Size-Separable Tile Self-Assembly: A Tight Bound for Temperature-1 Mismatch-Free Systems *

Andrew Winslow

Tufts University, Medford, MA 02155, USA.
awinslow@cs.tufts.edu

Abstract. We introduce a new property of tile self-assembly systems that we call *size-separability*. A system is size-separable if every terminal assembly is a constant factor larger than any intermediate assembly. Size-separability is motivated by the practical problem of filtering completed assemblies from a variety of incomplete "garbage" assemblies using gel electrophoresis or other mass-based filtering techniques.
Here we prove that any system without cooperative bonding assembling a unique mismatch-free terminal assembly can be used to construct a size-separable system uniquely assembling the same shape. The proof achieves optimal scale factor, temperature, and tile types (within a factor of 2) for the size-separable system. As part of the proof, we obtain two results of independent interest on mismatch-free temperature-1 two-handed systems.

Keywords: 2HAM, hierarchical, aTAM, glues, gel electrophoresis

1 Introduction

The study of theoretical tile self-assembly was initiated by the Ph.D. thesis of Erik Winfree [18]. He proved that systems of passive square particles (called *tiles*) that attach according to matching bonds (called *glues*) are capable of universal computation and efficient assembly of shapes such as squares. Soloveichik and Winfree [16] later proved that these systems are capable of efficient assembly of any shape, allowing for an arbitrary scaling of the shape, used to embed a roving Turing machine. In this original *abstract Tile Assembly Model (aTAM)*, tiles attach singly to a growing seed assembly.

An alternative model, called the *two-handed assembly model (2HAM)* [1,2,4,5], *hierarchical tile assembly model* [3,13], or *polyomino tile assembly model* [8,9], allows "seedless" assembly, where tiles can attach spontaneously to form large assemblies that may attach to each other. This seedless assembly was proved by Cannon et al. [2] to be capable of simulating any seeded assembly process, while also achieving more efficient assembly of some classes of shapes.

* A full version of this paper can be found at http://arxiv.org/abs/1404.7410

A generalization of the 2HAM called the *staged tile assembly model* introduced by Demaine et al. [4] utilizes sequences of *mixings*, where each mixing combines a set of *input assemblies* using a 2HAM assembly process. The products of the mixing are the *terminal assemblies* that cannot combine with any other assembly produced during the assembly process (called a *producible assembly*). This set of terminal assemblies can then be used as input assemblies in another mixing, combined with the sets of terminal assemblies from other mixings.

After a presentation by the author of work [19] on the staged self-assembly model at DNA 19, Erik Winfree commented that the staged tile assembly model has a unrealistic assumption: at the end of each mixing process, all producible but non-terminal assemblies are removed from the mixing. A similar assumption is made in the 2HAM model, where only the terminal assemblies are considered to be "produced" by the system.

Ignoring large producible assemblies is done to simplify the model definition, but allows unrealistic scenarios where "nearly terminal" systems differing from some terminal assembly by a small number of tiles are presumed to be eliminated or otherwise removed at the end of the assembly process. While filtering techniques, including well-known gel electrophoresis, may be employed to obtain filtering of particles at the nanoscale, such techniques generally lack the resolution to distinguish between macromolecules that differ in size by only a small amount.

Our results. In this work, we consider efficient assembly of shapes in the 2HAM model under the restriction that terminal assemblies are significantly larger than all non-terminal producible assemblies. We call a system *factor-c size-separable* if the ratio between the smallest terminal assembly and largest non-terminal producible assembly is at least c. Thus, high-factor size-separable systems lack large but non-terminal assemblies, allowing robust filtering of terminal from non-terminal assemblies in these systems.

Our main result is an algorithm for converting 2HAM systems of a special class into size-separable 2HAM systems. A 2HAM system $\mathcal{S} = (T, f, \tau)$ consists of a set of *tiles* T that attach by forming bonds according to their *glues* and a *glue-strength function* f, and two assemblies can attach if the total strength of the bonds formed meets or exceeds the temperature τ of the system. If a system is temperature-1 ($\tau = 1$), then any two assemblies can attach if they have a single matching glue. An assembly is said to be *mismatch-free* if no two coincident tile sides in the assembly or any assembly in the system have different glues. We prove the following:

Theorem 1. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with a mismatch-free unique terminal assembly A. Then there exists a factor-2 size-separable 2HAM system $\mathcal{S}' = (T', f', 2)$ with a unique mismatch-free finite terminal assembly A' such that $|\mathcal{S}'| \leq 8|\mathcal{S}|$ and A' has the shape of A scaled by a factor of 2.*

Along the way, we prove two results of independent interest on temperature-1 mismatch-free systems. The *bond graph* of an assembly A, denoted $G(A)$, is the dual graph of A formed by a node for each tile, and an edge between two tiles if they form a bond. We show that any system with a unique mismatch-free finite terminal assembly whose bond graph is not a tree can be made so without increasing the number of tile types in the system:

Lemma 7 (Tree-ification Lemma). *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with unique mismatch-free finite terminal assembly A. Then there exists a 2HAM system $\mathcal{S}' = (T', f', 1)$ with unique mismatch-free finite terminal assembly A' and $|\mathcal{S}'| \leq |\mathcal{S}|$, where A' has the shape of A and $G(A')$ is a tree.*

The proof of the Tree-ification Lemma yields a simple algorithm for obtaining $\mathcal{S}'$: while a cycle in $G(A)$ remains, remove a glue on this cycle from the tile type containing it. The challenge is in proving such a process does not disconnect $G(A)$, regardless of the glue and cycle chosen.

We also prove that the tile types used only once in a unique terminal assembly, called *1-occurrence tiles*, form a connected subgraph of $G(A)$. That is, these tiles taken alone form a valid assembly.

Lemma 9. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with unique mismatch-free finite terminal assembly A. Then the 1-occurrence tiles in A form a 1-stable subassembly of A.*

For some questions about temperature-1 systems, results have been far easier to obtain for mismatch-free systems than for general systems allowing mismatches. For instance, a lower bound of $2n - 1$ for the assembly of a $n \times n$ square by any temperature 1 aTAM system was conjectured by Rothemund and Winfree [15], and proved for mismatch-free systems by Maňuch, Stacho, and Stoll [10]. Meunier [11] was able to show the same lower bound for systems permitted to have mismatches under the assumption that the seed tile starts in the lower left of the assembly, and removing this restriction remains open. In a similar vein, Reif and Song [14] have shown that temperature-1 mismatch-free aTAM systems are not computationally universal, while the same problem for systems with mismatches permitted is a notoriously difficult problem that remains open, despite significant efforts [7,6,17,12].

In spite of such results, constructing high-factor size-separable versions of temperature-1 mismatch-free systems remains challenging. One difficulty lies in the partitioning the assembly into two equal-sized halves that will come together for the final assembly step. Note that for many assemblies, such a cutting is impossible (e.g. the right assembly in Figure 1). Even if such a cutting is possible, removing the bonds connecting the two halves by modifying the tiles along the boundary may require a large increase in the number of tile types of the system.

Another challenge lies in coping with cycles in the bond graph. Factor-2 size-separability requires that the last assembly step consists of two completely assembled halves of the unique terminal assembly attaching. Cycles in the bond graph (e.g. the left assembly in Figure 1) prevent communication between the tiles inside and outside of the cycles, risking the possibility that the portion of the assembly inside a cycle still has missing tiles as the exterior takes part in the supposed final assembly step.

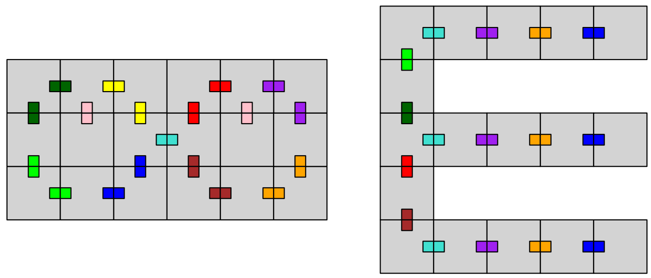

Fig. 1. Unique mismatch-free terminal assemblies of two different temperature-1 2HAM systems. Constructing high-factor size-separable versions of these systems is challenging due to the existence of cycles (left) and lack of equal-sized halves (right).

Loosely speaking, our approach is to first construct a version of A where the bond graph is a tree and a vertex cut of $G(A)$ consisting of a path of 1-occurrence tiles exists. This modified version of A is then scaled in size and temperature by a factor of 2, using special 2×2 macrotiles that only assemble along the boundary of the scaled assembly via mixed-strength bonds. Finally, the 1-occurrence tiles forming a vertex cut are given weakened glues such that only completely formed subassemblies on both sides of the cut can attach across the weak-glue cut.

2 Definitions

Here we give a complete set of formal definitions of tile self-assembly used throughout the paper. All of the definitions used are equivalent to those found in prior work on the two-handed tile assembly model, e.g. [1,2,3,13].

Assembly systems. In this work we study the *two-handed tile assembly model (2HAM)*, and instances of the model called *systems*. A 2HAM system $\mathcal{S} = (T, f, \tau)$ is specified by three parts: a *tile set* T, a *glue-strength function* f, and a *temperature* $\tau \in \mathbb{N}$.

The tile set T is a set of unit square *tiles*. Each tile $t \in T$ is defined by 4-tuple $t = (g_n, g_e, g_s, g_w)$ consisting of four *glues* from a set Σ of *glue types*, i.e. $g_n, g_e, g_s, g_w \in \Sigma$. The four glues g_n, g_e, g_s, g_w specify the glue types in Σ found on the north (N), east (E), south (S), and west (W) sides of t, respectively. Each glue also defines a *glue-side*, e.g. (g_n, N). Define $g_D(t)$ to be the glue on the side D of t, e.g. $g_{\mathrm{N}}(t) = g_n$.

The glue function $f : \Sigma^2 \rightarrow \mathbb{N}$ determines the strength of the *bond* formed by two coincident glue-sides. For any two glues $g, g' \in \Sigma$, $f(g, g') = f(g', g)$. A unique *null glue* $\varnothing \in \Sigma$ has the property that $f(\varnothing, g) = 0$ for all $g \in \Sigma$. In this work we only consider glue functions such that for all $g, g' \in \Sigma$, $f(g, g') = 0$ and if $g \neq \varnothing, f(g,g) > 0$. For convenience, we sometimes refer to a glue-side with the null glue as a side *without a glue*.

Configurations and assemblies. A *configuration* is a partial function $C : \mathbb{Z}^2 \rightarrow T$ mapping locations on the integer lattice to tiles. Define $L_D(x,y)$ to be the location in $\mathbb{Z}^2$ one unit in direction D from (x,y), e.g. $L_{\mathrm{N}}(0,0) = (0,1)$. For any pair of locations $(x,y), L_D(x,y) \in C$, the *bond strength* between the these tiles is $f(g_D(C(x,y)), g_{D^{-1}}(C(L_D(x,y))))$. If $g_D(C(x,y)) \neq g_{D^{-1}}(C(L_D(x,y)))$, then the pair of tiles is said to form a *mismatch*, and a configuration with no mismatches is *mismatch-free*. If $g_D(C(x,y)) = g_{D^{-1}}(C(L_D(x,y)))$, then the common glue and pair of directions define a *glue-side pair* $(g_D(C(x,y), \{D, D^{-1}\})$.

The *bond graph of* C, denoted $G(C)$, is defined as the graph with vertices $\mathrm{dom}(C)$ and edges $\{((x,y), L_D(x,y)) : f(g_D(C(x,y)), g_{D^{-1}}(C(L_D(x,y)))) > 0\}$. That is, the graph induced by the neighboring tiles of C forming positive-strength bonds.

A configuration C is a *τ-stable assembly* or an *assembly at temperature τ* if $\mathrm{dom}(C)$ is connected on the lattice and, for any partition of $\mathrm{dom}(C)$ into two subconfigurations C_1 and C_2, the sum of the bond strengths between tiles at pairs of locations $p_1 \in \mathrm{dom}(C_1)$, $p_2 \in \mathrm{dom}(C_2)$ is at least τ, the temperature of the system. Any pair of assemblies A_1, A_2 are equivalent if they are identical up to a translation by $\langle x, y \rangle$ with $x, y \in \mathbb{Z}$. The *size* of an assembly A is $|\mathrm{dom}(A)|$, and $t \in T$ is a *k-occurrence tile* in A if $|\{(x,y) \in \mathrm{dom}(A) : A(x,y) = t\}| = k$. The *shape* of an assembly is the polyomino induced by $\mathrm{dom}(A)$, and a shape is *scaled by a factor k* by replacing each cell of the polyomino with a $k \times k$ block of cells.

Two τ-stable assemblies A_1, A_2 are said to *assemble* into a *superassembly* A_3 if A_2 is equivalent to an assembly A_2' such that $\mathrm{dom}(A_1) \cap \mathrm{dom}(A_2') = \emptyset$ and A_3 defined by the union of the partial functions A_1 and A_2' is a τ-stable assembly. Similarly, an assembly A_1 is a *subassembly* of A_2, denoted $A_1 \subseteq A_2$, if A_2 is equivalent to an assembly A_2' such that $\mathrm{dom}(A_1) \subseteq \mathrm{dom}(A_2')$.

Producible and terminal assemblies. An assembly A is a *producible assembly* of a 2HAM system $\mathcal{S}$ if A can be assembled from two other producible assemblies or A is a single tile in T. A producible assembly A is a *terminal assembly* of $\mathcal{S}$ if A is producible and A does not assemble with any other producible assembly of $\mathcal{S}$.

We also consider *seeded* versions of some 2HAM systems, where an assembly is producible if it can be assembled from another producible assembly and a single tile of T. Note that for any temperature-1 2HAM system $\mathcal{S}$, the seeded version of $\mathcal{S}$ has the same set of terminal assemblies as $\mathcal{S}$.

If $\mathcal{S}$ has a single terminal assembly A, we call A the *unique terminal assembly (UTA)* of $\mathcal{S}$. In the case that $|A|$ is finite and mismatch-free, we further call A the *unique mismatch-free finite terminal assembly (UMFTA)* of $\mathcal{S}$.

Size-separability. A 2HAM system $\mathcal{S} = (T, f, \tau)$ is a *factor-c size-separable* if for any pair of producible assemblies A, B of $\mathcal{S}$ with A terminal and B not terminal, $|A|/|B| \geq c$. Since this ratio is undefined when $\mathcal{S}$ has infinite producible assemblies, we define such a system to have undefined size-separability. Every system with defined size-separability has factor-c size-separability for some $1 \leq c \leq 2$.

3 Tree-ification

First, we prove that any $\tau = 1$ system producing a unique terminal assembly can be converted into a system with another unique terminal assembly with the same shape but whose bond graph is a tree.

Lemma 1. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system. Every 1-stable assembly consisting of tiles in T is a producible assembly of $\mathcal{S}$.*

Lemma 2. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UTA A. Let a glue-side pair appear twice on a simple cycle of $G(A)$ between tiles t_1 and t_2, and t_3 and t_4. Then $|\{t_1, t_2, t_3, t_4\}| \neq 4$.*

Lemma 3. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UMFTA A. Let a glue-side pair appear twice on a simple cycle of $G(A)$ between tiles t_1 and t_2, and t_3 and t_4. Then $|\{t_1, t_2, t_3, t_4\}| \neq 2$.*

Lemma 4. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UTA A. Let a glue-side pair appear twice on a simple cycle of $G(A)$ between tiles t_1 and t_2, and t_3 and t_4. Then $|\{t_1, t_2, t_3, t_4\}| \neq 3$.*

Lemma 5. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UMFTA A. Then no glue-side pair appears twice on a simple cycle of $G(A)$.*

Lemma 6. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UMFTA A. Let (g, p) be the glue-side pair of an edge e in $G(A)$. Then if e lies on a simple cycle in $G(A)$, all edges with glue-side pair (g, p) lie on simple cycles of $G(A)$.*

Lemma 7 (Tree-ification Lemma). *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UMFTA A. Then there exists a 2HAM system $\mathcal{S}' = (T', f', 1)$ with UMFTA A' and $|\mathcal{S}'| \leq |\mathcal{S}|$, where A' has the shape of A and $G(A')$ is a tree.*

4 1-Occurrence Tile Types

In addition to tree-ification, we also make use of the existence of *1-occurrence tile types*: tile types that appear only once in the terminal assembly of the system.

Lemma 8. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UMFTA A with $G(A)$ a tree and $|A| \geq 2$. Then A has at least two 1-occurrence tiles.*

Lemma 9. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UMFTA A. Then the 1-occurrence tiles in A form a 1-stable subassembly of A.*

Lemma 10. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UTA A. For any glue-side pair (g, p) occurring between a pair of 1-occurrence tiles in A, (g, p) occurs only once in A.*

Lemma 11. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UMFTA A with $G(A)$ a tree. For any tile $t \in T$, the simple path in $G(A)$ between any two occurrences of t uses the same glue-side of t on both occurrences.*

Lemma 12. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UMFTA A with $G(A)$ a tree. Let edges $e, e' \in G(A)$, with e' between a pair of 1-occurrence tiles. Then there exists a second 2HAM system $\mathcal{S}' = (T', f', 1)$ with $|T'| \leq 2|T|$ and UMFTA A' with $G(A') = G(A)$ and the unique path from e' to e in $G(A')$ consisting entirely of 1-occurrence tiles in A'.*

5 A Size-Separable Macrotile Construction

A simple barrier to general high-factor size-separability is the fact that any system with a tree-shaped unique terminal assembly A cannot be factor-c size-separable for any $c > 1 + 1/|A|$. A more subtle challenge is how to partition assemblies into equal-sized 1-stable halves that will come together in the final assembly step. We resolve both of these issues by creating a temperature-2 2HAM system with a unique terminal assembly whose shape is the shape of A scaled by a factor of 2, and whose bond graph has an edge cut of two temperature-1 bonds that partitions $G(A)$ into two subgraphs of equal size.

Lemma 13. *Let $\mathcal{S} = (T, f, \tau)$ be a 2HAM system with P and P' producible assemblies of $\mathcal{S}$ with P a proper subassembly of P'. Then P is not a terminal assembly.*

Lemma 14. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UMFTA A with $G(A)$ a tree. Then there exists a 2HAM system $\mathcal{S}' = (T', f', 2)$ with UMFTA A' and $|\mathcal{S}'| \leq 4|\mathcal{S}|$ such that A' has the shape of A scaled by a factor of 2.*

Proof. We start by describing common properties of all occurrences of each tile type $t \in T$. Since $G(A)$ is a tree, Lemmas 8 and 9 imply that there exists an edge e' in $G(A)$ between two 1-occurrence tiles and Lemma 11 implies that any path between two occurrences of t use the same glue-side pair. So any breadth-first search $G(A)$ starting at a 1-occurrence tile incident to e' visits all occurrences of t exactly once, and all via incoming edges from the same side of t. Then since A is mismatch-free, if a direction is applied to each edge of $G(A)$ according to the direction of traversal during the breadth-first search, all occurrences of t have the same set of incoming and outgoing edges. So all occurrences of t have their corners visited in the same order during a traversal of the boundary of A.

We use these conditions to construct unique macrotile versions of each tile type according to their incoming and outgoing edges induced by the breadth-first search starting at e'. All possible macrotile constructions (up to symmetry) are shown in Figure 2. For each glue-side pair in the original system, we use two glue-side pairs in the scaled system, one with strength-2 and the other with strength-1. The glue-side pair visited first in the counterclockwise traversal of the boundary starting at e' has strength 2, while the other pair has strength 1.

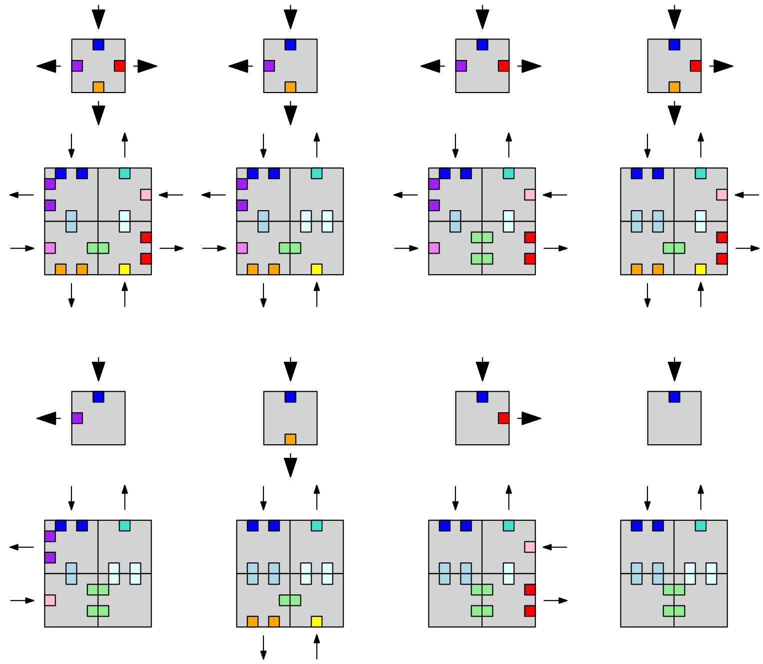

Fig. 2. The individual tiles enumerate (up to symmetry) all combinations of incoming and outgoing edges (large arrows) induced by a breath-first search of $G(A)$. The corresponding 2×2 macrotiles are used in the proof of Theorem 1 to construct a temperature-2 system that carries out the assembly of A at scale 2 in the order that the tiles appear along the boundary (small arrows). All internal macrotile glues are unique to the tile type, while all external macrotile glues correspond to the glues found on the surface of the inducing tile.

There are also three glues internal to each macrotile attaching each pair of adjacent tiles forming a macroedge whose corresponding edge of the tile either has an outgoing edge induced by the breadth-first search, or has no edge. These glues are unique to the macrotile type. The strengths of each of these glues is determined by whether the closest macroside has glues. If not, then the glue is strength-2, otherwise the glue is strength-1.

If the closest macroside does not contain a glue, then the strength-2 internal glue is necessary to allow assembly to continue along the boundary of the assembly in the counterclockwise direction (e.g. from the northwest to the southwest tile). If the closest macroside does contain a glue, then the strength-1 internal glue prevents the placement of the next tile in the macrotile (e.g. the southwest tile after the northwest tile) until a second tile from an adjacent macrotile (e.g. the southeast tile of the macrotile to the west) has been placed. As a result, no pair of tiles in a macrotile can be present in a common assembly unless all tiles between them along a counterclockwise traversal of the boundary of the macrotile assembly are also present.

Scaled assembly of A. We claim that this scaled version of the system has a unique terminal assembly A' obtained by replacing each tile in the original unique terminal assembly with the corresponding 2×2 macrotile. First we prove that any subassembly of A' corresponding to a subtree of $G(A)$ is producible. A subtree of size 1 corresponds to a leaf node, the lower-rightmost case in Figure 2, and is clearly producible. For larger subtrees, the assembly can be formed by combining the 4 tiles of the root macrotile to the (up to) three subtrees assemblies. Grow the assembly in counterclockwise order around the boundary, attaching either a subtree assembly (if the macroside has a glue) or the next tile of the root macrotile. In both cases, placing the second root tile along the macroedge is possible, as either the internal glue shared with the previous root tile is strength-2 or a second glue is provided by the subtree assembly. Then by induction, the assembly A' corresponding to the subtree rooted at the root of the breadth-first search is producible.

By construction, A' is terminal because it corresponds to a mismatch-free terminal assembly in the original system that necessarily has no exposed glues. So A' is a terminal assembly of the scaled system. Next, we prove that A' is the unique terminal assembly of the system.

Terminal assembly uniqueness. We start by proving that every producible assembly can positioned on a 2×2 macrotile *grid*, where every tile in the southwest corner is a southwest tile of some macrotile, every tile in the northwest corner is the northwest tile of some macrotile, etc. Start by noticing that each glue type appears coincident to only one of 12 edges of the grid: the 4 internal edges of each macrotile, and the 8 external edges. Suppose there is some smallest producible assembly that does not lie on a grid. Then this assembly must be formed by the attachment of two smaller assemblies that do lie on grids, and whose glues utilized in the attachment are coincident to only one of 12 edges of the grid. So if these assemblies are translated to have coincident matching glue sides, then their grids must also be aligned and the assembly resulting from their attachment also lies on the grid, a contradiction.

Let A'_p be a producible assembly of the macrotile system that is not A'. Construct an assembly A_p of the original input system $\mathcal{S}$ in the following way: replace each macrotile region with a single tile corresponding to one of the tiles in the macrotile region. If such a replacement is unambiguous, meaning that all tiles in each macrotile region belong to a common macrotile, then the resulting assembly is a 1-stable (and thus producible) assembly of $\mathcal{S}$.

We also claim that such a replacement is always unambiguous. Suppose, for the sake of contradiction, that there is some A'_p such that replacement is ambiguous. The ambiguity must be due to two tiles in the same macrotile region bonded via external strength-2 glues on *different* macrosides to tiles in adjacent macrotiles, since no macroside has two strength-2 glues (see Figure 2). So there is some path in $G(A'_p)$ from the external glue of one of of these tiles to the external glue to the other consisting of length-2 and length-3 subpaths through other macrotile regions, each consisting of tiles of a common macrotile. So this path can be unambiguously replaced with a path from tiles in $\mathcal{S}$ from one side of a tile location to the other side, with some tile of $\mathcal{S}$ able to attach at this location. But this yields a producible assembly of $\mathcal{S}$ (and thus a subassembly of A) with a cycle, a contradiction. Since constructing A_p from A'_p is always unambiguous, and A_p is a subassembly of A, A'_p is a subassembly of A'. Then by Lemma 13, A'_p is not terminal. □

Theorem 1. *Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UMFTA A. Then there exists a factor-2 size-separable 2HAM system $\mathcal{S}' = (T', f', 2)$ with UMFTA A' and $|\mathcal{S}'| \leq 8|\mathcal{S}|$. Furthermore, A' has the shape of A scaled by a factor of 2.*

Proof (Sketch). We modify the construction used in the proof of Lemma 14 in two ways. First, we use Lemma 12 to create a path of 1-occurrence tiles in A that partitions A into two 1-stable subassemblies, each containing a contiguous, equal-sized half of the boundary of A. This is done to the original system $\mathcal{S}$, before the macrotile conversion is performed. Second, after constructing the macrotile system $\mathcal{S}'$, we modify some of the glues of the macrotiles corresponding to this path of 1-occurrence tiles to give the unique terminal assembly A' of the macrotile system a 2-edge cut. These two edges occur at opposite ends of the path of 1-occurrence tiles, and reducing their strength to 1 enforces that A' can only assemble from two equal-sized halves. □

6 Open Problems

For temperature-1 systems with mismatch-free unique terminal assemblies, our result is nearly as tight as possible. Scaling to at least a factor of 2 and using temperature of at least 2 are both necessary, since any temperature-1 system or system with a tree-shaped assembly is at most factor-$(1 + 1/|A|)$ size-separable. The only remaining opportunity for improvement is to reduce the number of tile types used to less than $8|\mathcal{S}|$.

We contend that our result is a first step in understanding what is possible in size-separable systems, and a large number of open problems remain. Perhaps the most natural problem is to extend this result to the same set of systems, except permitting mismatches. We conjecture that a similar result is possible there:

Conjecture 1. Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with unique finite terminal assembly A. Then there exists factor-2 size-separable system $\mathcal{S}' = (T', f', 2)$ with a unique finite terminal assembly A' and $|\mathcal{S}'| = O(\mathcal{S})$. Furthermore, A' has the shape A scaled by a factor of $O(1)$.

Extending the result to mismatch-free systems at higher temperatures also is of interest because these systems are generally capable of much more efficient assembly. Soloveichik and Winfree [16] prove that one can construct a temperature-2 system that uses an optimal number of tiles (within a constant factor) to construct any shape, provided one is allowed to scale the shape by an arbitrary amount, and it is likely their construction can be modified to be factor-2 size-separable. However, it remains open to achieve high-factor size-separable systems at temperature 2 using only a small scale factor.

Conjecture 2. Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with UMFTA A. Then there exists factor-2 size-separable system $\mathcal{S}' = (T', f', 2)$ with a unique terminal assembly A' and $|\mathcal{S}'| = O(\mathcal{S})$. Furthermore, A' has the shape A scaled by a factor of $O(1)$.

In the interest of applying size-separability to system in the staged model of tile self-assembly, we pose the problem of developing size-separable systems with multiple terminal assemblies. Of course, one can construct systems where the smallest terminal assembly is less than half the size of the largest terminal assembly, ensuring that the system cannot even be factor-1 size-separable. But given a system whose ratio of smallest to largest terminal assembly is c, is a size-separable system with optimal factor $\frac{2}{c}$ always possible?

Conjecture 3. Let $\mathcal{S} = (T, f, 1)$ be a 2HAM system with finite terminal assemblies $A_1, A_2, \ldots, A_k$ with A_1 and A_k the smallest and largest terminal assemblies. Then there exists factor-$|A_k|/|A_1|$ size-separable system $\mathcal{S}' = (T', f', 2)$ with $|\mathcal{S}'| = O(\mathcal{S})$ and mismatch-free terminal assemblies $A'_1, A'_2, \ldots, A'_k$ where A'_i has the shape of A_i scaled by a factor of $O(1)$.

We close by conjecturing that not every system can be made size-separable by paying only a constant factor in scale and tile types. We ask for an example of such a system:

Conjecture 4. There exists a 2HAM system $\mathcal{S} = (T, f, \tau)$ with a unique finite terminal assembly A such that any factor-2 size-separable system $\mathcal{S}' = (T', f', \tau')$ with unique finite terminal assembly A' with the shape of A either has $|\mathcal{S}'| \geq 100|\mathcal{S}|$ or the scale of A' is at least 100.

Acknowledgments

We thank the anonymous UCNC reviews for their comments that improved the presentation and correctness of the paper.

References

1. Z. Abel, N. Benbernou, M. Damian, E. D. Demaine, M. L. Demaine, R. Flatland, S. D. Kominers, and R. Schweller. Shape replication through self-assembly and RNase enzymes. In *Proceedings of the 21st ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1045–1064, 2010.
2. S. Cannon, E. D. Demaine, M. L. Demaine, S. Eisenstat, M. J. Patitz, R. T. Schweller, S. M. Summers, and A. Winslow. Two hands are better than one (up to constant factors): Self-assembly in the 2HAM vs. aTAM. In *STACS 2013*, volume 20 of *LIPIcs*, pages 172–184. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2013.
3. H. Chen and D. Doty. Parallelism and time in hierarchical self-assembly. In *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1163–1182, 2012.

4. E. D. Demaine, M. L. Demaine, S. P. Fekete, M. Ishaque, E. Rafalin, R. T. Schweller, and D. L. Souvaine. Staged self-assembly: nanomanufacture of arbitrary shapes with $O(1)$ glues. *Natural Computing*, 7(3):347–370, 2008.
5. D. Doty, M. J. Patitz, D. Reishus, R. T. Schweller, and S. M. Summers. Strong fault-tolerance for self-assembly with fuzzy temperature. In *Foundations of Computer Science (FOCS)*, pages 417–426, 2010.
6. D. Doty, M. J. Patitz, and S. M. Summers. Limitations of self-assembly at temperature 1. In R. Deaton and A. Suyama, editors, *DNA 15*, volume 5877 of *LNCS*, pages 35–44. Springer Berlin Heidelberg, 2009.
7. J. I. Lathrop, J. H. Lutz, M. J. Patitz, and S. M. Summers. Computability and complexity in self-assembly. In A. Beckmann, C. Dimitracopoulos, and B. Löwe, editors, *Logic and Theory of Algorithms*, volume 5028 of *LNCS*, pages 349–358. Springer Berlin Heidelberg, 2008.
8. C. Luhrs. Polyomino-safe DNA self-assembly via block replacement. In A. Goel, F. C. Simmel, and P. Sosik, editors, *DNA 14*, volume 5347 of *LNCS*, pages 112–126. 2009.
9. C. Luhrs. Polyomino-safe DNA self-assembly via block replacement. *Natural Computing*, 9(1):97–109, 2010.
10. J. Maňuch, L. Stacho, and C. Stoll. Journal of computational biology. *Journal of Computational Biology*, 16(6):841–852, 2010.
11. P.-E. Meunier. The self-assembly of paths and squares at temperature 1. Technical report, arXiv, 2014. http://arxiv.org/abs/1312.1299.
12. P.-E. Meunier, M. J. Patitz, S. M. Summers, G. Theyssier, A. Winslow, and D. Woods. Intrinsic universality in tile self-assembly requires cooperation. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 752–771, 2014.
13. J. Padilla, M. J. Patitz, R. Pena, R. T. Schweller, N. C. Seeman, R. Sheline, S. M. Summers, and X. Zhong. Asynchronous signal passing for tile self-assembly:fuel efficient computation and efficient assembly of shapes. In G. Mauri, A. Dennunzio, L. Manzoni, and A. E. Porreca, editors, *Unconventional Computation and Natural Computation (UCNC)*, volume 7956 of *LNCS*, pages 174–185. Springer Berlin Heidelberg, 2013.
14. J. Reif and T. Song. The computation complexity of temperature-1 tilings. Technical report, Duke University, 2014.
15. P. W. K. Rothemund and E. Winfree. The program-size complexity of self-assembled squares (extended abstract). In *Proceedings of ACM Symposium on Theory of Computing (STOC)*, pages 459–468, 2000.
16. D. Soloveichik and E. Winfree. Complexity of self-assembled shapes. *SIAM Journal on Computing*, 36(6):1544–1569, 2007.
17. S. M. Summers. *Universality in algorithm self-assembly*. PhD thesis, Iowa State University, 2010.
18. E. Winfree. *Algorithmic Self-Assembly of DNA*. PhD thesis, Caltech, 1998.
19. A. Winslow. Staged self-assembly and polyomino context-free grammars. In D. Soloveichik and B. Yurke, editors, *DNA 19*, 2013.