction does not loop, even though cases 1 and 2a seem to make the assembly smaller. However, all the paths we grow in this construction are right-priority paths. This means that after these cases happen, the same points A_i will appear again, but the path that we can build to these repeated A_i s will be “less right-priority” than the original ones, that is, they will turn left earlier than the original path we grew to reach A_i .

We examine what happens to these “broken paths” in greater detail now. When a path is broken, and we branch and grow a new one to reach some A_j , that is a repetition of an earlier A_i , one of three things could happen:

1. Either P_j is fragile, or its contraction vector is to the left. In this case, nothing can happen to earlier broken parts: either they are to the left of this branch, in which case they are trapped, or they are to the right, but even if a collision happens between the early restart on P_j , and these parts, this collision is either with the original fragile part, which still keeps it broken, or between the early restart on P_i , which still encloses P_i .
2. In the case where P_j cannot be broken, that is, in case 2a, the construction will continue. If P_j 's early restart crashes into something, this crash does not affect any of the arguments.

□

Corollary 4.6. *Let $\mathcal{T} = (T, \sigma, 1)$ be a tile assembly system whose terminal assemblies are all of domain $\{0, \dots, n-1\}^2$, and such that $|\text{dom}(\sigma)| = 1$. Then $|T| \geq n$.*

Proof. Let Σ be the position of the seed. We can use the technique of Theorem 4.2 on all the rectangles with diagonals (ΣA) , (ΣB) , (ΣC) and (ΣD) . At least one is of width and height at least n . □

5 Future work

The next step, in proving the fully general conjecture, is to extend Lemma 4.4 to the case where the seed can be anywhere in the square. The reason why it does not apply to that case, is that a “lower restart” is not necessarily an “early restart”; indeed, let P be a path from O to D , and Q be a path from O to B . If P_i and Q_j are of the same type for some i and j , then restarting P_i from Q_j does not result in a competition for growth between this “lower branch” and Q , since the tiles of Q , to the north of Q_j , were grown before Q_j .

Therefore, despite Corollary 4.6, the general question remains open:

Open Problem 1. Is there a tile assembly system with less than $2n - 1$ tile types, that can assemble a filled square at temperature 1, starting from a single tile anywhere in the square?

Even though the recent results about its intrinsic universality [14], and the present paper, have made significant advances in that direction, the exact computational power of temperature 1 systems is still completely unknown. Moreover, the existence of single tileset simulating, at temperature 1, any other temperature 1 tile assembly system, is still open. The following open problem is particularly puzzling, especially in regard of the impressive results of [5], showing that three-dimensional temperature 1 tile assembly are capable of Turing computation:

Open Problem 2. Is there a tileset T , and an undecidable first order formula F on the following language:

- Constants $T \cup \mathbb{Z}^2$
- The addition on \mathbb{Z}^2
- The ternary relation $\rightarrow^{(T, \sigma, 1)}$ defined in section 2 (σ is an argument of the relation)
- And a ternary relation \mathcal{R} , where $\mathcal{R}(\alpha, (x, y), t)$ if and only if $\alpha(x, y) = t$ (in the notations of this paper).

Moreover, Theorem 3.1 is the first two-dimensional construction at temperature 1, with less tile types than its Manhattan diameter, in the general aTAM. On the other hand, our solution to the original conjecture relies heavily on the fact that our squares need to be completely filled. This leaves the following question open:

Open Problem 3. For all n , is there a tile assembly system $\mathcal{T}_n = (T_n, \sigma_n, 1)$, with $T_n < 2n - 1$, whose terminal assemblies all contain at least a square frame, that is, such that for all $\alpha \in \mathcal{A}_\square[\mathcal{T}_n]$, $(\{0\} \times \{0, \dots, n-1\}) \cup (\{0, \dots, n-1\} \times \{n-1\}) \cup (\{n-1\} \times \{0, \dots, n-1\}) \cup (\{0, \dots, n-1\} \times \{0\}) \subseteq \text{dom}(\alpha)$?

The initial construction of a square with $2n - 1$ tile types, by Rothmund and Winfree, used a fairly simple design. Our result shows that this is optimal, but it does not discard the possibility of other designs:

Open Problem 4. Is there a way to self-assemble a square of size $n \times n$ at temperature 1 with $2n - 1$ tile types, that is not a trivial variation of Rothmund and Winfree’s “comb” design (Figure 1)? How many are there?

6 Acknowledgments

I thank Damien Woods and David Doty for friendly support and discussions, Erik Winfree and his group for hosting me during this research, and Paul Rothmund for encouragements. I would like to thank Damien Woods again for taking me with him on the journey [29]. Section 1 of this paper is partially stolen from [14], I would like to thank my co-authors on that paper.

References

- [1] Leonard Adleman, Qi Cheng, Ashish Goel, and Ming-Deh Huang. Running time and program size for self-assembled squares. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, pages 740–748, Hersonissos, Greece, 2001.
- [2] Leonard M. Adleman, Qi Cheng, Ashish Goel, Ming-Deh A. Huang, David Kempe, Pablo Moisset de Espanés, and Paul W. K. Rothmund. Combinatorial optimization problems in self-assembly. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, pages 23–32, 2002.
- [3] Nathaniel Bryans, Ehsan Chiniforooshan, David Doty, Lila Kari, and Shinnosuke Seki. The power of nondeterminism in self-assembly. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA ’11*, pages 590–602. SIAM, 2011.

- [4] Harish Chandran, Nikhil Gopalkrishnan, and John Reif. Tile complexity of approximate squares. *Algorithmica*, 66(1):1–17, 2013.
- [5] Matthew Cook, Yunhui Fu, and Robert T. Schweller. Temperature 1 self-assembly: deterministic assembly in 3D and probabilistic assembly in 2D. In *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 570–589, 2011.
- [6] Erik D. Demaine, Matthew J. Patitz, Trent A. Rogers, Robert T. Schweller, Scott M. Summers, and Damien Woods. The two-handed tile assembly model is not intrinsically universal. In *ICALP: 40th International Colloquium on Automata, Languages and Programming*, volume 7965 of *LNCS*, pages 400–412, Riga, Latvia, July 2013. Springer.
- [7] David Doty, Jack H. Lutz, Matthew J. Patitz, Robert T. Schweller, Scott M. Summers, and Damien Woods. The tile assembly model is intrinsically universal. In *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science*, pages 439–446, October 2012.
- [8] David Doty, Jack H. Lutz, Matthew J. Patitz, Scott M. Summers, and Damien Woods. Intrinsic universality in self-assembly. In *Proceedings of the 27th International Symposium on Theoretical Aspects of Computer Science*, pages 275–286, 2009.
- [9] David Doty, Matthew J. Patitz, and Scott M. Summers. Limitations of self-assembly at temperature 1. *Theoretical Computer Science*, 412(1–2):145–158, 2011.
- [10] Kenichi Fujibayashi, Rizal Hariadi, Sung Ha Park, Erik Winfree, and Satoshi Murata. Toward reliable algorithmic self-assembly of DNA tiles: A fixed-width cellular automaton pattern. *Nano Letters*, 8(7):1791–1797, 2007.
- [11] James I. Lathrop, Jack H. Lutz, Matthew J. Patitz, and Scott M. Summers. Computability and complexity in self-assembly. *Theory Comput. Syst.*, 48(3):617–647, 2011.
- [12] Kyle Lund, Anthony T. Manzo, Nadine Dabby, Nicole Micholotti, Alexander Johnson-Buck, Jeanetter Nangreave, Steven Taylor, Renjun Pei, Milan N. Stojanovic, Nils G. Walter, Erik Winfree, and Hao Yan. Molecular robots guided by prescriptive landscapes. *Nature*, 465:206–210, 2010.
- [13] Ján Maňuch, Ladislav Stacho, and Christine Stoll. Two lower bounds for self-assemblies at temperature 1. *Journal of Computational Biology*, 17(6):841–852, 2010.
- [14] Pierre-Etienne Meunier, Matthew J. Patitz, Scott M. Summers, Guillaume Theyssier, Andrew Winslow, and Damien Woods. Intrinsic universality in tile self-assembly requires cooperation. *CoRR*, abs/1304.1679, 2013.
- [15] Matthew J. Patitz, Robert T. Schweller, and Scott M. Summers. Exact shapes and Turing universality at temperature 1 with a single negative glue. In *DNA 17: Proceedings of the Seventeenth International Conference on DNA Computing and Molecular Programming*, LNCS, pages 175–189. Springer, September 2011.
- [16] Lulu Qian and Erik Winfree. Scaling up digital circuit computation with DNA strand displacement cascades. *Science*, 332(6034):1196, 2011.
- [17] Lulu Qian, Erik Winfree, and Jehoshua Bruck. Neural network computation with DNA strand displacement cascades. *Nature*, 475(7356):368–372, 2011.

- [18] Paul W. K. Rothemund. Folding DNA to create nanoscale shapes and patterns. *Nature*, 440(7082):297–302, March 2006.
- [19] Paul W. K. Rothemund and Erik Winfree. The program-size complexity of self-assembled squares (extended abstract). In *STOC '00: Proceedings of the thirty-second annual ACM Symposium on Theory of Computing*, pages 459–468, Portland, Oregon, United States, 2000. ACM.
- [20] Paul W.K. Rothemund, Nick Papadakis, and Erik Winfree. Algorithmic self-assembly of DNA Sierpinski triangles. *PLoS Biology*, 2(12):2041–2053, 2004.
- [21] Georg Seelig, David Soloveichik, David Yu Zhang, and Erik Winfree. Enzyme-free nucleic acid logic circuits. *science*, 314(5805):1585–1588, 2006.
- [22] Nadrian C. Seeman. Nucleic-acid junctions and lattices. *Journal of Theoretical Biology*, 99:237–247, 1982.
- [23] Michael Sipser. *Introduction to the Theory of Computation*. International Thomson Publishing, 1st edition, 1996.
- [24] David Soloveichik and Erik Winfree. Complexity of self-assembled shapes. *SIAM Journal on Computing*, 36(6):1544–1569, 2007.
- [25] Hao Wang. Proving theorems by pattern recognition – II. *The Bell System Technical Journal*, XL(1):1–41, 1961.
- [26] Bryan Wei, Mingjie Dai, and Peng Yin. Complex shapes self-assembled from single-stranded DNA tiles. *Nature*, 485(7400):623–626, 2012.
- [27] Erik Winfree. *Algorithmic Self-Assembly of DNA*. PhD thesis, California Institute of Technology, June 1998.
- [28] Erik Winfree, Furong Liu, Lisa A. Wenzler, and Nadrian C. Seeman. Design and self-assembly of two-dimensional DNA crystals. *Nature*, 394(6693):539–44, 1998.
- [29] Damien Woods. Intrinsic universality and the computational power of self-assembly. *arXiv preprint arXiv:1309.1265*, 2013.
- [30] Bernard Yurke, Andrew J Turberfield, Allen P Mills, Friedrich C Simmel, and Jennifer L Neumann. A DNA-fuelled molecular machine made of DNA. *Nature*, 406(6796):605–608, 2000.