# The Tiling Enigma: covering a grid with colored simple polygons*

Arthur van Goethem[1], Irina Kostitsyna[2], Marc van Kreveld[3], Wouter Meulemans[1], Max Sondag[1], and Jules Wulms[1]

[1] Dept. of Mathematics and Computer Science, TU Eindhoven
{a.i.v.goethem | w.meulemans | m.f.m.sondag | j.j.h.m.wulms}@tue.nl
[2] Computer Science Department, Université Libre de Bruxelles
irina.kostitsyna@ulb.ac.be
[3] Dept. of Information and Computing Sciences, Utrecht University
m.j.vankreveld@uu.nl

**Abstract.** Motivated by a new way of visualizing hypergraphs, we study the following problem. Consider a regular grid and a set of colors $\chi$. Each cell $s$ in the grid is assigned a subset $\chi_s \subseteq \chi$ and should be partitioned such that for each color $c \in \chi_s$ at least one region in the cell is identified with $c$. Cells assigned the empty set remain white. We focus on the case where $\chi = \{\text{red}, \text{blue}\}$. Is it possible to partition each cell such that for the complete grid the resulting red and blue form two connected components? We analyze the combinatorial properties and derive a necessary and sufficient condition for such a *tiling*. We show that if a tiling exists, there also exists a tiling with bounded complexity per cell. This tiling has at most five colored regions per cell if the grid contains white cells, and at most two colored regions per cell if it does not.

## 1 Introduction

Hypergraphs are a powerful structure to represent unordered set systems. In general, there are a number of elements (vertices of the hypergraph) and a number of different subsets over these elements (the hyperedges of the graph). The purpose of visualizing hypergraphs is to clarify the various set relations between the hyperedges. There are, roughly speaking, two strands of hypergraph visualizations: those where the position of the elements is fixed (e.g. [2,7,8,15]), and those where the positions can be chosen by the layout algorithm (e.g. [10,19,18]). For a more detailed overview and in-depth classification of set visualization methods we refer to the survey by Alsallakh et al. [4]. Though some methods aim to overcome layout complexity by replicating elements (e.g. [3,10]), we focus on a visualization using a single representation for each element.

**Contributions.** We investigate the combinatorial and algorithmic properties of renderings of a hypergraph using disjoint polygons. In our setting, each element

* Research for this paper was initiated at the 2nd Workshop on Applied Geometric Algorithms (AGA 2017) in Vierhouten, the Netherlands, supported by the Netherlands Organisation for Scientific Research (NWO) under project no. 639.023.208.

is represented by a small simple polygon (typically, a square or circle) and we render each set as a polygon that overlaps exactly the elements that belong to the set. Motivated by moving towards a set visualization in a geographic small multiples or grid map (see e.g. [14,22]), we specifically study the variant where each element has a fixed location, being a cell in a unit-grid. As an initial exploration we focus on the 2-color case, where each cell is either red, blue, both (purple) or uncolored (white). We derive a necessary and sufficient condition to efficiently recognize whether an instance can be represented with two simple polygons. For solvable instances, we bound the complexity of polygons within a single square by a small constant and show that these bounds are tight. Complete proofs for omissions and proof sketches may be found in the appendix.

**Related work.** In theoretic research on drawing hypergraphs (e.g. [6,12]), the (often implicit) assumption is that the representations of two sets may cross at common vertices. These intersections and crossings are not deemed problematic as most visual encodings rely on the *nesting* of intersecting polygons—such as also present in the prototypical Venn and Euler diagrams [5]—to identify set memberships. Nesting, however, gives a strong visual cue of containment and may result in misleading visual representations implying containment relationships between hyperedges. Instead we focus on a visualization that relies on *disjoint membership* encoding. A rendering style closer to our approach is the technique of Kelp Diagrams [8], which does not use nesting either. KelpFusion [15] extended on the idea of Kelp Diagrams further reducing visual clutter.

One of the most well-established quality criterions of graph drawings is planarity (see e.g. [16,17]). When nested encodings are used, a planar drawing relates to finding a planar support [6]: a planar (regular) graph such that the vertices of each hyperedge induce a connected subgraph in the support. Deciding whether a planar support exists is possible for some simple support classes (see [6] for a discussion), but is already NP-hard for 2-outerplanar support graphs [6]. Hypergraph supports without planarity constraints but instead optimizing for total graph length is NP-hard, but approximation algorithms exist [1,11].

For the disjoint membership encoding, the problem is different since we cannot cross two set representations at a common element. It thus intuitively lies closer to the edge-based drawings [13] or the equivalent Zykov representation [21], for which notions of planarity follow readily from the standard notion for regular graphs. However, as is illustrated in Fig. 1, our disjoint encoding is stronger as it can visualize some hypergraphs that are not Zykov-planar, whereas any Zykov-planar hypergraph admits a drawing in our disjoint encoding. We can thus use elements to "pass in between" the representations of other sets, though not as flexibly as is allowed when searching for planar supports: the polygons must remain disjoint.

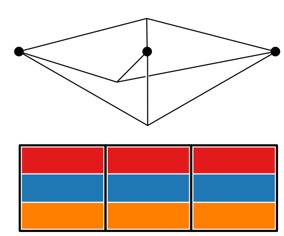


**Fig. 1.** A hypergraph that is not Zykov-planar (top) but has a disjoint-polygons drawing (bottom).

## 2 Preliminaries

**Colored grid.** We define a *$k$-colored grid $\Gamma$* as a unit-grid, in which each cell $s$ has a set of associated colors $\chi_s \subseteq \{1, \ldots, k\}$. A *fully* $k$-colored grid is the case where $\chi_s \neq \emptyset$ for all squares $s$. Throughout this paper, we primarily investigate 2-colored grids and use *colored graph* to refer to the 2-colored case, unless indicated otherwise. In the 2-colored case, we refer to the two colors as *red* ($r$) and *blue* ($b$); squares for which $\chi_s = \{r, b\}$ are called *purple* ($p$). Finally, squares without associated colors are *white*.

**Region.** A region is a maximal set of cells that have the same color assignment ($r$, $b$, or $p$) and where every cell $c$ in the region is connected via adjacent cells to every other cell $c'$ in the region. Cells are considered adjacent if they are horizontally or vertically adjacent.

**Tiling.** A tile $\tau_s$ for cell $s$ (with $\chi_s \neq \emptyset$) maps each color $c \in \chi_s$ to a (possibly disconnected) polygon $\tau_s(c)$ such that these partition the cell: that is, $\bigcup_{c \in \chi_s} \tau_s(c) = s$ and $\tau_s(c_1) \cap \tau_s(c_2) = \emptyset$ for colors $c_1 \neq c_2$. A *tiling* of a $k$-colored grid consists of a tile for each square with at least one associated color. We call a tiling *connected* if each color forms a connected polygon: that is, $\bigcup_{s \in \Gamma \wedge c \in \chi_s} \tau_s(c)$ is a connected polygon for each color $c \in \{1, \ldots, k\}$. For this definition of connected, two squares sharing only a corner are *not* connected. Our primary interest is in connected tilings: in the remainder, we use tiling to indicate a connected tiling.

## 3 Testing for a tiling

In this section we show how to test whether a 2-colored grid admits a tiling. As all completely red, blue, and white tiles are fixed, finding a tiling equals finding partitions of purple regions that ensure both the resulting red and blue polygon are connected. We model the 2-colored grid using two colored graphs, $G_r$ and $G_b$, representing the connectivity options for the red and blue polygon. Every intersection between and edge of $G_r$ and $G_b$ indicates a choice to connect either the blue or the red polygon through (part of a) purple region. Properties of these graphs allow us to test for admissibility of a tiling. We will prove that when these graphs are duals of each other, the corresponding grid $\Gamma$ admits a tiling.

We first show how to model simple purple regions (Section 3.1) as two colored graphs, and then extend it to general purple regions (Sections 3.2 and 3.3).

### 3.1 Purple regions with a simple adjacency structure

In this section we assume that each purple region is simply-connected and thus has no holes; we will remove this assumption in Section 3.3. The neighboring red and blue regions of a purple region $P$ form an ordered cyclic list as they appear along the boundary of $P$. Note that there can be duplicates in this list as the same red or blue region can touch $P$ multiple times. Let $\kappa(P)$ denote

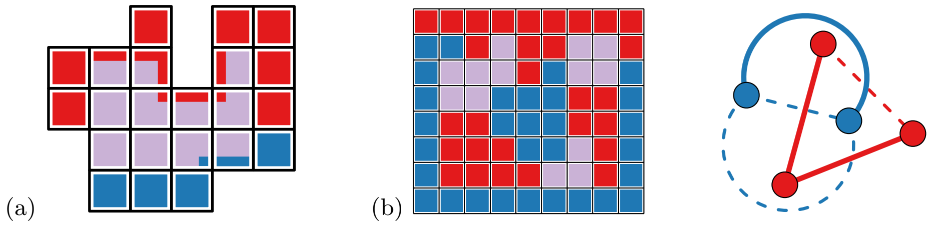



**Fig. 2.** A purple region with multiple adjacent blue and red regions.

the number of times the adjacent red and blue regions change their color in the cyclic order around $P$. Note, that $\kappa$ is always even. Consecutive (not necessarily distinct) neighboring regions of the same color can always be safely connected via the boundary of $P$ into a single connected region without restricting the connectivity options for the other color (see Fig. 2(a)).

Any purple region with $\kappa(P) = 2$ can easily be solved. W.l.o.g., assume these are handled, consecutive neighboring regions of the same color are connected, and any remaining purple region has at least four neighbors of alternating colors.

We construct the embedded graphs, $G_r$ and $G_b$, as follows: every red region is a vertex in $G_r$ and every blue region is a vertex in $G_b$. Assume, for now, that every purple region has exactly four (not necessarily distinct) neighbors (in Section 3.2 we will remove this assumption). For every purple region we create a red and a blue intersecting edge, the red edge connects the red vertices corresponding to the adjacent red regions, and the blue edge connects the corresponding blue vertices (see Fig. 2(b)). If the same red or blue region touches the purple region twice, the connecting edge will be a loop edge. Note that every blue edge intersects exactly one red edge, every red edge intersects exactly one blue edge, and the graphs $G_r$ and $G_b$ are plane by construction. Using the following lemma we prove the exact characteristic of graphs $G_r$ and $G_b$ for a 2-colored grid $\Gamma$ to admit a tiling.

**Lemma 1 ([9,20]).** *Let $G$ be a graph, $G^*$ its dual and $T$ a spanning tree of $G$. Let $T^*$ be the set of dual edges $\{e^* \mid e \notin T\}$. Then $T^*$ is a spanning tree of $G^*$.*

**Theorem 1.** *A 2-colored grid $\Gamma$ with simply-connected purple regions, where $\kappa(P) \leq 4$ for every purple region $P$, admits a tiling if and only if the corresponding $G_r$ and $G_b$ are each other's exact duals: there is exactly one blue vertex in every red face and there is exactly one red vertex in every blue face.*

*Proof (sketch).* We prove that if $\Gamma$ admits a tiling then graphs $G_r$ and $G_b$ are each other's duals using a counting argument. We count the number of edges needed to connect all red and blue regions, and use Euler's formula to show the number of red faces must be equal to the number of blue vertices, and vice versa.

The argument for the other direction immediately follows from Lemma 1. □

It immediately follows how we can find the connections through all grid-cells that result in a tiling. Since we need to have two dual graphs to find a tiling, Lemma 1 readily implies that we can pick any spanning tree on $G_b$ which then implies a spanning tree on $G_r$ (or the other way around).

### 3.2 Spiderweb gadgets

In the previous section we have shown how to solve the problem if each purple region has at most two adjacent blue and two adjacent red regions. In this section we will show how we can solve the problem when purple regions are still simply-connected but may have more adjacent red and blue regions.

As in the previous section, we construct graphs $G_r$ and $G_b$ by adding a vertex to $G_r$ for every red region, and a vertex to $G_b$ for every blue region. For every purple region we construct a *spiderweb* gadget and insert it into the graphs in such a way that Theorem 1 can be applied to find a solution.

Observe that given a purple region $P$, it is sufficient to know how regions adjacent to $P$ are connected to each other in a solution to generate a tiling for region $P$. We will show how to do this in Section 4. The problem can thus be reduced to finding the topology of the adjacent regions.

A spiderweb gadget $S$ of a purple region $P$ with $k$ red and $k$ blue alternating adjacent regions is defined as follows. It consists of $\lfloor k/2 \rfloor + 1$ levels (refer to Fig. 3). Each level consists of a cycle of $k$ vertices, with the exception of the outermost and the innermost levels. The outermost level has $k$ (blue) vertices without any edges between them, while the innermost level consists of only a single vertex. We label the levels from outside to inside by numbers from 0 to $\lfloor k/2 \rfloor$. On even levels the vertices are blue and are labeled with even numbers from 0 to $2k-2$ clockwise. On odd levels the vertices are red and are labeled with odd numbers 1 to $2k-1$ clockwise.

Each vertex of level $\ell$ with $0 \le \ell < \lfloor k/2 \rfloor - 2$ is connected to the vertex with the same label on level $\ell + 2$. The single vertex of level $\lfloor k/2 \rfloor$ is connected to all the vertices of level $\lfloor k/2 \rfloor - 2$. This gives us $2k$ paths starting from the outermost levels 0 or 1 to levels $\lfloor k/2 \rfloor - 1$ or $\lfloor k/2 \rfloor$, which we will call *spokes* and will refer to by the label of the corresponding vertices. We embed the resulting two connected components in such a way that they are each other's dual graph, by making sure that we get a proper clockwise numbering on the vertices of the two outermost levels (see Fig. 3). The vertices on levels 0 and 1 represent respectively the blue and red regions around the purple region $P$ and respect the adjacency order around $P$. If the same red or blue region has multiple points of adjacency with $P$, we merge the corresponding vertices of the spiderweb gadget so they are consistent with the topology of the nested neighboring regions of $P$.

In a spiderweb gadget $S$ corresponding to a purple region $P$ we define *bridging paths* that respect a proper tiling $T$ as follows: let $u$ and $v$ be two vertices in $S$ that represent two blue regions that are connected by a proper tiling $T$ through $P$. Assume that the clockwise distance from $u$ to $v$ is not greater than $k$, that is, if $u$ has label $x$ then $v$ has label $(x + 2i) \mod 2k$ for some $1 \le i \le \lfloor k/2 \rfloor$. To connect $u$ and $v$ with a bridging path, we start from $u$ on level 0, go down to level $2\lfloor (i+1)/2 \rfloor$ along the spoke $x$, take a shortest path within the level $2\lfloor (i+1)/2 \rfloor$ from the vertex with label $x$ to the vertex with label $(x+2i) \mod 2k$, and move up the spoke $(x + 2i) \mod 2k$ to vertex $v$ at level 0. If there are two possible shortest paths in level $2\lfloor (i+1)/2 \rfloor$, we take the clockwise path.

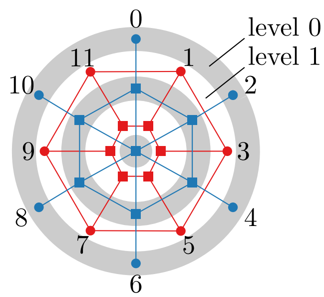



**Fig. 3.** Spiderweb gadget for $k = 6$: three blue levels with indices 0, 2, 4, and two red levels with indices 1, 3.

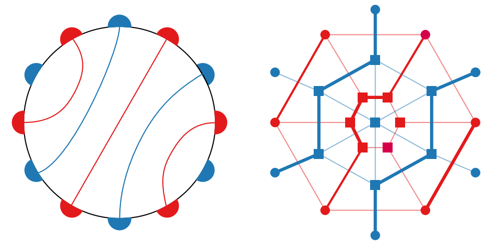


**Fig. 4.** A topological tiling and the corresponding paths through a spiderweb gadget.

The same kind of path can be constructed for a pair of red vertices, but starting from level 1, going to level $2\lfloor i/2 \rfloor + 1$, and moving back to level 1. We now show that connecting different blue and red regions using bridging paths within the spiderweb gadgets will result in blue trees and red trees, such that no pair of a blue and a red edge intersect (see Fig. 4 for an example).

By performing a case analysis on the possible red and blue pairs of adjacent regions to be connected, we can prove that the following lemma holds.

**Lemma 2.** *A red and a blue bridging paths in spiderweb gadget $S$, which respect a tiling $T$, can never intersect.*

We can now prove that a spiderweb gadget correctly represents a mixed region with adjacent red and blue regions. The following theorem can be proved using Lemma 2 similarly to the proof of Theorem 1.

**Theorem 2.** *A 2-colored grid $\Gamma$ with simply-connected purple regions admits a tiling if and only if the corresponding $G_r$ and $G_b$ are each other's exact duals: there is exactly one blue vertex in every red face and there is exactly one red vertex in every blue face.*

### 3.3 Higher-genus purple regions

A grid may also contain purple components of a higher genus (see Fig. 5(a)). We can reduce the genus of a component without affecting the solvability of the instance. For simplicity of explanation we assume a region of genus one. A *cross-annulus connection* $x$ has one endpoint $x_i$ on the inside of the annulus and one endpoint $x_o$ on the outside of the annulus. Let $C_S$ be the set of cross-annulus connections in a given solution $S$. We assume that the annulus is not degenerate and both inside and outside the annulus red and blue components exist.

**Lemma 3.** *If a solution $S$ exists with two adjacent cross-annulus connections $x, y$ of the same color, possibly separated by non-crossing connections, then there also exists a solution $S'$ where $C_{S'} = C_S \setminus \{y\}$.*

**Lemma 4.** *If there exists a solution $S$ with $|C_S| > 3$ and all cross-annulus connections are alternating in color, then there also exists a solution $S'$ with $|C_S| - 2$ cross-annulus connections.*

*Proof.* Let $u, v, x, y$ be four consecutive cross-annulus connections. W.l.o.g. assume $u$ and $x$ are red and $v$ and $y$ are blue. We remove $v$ and $x$ from the solution separating both the red and blue into two connected components. It cannot be the case that the connected components are both on the outside (/inside) of the annulus (see Fig. 5(b)). If so then both $u$ and $x$, as well as $v$ and $y$, would be connected through the inside of the annulus as there is no cross-annulus connection between $u$ and $y$. However, only one of these connections can exist as any connection from $v_i$ to $y_i$ separates $u_i$ and $x_i$. As the disconnected components for red and blue are on opposite sides of the annulus we can reconnect them to $u$ respectively $y$ without mutually interfering. A similar argument to Lemma 3 shows this can be done without invalidating the solution. □

**Corollary 1.** *If a solution exists, then a solution also exists that has at exactly one red and one blue connection across each annulus.*

**Lemma 5.** *If a solution exists with one red and one blue cross-annulus connection, then any cross-annulus placement of these connections is part of a solution.*

*Proof.* Let $x$ be the blue cross-annulus connection and $y$ the red cross-annulus connection. We show we can freely move each of the endpoints of the cross-annulus connection. W.l.o.g. assume that $x$ is not counter-clockwise adjacent to $y$ on the outer annulus ring. Let $c$ be the blue region that is counter-clockwise adjacent to $y_o$. Region $c$ may have several incoming blue connections that enter $c$ counter-clockwise along the annulus (see Fig. 6(a)). As the blue solution contains no cycles, removing each of these connections would break the solution in distinct connected components. Hence, we can rewire the blue connectors inside the annulus to connect these connected components in sorted order around the annulus, resulting in only one connection $c_b = (c', c)$ to $c$. Using a similar argument we can also rewire the red connectors crossing $c$ and ending at $y_o$ to ensure only one red connector $c_r$ crosses $c$.

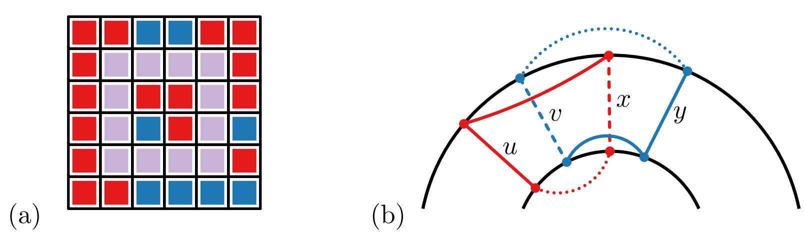



**Fig. 5.** (a) An annulus-type purple region with adjacent blue and red regions both on the outside and the inside. (b) By adding edges $(x_1, v_1)$ and $(y_2, u_2)$ we reconnect the disconnected components formed be removing cross-annulus connections $x$ and $y$.

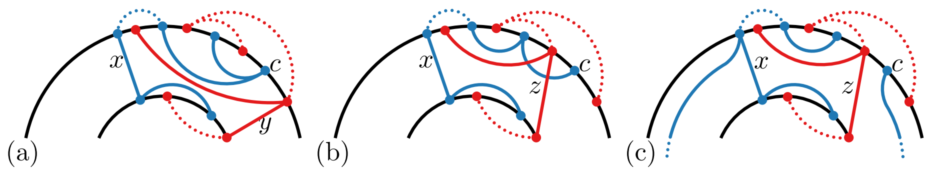



**Fig. 6.** (a) Initial configuration with several connectors crossing $z_o$. (b) Rewiring the red and blue connectors ensures only one red and blue connector crosses $z_o$. (c) Introducing $z$ and rerouting the intersecting red connectors results in only one blue connector that still causes intersections. (d) As the blue disconnected component cannot be enclosed by the newly created red connector, we can always connect it back to $x$.

Remove $y$ and insert a new red cross-annulus connector $z = (y_i, z_o)$ where $z_o$ is one region counter-clockwise from $y_o$. The connector $z$ can only intersect $c_r$ and $c_b$. Removing $c_r$ results in two red connected components, one of which contains $z$. As $c_r$ intersected $z$, it is possible to connect the disconnected component to $z$ while only intersecting $c_b$ (see Fig. 6(b)). Removing $c_b$ results in two blue connected components, one of which contains $x$. We prove that the disconnected component are separated by $x$ and $z$.

W.l.o.g assume that the disconnected components for red and blue are both clockwise between $x$ and $z$. Region $z_0$ must be connected to $y_o$ through the outside as $c_b$ blocked any connection from the inside. Region $c'$ cannot be connected clockwise to $x_0$ due to $z$ and it cannot be connected counter-clockwise to $x_0$ as it must be disconnected. Thus $c'$ must be connected to $x_0$ via the outside of the annulus. This cannot both be true and thus $x$ and $z$ are not both clockwise between $x$ and $z$. The components are thus separated by $x$ and $z$.

Therefore we can safely reconnect the disconnected blue component through the annulus to $x$ (see Fig. 6(c)). Repeatedly moving the end-point of one of the cross-annulus connections allows the creation of any configuration of the two red and blue cross-annulus connections without invalidating the solution. □

**Theorem 3.** *If a solution exists, then there also exists a solution with exactly one red and one blue cross-annulus connection starting from any two regions on the inner annulus and connecting to any two regions on the outer annulus.*

From Theorem 3 it follows that we can cut the annulus open to reduce the genus of the purple region by one without changing the solvability of the problem. After cutting open all the annuli and replacing all purple regions with spiderweb gadgets, we can then solve the problem using Theorem 1.

## 4 Optimizing tiles

As shown, not all colored grids admit any tiling. Here we investigate the design of the tiles themselves, based on the assumption that some tiling is possible as tested in the previous section. To this end, we define the complexity of a tile as the number of connected components, see Fig. 7. The complexity of a tiling is the maximal complexity of any of its tiles. We use the notation $t$-tile ($t$-tiling) to denote a tile (tiling) of complexity $t$.

Assuming some tiling exists, we prove in this section that a 5-tiling is sufficient in general and a 2-tiling is even sufficient if there are no white cells.

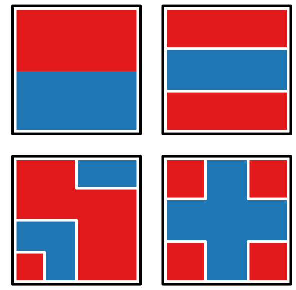


**Fig. 7.** Tiles with complexity 2, 3, 4 and 5.

### 4.1 Ensuring a 5-tiling

We prove that a valid tiling for a colored grid (possibly including white cells) can be redrawn to include no more than three colored maximal segments along each side of a tile. Using this we show that a 5-tiling can always be achieved.

**Lemma 6.** *If a 2-colored grid admits a tiling, then it admits a tiling where each tile has at most 3 intervals of alternating red and blue along each side of the tile.*

*Proof (sketch).* Assume that a tile $t$ has at least 4 intervals of alternating red and blue on the left-side of $t$. As the tiling is valid, both blue (/red) intervals are connected in the tiling. For each interval we identify whether the path exiting or entering $t$ connects to the other interval of the same color (see Fig. 8(a)). It cannot be the case that the red and blue path leave the border of $t$ in the same direction at both the middle two intervals (see Fig. 8(b)). To reduce the number of intervals, we recolor the interval by shortcutting both the blue and the red component inside $t$ (see Fig. 8(c)). □

**Theorem 4.** *If a general 2-colored grid admits a tiling, then it admits a 5-tiling.*

*Proof (sketch).* Using Lemma 6 we know that there are at most three alternatingly colored intervals along each side of $t$. The intervals at the corners must have the same colors. If we have more than five connected components, any

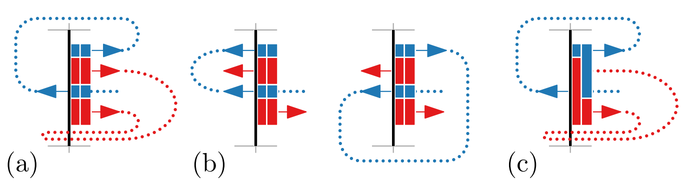



**Fig. 8.** (a) Identify per interval if the path connecting to the other interval of the same color enters or exits $t$. (b) When the connecting path from the middle blue interval exits $t$, it separates the left side of the top red interval from the bottom red interval. (c) Short-cutting inside the cell reduces the number of maximal colored segments without affecting the rest of the solution.

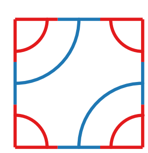
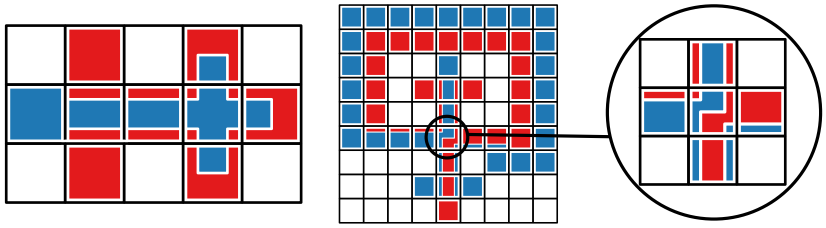


**Fig. 9.** Connected components in a tile with more than five connected components.

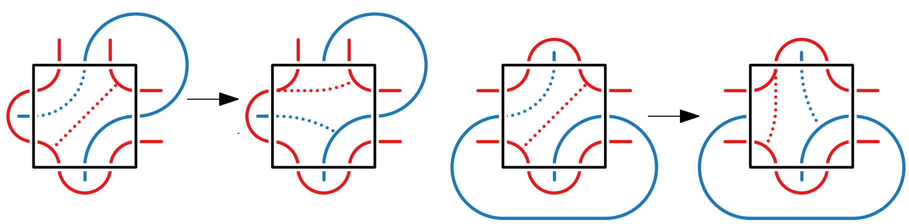


(a) (b)

**Fig. 10.** (a) Colored grid requiring a 5-tiling. (b) Colored grid requiring a 4-tiling with two connected components of both colors in the same cell.

component that only has one continuous interval along a single side of $T$ can be removed without affecting the solution. Thus, each remaining connected component can be identified with at least two continuous intervals along the border of $t$. To maximize the number of connected components we need to connect all corners and pairs of adjacent blue connected intervals along the border (see Fig 9). The maximum achievable number of connected components is then five. □

The upper-bound on tile complexity is tight as a 5-tile may be required when the grid includes white cells (see Fig. 10(a)). There does not, however, exist a 5-tiling that has at least two connected components of either color. As is clear from Theorem 4 there is only one other alternative to create a 5-tiling. This 5-tiling only has two legal configurations and both configuration can be simplified to a 4-tiling (see Fig. 11). A 4-tile containing two connected components of either color is possible though (see Fig. 10(b)).



(a) (b)

**Fig. 11.** (a-b) There are two configurations for a 5-tile where both colors have at least two connected components. Both possible configuration can be simplified to a 4-tile.

### 4.2 Ensuring a 2-tiling

We show that a fully 2-colored grid (without white cells) even admits a 2-tiling, provided it admits any tiling. As an intermediate step, we first prove that a tiling exists that only using one blue component in any tile.

**Lemma 7.** *If a fully 2-colored grid admits a tiling, then it admits a tiling in which each tile has at most one blue connected component.*

*Proof (sketch).* Since the grid admits a tiling, we need to only retile every purple region as to ensure that the lemma holds. To do so, we treat them one by one, creating a spanning forest over the purple squares that connects the square-centers of adjacent squares. This ensures that each tile has exactly one blue connected component inside, but it may disconnect the blue region. However, since we know that a tiling exists and the current solution maps to some forest in $G_b$, we know by Lemma 1 that its dual $G_r$ has a cycle around each tree in the forest. Hence, we can add arbitrary connections between unconnected blue components to make it all connected again, without disconnecting the red. □

The above construction relies on the alternation of the blue and red components along the boundary of $P$. As there are no white cells we can guarantee this alternating pattern after connecting all blue (and red) corners. Indeed, the higher complexity with white cells is caused by long connections along a purple region's boundary that are needed to achieve this alternating pattern for a partially colored grid (e.g., Fig. 10).

**Theorem 5.** *If a fully* 2*-colored grid admits a tiling, then it admits a* 2*-tiling.*

*Proof (sketch).* Since the fully 2-colored grid admits a tiling, Lemma 7 implies that there is a tiling $T$ where every purple tile only has a single blue component. After removing red components that only connects to one neighboring tile and recoloring a red corners to blue if the adjacent grid vertex has 4 red corners surrounding it, the tile is in one of the following four cases.

1 There are two red corners $r_1$ and $r_2$ on the same side of the cell. There is a path through the grid connecting the red corners where the path exits the current cell via the same side or adjacent sides. (see Fig. 12(a)).
2 There are two red corners $r_1$ and $r_2$ on the same side of the cell. The connecting path exits the cell via opposite sides of the cell (see Fig. 12(b)). Blue connects only downwards in the cell below.
3 There are two red corners $r_1$ and $r_2$ that do not share a common side of the cell. In this case the other corners are blue, otherwise one of the two previous cases applies (see Fig. 13). Furthermore, either $p_1$ or $p_2$ is blue.

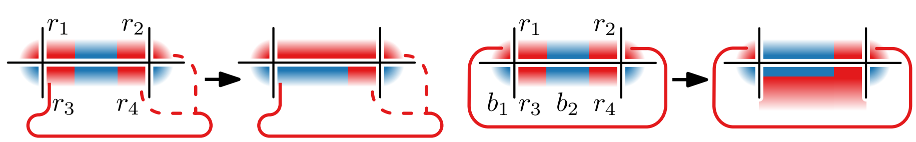



**Fig. 12.** (a) Two red corners along the same cell side connected via adjacent (or the same) sides of the cell. The tile complexity can be reduced by connecting the red corners. (b) Two red corners along the same cell side connected via opposite sides. Tile complexity can be reduced by connecting both red corners via one side.

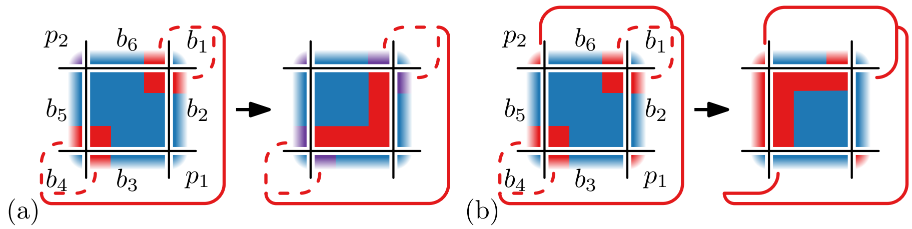



**Fig. 13.** Two diagonally positioned red corners. The complexity of the tile can be reduced by introducing a red $L$-shape that connects all the red. (a) Reduced complexity if either $p_1$ or $p_2$ was blue. (b) Reduced complexity if both $p_1$ and $p_2$ were red.

4 There are two red corners $r_1$ and $r_2$ that do not share a common side of the cell. Furthermore, both $p_1$ and $p_2$ are red.

We then use a set of reduction rules which are illustrated in Fig. 12 and Fig. 13 to reduce the complexity of the tile. The reductions rules and their proofs are contained in the appendix. Repeated application of the reduction rules, interlaced with the reduction of the number of red components in a cell, results in a 2-tiling. □

## 5 Conclusion

We took the first steps towards investigating a disjoint-polygons representation for visualizing set memberships (hypergraphs). We investigated the 2-color version in which each element is positioned as a cell in a (unit-)grid. We showed how to test whether a disjoint-polygons representation is possible for a given 2-colored grid. Moreover, we proved that if such a representation is possible, then we can also bound the complexity of the corresponding "tiles" (the coloring of a single cell). Each tile requires at most five connected colored regions, and even only two regions are required when no white cells are present in the grid.

**Future work.** Clearly, there are still myriad options for further exploration.We have not touched upon variants with more colors: does our approach readily generalize? Considering the restrictions already in the studied 2-color variant, introducing more colors may cause many practical scenarios to have no solution.

If we allow the rearrangement of elements, the 2-color variant becomes trivial: any hypergraph with two hyperedges is representable. As the disjoint-polygons encoding can represent more than the edge-based drawings but cannot represent all planar supports, this question remains interesting for multiple colors (sets).

Finally we may consider the situation where some cells have no assigned set of colors but may be tiled using any subset of the colors.

**Acknowledgments.** The authors would like to thank Jason Dykes for fruitful discussions at an early stage of this research.

## References


1. H. Akitaya, M. Löffler, and C. Tóth. Multi-Colored Spanning Graphs. In *Int. Symp. on Graph Drawing and Network Visualization*, pages 81–93. Springer, 2016.
2. B. Alper, N. Henry Riche, G. Ramos, and M. Czerwinski. Design study of LineSets, a novel set visualization technique. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2259–2267, 2011.
3. B. Alsallakh, W. Aigner, S. Miksch, and H. Hauser. Radial Sets: Interactive visual analysis of large overlapping sets. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2496–2505, 2013.
4. B. Alsallakh, L. Micallef, W. Aigner, H. Hauser, S. Miksch, and P. Rodgers. The state of the art of set visualization. *CGF*, 35(1):234–260, 2016.
5. M. Baron. A note on the historical development of logic diagrams: Leibniz, euler and venn. *Mathematical Gazette*, 53(384):113–125, 1969.
6. K. Buchin, M. van Kreveld, H. Meijer, B. Speckmann, and K. Verbeek. On planar supports for hypergraphs. *J. of Graph Algorithms and Appl.*, 15(4):533–549, 2011.
7. C. Collins, G. Penn, and S. Carpendale. Bubble Sets: Revealing set relations with isocontours over existing visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1009–1016, 2009.
8. K. Dinkla, M. van Kreveld, B. Speckmann, and M. Westenberg. Kelp Diagrams: Point set membership visualization. *CGF*, 31(3pt1):875–884, 2012.
9. D. Eppstein, G. Italiano, R. Tamassia, R. Tarjan, J. Westbrook, and M. Yung. Maintenance of a minimum spanning forest in a dynamic plane graph. *J. of Algorithms*, 13(1):33–54, 1992.
10. N. Henry Riche and T. Dwyer. Untangling Euler Diagrams. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1090–1099, 2010.
11. F. Hurtado, M. Korman, M. van Kreveld, M. Löffler, V. Sacristán, R. Silveira, and B. Speckmann. Colored Spanning Graphs for Set Visualization. In *Int. Symp. on Graph Drawing*, pages 280–291. Springer, 2013.
12. M. Kaufmann, M. van Kreveld, and B. Speckmann. Subdivision drawings of hypergraphs. In *Proc. of the 16th Int. Symp. on Graph Drawing (GD 2008)*, LNCS 5417, pages 396–407, 2009.
13. E. Mäkinen. How to draw a hypergraph. *Int. J. of Computer Mathematics*, 34:177–185, 1990.
14. W. Meulemans, J. Dykes, A. Slingsby, C. Turkay, and J. Wood. Small multiples with gaps. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):381–390, 2017.
15. W. Meulemans, N. Henry Riche, B. Speckmann, B. Alper, and T. Dwyer. KelpFusion: A hybrid set visualization technique. *IEEE Transactions on Visualization and Computer Graphics*, 19(11):1846–1858, 2013.
16. H. Purchase. Metrics for graph drawing aesthetics. *J. of Visual Languages & Computing*, 13(5):501–516, 2002.
17. H. Purchase, R. Cohen, and M. James. On planar supports for hypergraphs. In *Proc. of the Symp. on Graph Drawing (GD 1995)*, LNCS 1027, pages 435–446, 1996.
18. P. Simonetto and D. Auber. Visualise undrawable euler diagrams. In *Proc. of the 12th Conference on Information Visualisation*, pages 594–599, 2008.
19. P. Simonetto, D. Auber, and D. Archambault. Fully automatic visualisation of overlapping sets. *CGF*, 28(3):967–974, 2009.
20. W. Tutte. *Graph Theory.* Addison-Wesley, Menlo Park, CA, 1984.

21. T. Walsh. Hypermaps versus bipartite maps. *J. of Comb. Theory*, 18:155–163, 1975.
22. J. Wood and J. Dykes. Spatially ordered treemaps. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1348–1355, 2008.

## A Full proofs

**Theorem 1.** *A 2-colored grid $\Gamma$ with simply-connected purple regions, where $\kappa(P) \leq 4$ for every purple region $P$, admits a tiling if and only if the corresponding $G_r$ and $G_b$ are each other's exact duals: there is exactly one blue vertex in every red face and there is exactly one red vertex in every blue face.*

*Proof.* Suppose $\Gamma$ admits a tiling $T$. Then there exists a tiling $T'$ where, for any purple region $P$, all consecutive neighboring regions of the same color are connected through the boundary of $P$. Let $n_r$ and $n_b$ denote the the number of vertices in $G_r$ (red regions) and in $G_b$ (blue regions) respectively. We observe that the number of purple regions with $\kappa = 4$, the number of edges in $G_r$, the number of edges in $G_b$, and the number of intersections is the same, say, $e$. By construction, $n_b \geq f_r$ and $n_r \geq f_b$. The purple regions with $\kappa = 4$ can be tiled to connect two adjacent red regions or two adjacent blue regions but never both. As there are $n_r$ red regions, at least $n_r - 1$ red edges are needed to connect all red vertices (regions). And similarly, at least $n_b - 1$ edges are needed to connect all blue vertices (regions). Thus, if $\Gamma$ admits tiling $T'$ then $e \geq n_r + n_b - 2$.

On the other hand, by Euler's formula, $n_r - e + f_r = 2$ and $n_b - e + f_b = 2$. Combining these equations and $n_b \geq f_r$, $n_r \geq f_b$, we derive that $e \leq n_r + n_b - 2$. Thus, $e = n_r + n_b - 2$, and $n_b = f_r$ and $n_r = f_b$. As each red edge intersects exactly one blue edge and vice versa, graphs $G_r$ and $G_b$ are each other's duals.

For the other direction, assume that $G_r$ and $G_b$ are dual graphs. By Lemma 1 there exist two non-intersecting spanning trees of $G_r$ and $G_b$ that will specify which adjacent regions of every purple region are to be connected to form connected red and blue polygons. Assuming that a tiling can be constructed, given the connectivity information for every purple region (which will be shown in Section 4), a 2-colored grid admits a tiling if and only if $G_r$ and $G_b$ are dual graphs. ◻

**Lemma 2.** *A red and a blue bridging paths in spiderweb gadget $S$, which respect a tiling $T$, can never intersect.*

*Proof.* Let $u$ and $v$ be two vertices that represent two blue regions that are connected in $T$. Let $u$ be on level 0 of spoke $x$ and $v$ be on level 0 of spoke $(x+2i) \bmod 2k$ for some $1 \leq i \leq \lfloor k/2 \rfloor$. We connect $u$ and $v$ by a bridging path as described above, which will go down to level $2\lfloor (i+1)/2 \rfloor$. Any connection of two red regions in $T$ through the purple region $P$ cannot cross the connection between the regions that $u$ and $v$ represent in $T$. This means that the corresponding vertices must both lie on the same side of the bridging path between $u$ and $v$. Denote these vertices as $u'$ and $v'$, and let $u'$ have label $y$ and $v'$ have label $(y + 2j) \mod 2k$ for some $1 \leq j \leq \lfloor k/2 \rfloor$. The bridging path goes down to level $2\lfloor j/2 \rfloor + 1$. There can be several cases:

- When $u'$ and $v'$ lie in between $u$ and $v$ moving in the clockwise order, we have that $j < i$. Thus, the two bridging paths cannot intersect, as the path from $u'$ to $v'$ goes to level $2\lfloor j/2 \rfloor + 1 < 2\lfloor (i + 1)/2 \rfloor$.

- When $u'$ and $v'$ lie in between $v$ and $u$ moving in the clockwise order, we have that either $j > i$ or the two bridging paths lie on the opposite sides of the vertex on the innermost level. In the later case the two bridging paths cannot intersect as they are separated by the innermost level, and in the former case, the two bridging paths cannot intersect as the path from $u'$ to $v'$ goes to level $2\lfloor j/2 \rfloor + 1 > 2\lfloor (i+1)/2 \rfloor$. $\square$

**Theorem 2.** *A 2-colored grid $\Gamma$ with simply-connected purple regions admits a tiling if and only if the corresponding $G_r$ and $G_b$ are each other's exact duals: there is exactly one blue vertex in every red face and there is exactly one red vertex in every blue face.*

*Proof.* The proof is similar to the proof of Theorem 1. Suppose $\Gamma$ admits a tiling $T$. Then there exists a tiling $T'$ where, for any purple region $P$, all consecutive neighboring regions of the same color are connected through the boundary of $P$. We can find non-intersecting trees in every spiderweb gadget $S$ corresponding to a purple region $P$ that connect all pairs of vertices of levels 0 and 1 in the same way as $T$. First, we create the bridging paths from Lemma 2 for every pair of regions of the same color connected in $T$ through $P$. The bridging paths do not create cycles in $S$, thus, every vertex in $S$ that is still not connected to the levels 0 or 1 can be connected to them by growing non-intersecting spanning forests from the vertices on the levels 0 and 1.

For the other direction, assume that $G_r$ and $G_b$ are dual graphs. By Lemma 1 there exist two non-intersecting spanning trees of $G_r$ and $G_b$. These spanning trees provide the decisions of which adjacent regions of every purple region to connect, and the topology of the connections. Assuming that a tiling can be constructed, given this information (which will be shown in Section 4), and because every spiderweb gadget consists of dual graphs (if we identify the outermost level to a single blue vertex), a 2-colored grid admits a tiling if and only if $G_r$ and $G_b$ are dual graphs. $\square$

**Lemma 3.** *If a solution $S$ exists with two adjacent cross-annulus connections $x, y$ of the same color, possibly separated by non-crossing connections, then there also exists a solution $S'$ where $C_{S'} = C_S \setminus \{y\}$.*

*Proof.* Removing $y$ from the solution causes $y_i$ and $y_o$ to be in different connected components. W.l.o.g. assume $y_o$ is in the same connected component as $x$. We connect $y_i$ to $x_i$. As $x$ and $y$ were adjacent cross-annulus connections, any connections already present between $x$ and $y$ cannot cross the annulus and hence we can safely connect $x$ and $y_2$ (see Fig. 5(b)). $\square$

**Corollary 1.** *If a solution exists, then a solution also exists that has at exactly one red and one blue connection across each annulus.*

*Proof.* By Lemma 4 we can find a solution with at most three cross-annulus connections. By Lemma 3 a solution with three cross-annulus connections can be reduced to a solution with two cross-annulus connections.

**Theorem 4.** *If a general 2-colored grid admits a tiling, then it admits a 5-tiling.*

*Proof.* Assume there exists a tile $t$ that has at least six connected components. Using Lemma 6 we know that there are at most three alternatingly colored intervals along each side of $t$. If there are intervals of different colors surrounding a corner, then we extend one interval on the inside of $t$ around the corner to get four intervals on one side and apply Lemma 6 to reduce back to at most three intervals. After doing this, both intervals surrounding the corner of $t$ must be the same color as there must be at least an epsilon part connected on the inside of $T$. W.l.o.g. we assume all corners are red. As we have more than five connected components, any component that only has one continuous interval along a single side of $T$ can be removed without affecting the solution. Removing this connected component reduces the number of connected components by at most one and consequently we still have at least one connected component of either color in $t$. Each remaining connected component can be identified with at least two continuous intervals along the border of $t$. Furthermore, as all corners must be connected and the blue connected components must be separated we must connect the blue intervals that are adjacent along the border in the resulting subdivision (see Fig 9). However there are two blue components and two red components that are not yet separated and we can separate at most one pair. Hence, the maximum number of connected components possible is five. □

**Lemma 6.** *If a 2-colored grid admits a tiling, then it admits a tiling where each tile has at most 3 intervals of alternating red and blue along each side of the tile.*

*Proof.* Assume that a tile $t$ has at least 4 intervals of alternating red and blue on the left-side of $t$. We consider the top-most four intervals that are inside $t$ and adjacent to the left-side of $t$ and we assume these intervals are ordered blue, red, blue, red. As the tiling is valid, both blue (/red) intervals are connected in the tiling and since there can be no cycles there exists exactly one path connecting them. For each interval we identify whether the path exiting or entering $t$ connects to the other interval of the same color (see Fig. 8(a)). It cannot be the case that the red and blue path leave the border of $t$ in the same direction at both the middle two intervals (see Fig. 8(b)). Assume w.l.o.g. that the middle blue connecting path exits $t$. Any path connecting the blue intervals must separate the left side of the top red interval from the bottom red interval. Hence the connecting path from the middle red interval cannot also exit $t$.

To reduce the number of intervals, we recolor the interval of the color whose connecting path exits the current tile (blue in Fig. 8(a)) to the other color along the boundary of $t$. To keep blue connected and remove the created red cycle, we move the other blue interval an epsilon distance inside $t$ and stretch it over the middle red interval. (see Fig. 8(c)). This reduces the number of intervals on the boundary of $t$ by two and can be repeated as necessary without affecting the validity of the solution. □

**Lemma 7.** *If a fully 2-colored grid admits a tiling, then it admits a tiling in which each tile has at most one blue connected component.*

*Proof.* Let $T$ be a tiling admitted by a fully 2-colored grid $\Gamma$. Any tile in $T$ that is uni-colored trivially satisfies our lemma. Hence, we consider a purple region $P$ and show how to retile it to ensure each tile has exactly one blue connected component. Let $T'$ be the retiled solution, which is identical to $T$ except for the retiled cells in $P$.

As there are no white cells, the blue and red components incident to $P$ must alternate around its boundary. Exceptions are the convex corners of $P$. At the convex corners there may be two blue cells adjacent to the same purple cell, separated by a red or purple cell (see Fig. 14(a) top-left). As $\Gamma$ admits a tiling these blue components must be connected through $P$. If both components are connected through the outside of $P$ then they isolate the red/purple corner cell from $P$. We refer to these connections through $P$ as the blue corners of $P$.

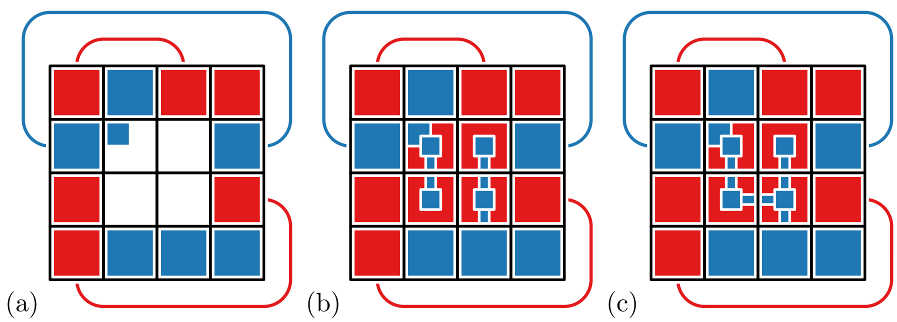



**Fig. 14.** A purple region of four cells to be retiled, including the direct neighborhood and topological connections in the original tiling $T$ outside $P$. (a) The top-left blue corner must be present. (b) Construction of the spanning forest. (c) Connecting two blue components breaks a cycle in the red component.

We construct a new tiling for $P$ as follows. First we place a small blue rectangle at the center of every cell in $P$. We then reconnect the blue corners of $P$ by filling the appropriate corners from this center rectangle. On these centers we build a spanning forest by connecting the centers of adjacent cells. This spanning forest is built such that each tree in the forest is rooted in a distinct blue component (see Fig. 14(b)). By construction, all red in $P$ forms a single component and thus red is connected. If there was only one blue component, then we are done as blue is connected and forms a single component inside every cell in $P$.

Assume there are two or more blue components. Since we know that a tiling exists, any blue component is separated from the other components by a red cycle. This is most easily seen in the constructed graphs $G_b$ and $G_r$ (see Theorem 1): our constructed forest maps to a forest in $G_b$ and by Lemma 1 we can arbitrarily complete this into a tree while ensuring that $G_r$ has a tree as well. Hence, we can pick two arbitrary components that have adjacent cells in $P$ and connect these (see Fig. 14(c)). This reduces the number of blue components by one while keeping the red connected as the connection only cuts a red cycle.

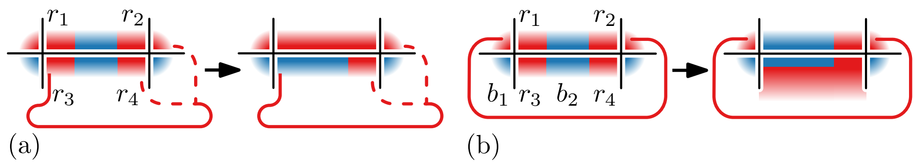



**Fig. 15.** (a) Two red corners along the same cell side connected via adjacent (or the same) sides of the cell. The tile complexity can be reduced by connecting the red corners. (b) Two red corners along the same cell side connected via opposite sides. Tile complexity can be reduced by connecting both red corners via one side.

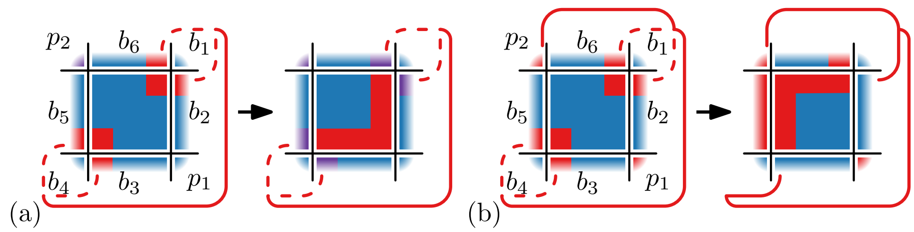



**Fig. 16.** Two diagonally positioned red corners. The complexity of the tile can be reduced by introducing a red $L$-shape that connects all the red. (a) Reduced complexity if either $p_1$ or $p_2$ was blue. (b) Reduced complexity if both $p_1$ and $p_2$ were red.

We repeat the above for every purple region, to ensure the eventual tiling has a single component for blue in every cell. □

**Theorem 5.** *If a fully* 2*-colored grid admits a tiling, then it admits a* 2*-tiling.*

*Proof.* If there is a solution for a fully 2-colored grid $\Gamma$, then by Lemma 7 there is a solution $T$ where every purple cell only has a single blue component. That is, every purple cell only has red along the sides or corners. All red components must include at least one corner by construction of Lemma 7. If a cell has a red component that only connects to one neighboring cell then we retract it such that it is only adjacent to a single corner. While a cell has more than one red component we can fully remove any such red component from the cell.

When four adjacent cells have a red corner around the same grid vertex, then we can color one of them blue, if it is the corner of a purple cell with another distinct red polygon in it. After repeated application of the above, any purple cell with multiple red components is in one of four cases:

1 There are two red corners $r_1$ and $r_2$ on the same side of the cell. There is a path through the grid connecting the red corners where the path exits the current cell via the same side or adjacent sides. (see Fig. 12(a)).
2 There are two red corners $r_1$ and $r_2$ on the same side of the cell. The connecting path exits the cell via opposite sides of the cell (see Fig. 12(b)). Blue connects only downwards in the cell below.

3 There are two red corners $r_1$ and $r_2$ that do not share a common side of the cell. In this case the other corners are blue, otherwise one of the two previous cases applies (see Fig. 13). Furthermore, either $p_1$ or $p_2$ is blue.
4 There are two red corners $r_1$ and $r_2$ that do not share a common side of the cell. Furthermore, both $p_1$ and $p_2$ are red.

We can reduce the complexity of each tile the following reduction rules:

1 Connect $r_1$ and $r_2$, and remove $r_3$ to break the red cycle (see Fig. 12(a)).
2 Connect $r_3$ and $r_4$, and connect $b_1$ and $b_2$ between $r_1$ and $r_3$. We remove $r_1$ as it has become useless by connecting $b_1$ and $b_2$. Similarly, we fully color the rest of the bottom cell red (see Fig. 12(b)). Note that the only option is for the blue component in the bottom cell to connect solely to the cell below it. Assume that it also connected to the left cell. Then the left red component must also connect to the left cell. When this red component connects two different sides of this cell then it envelopes the blue corner and this cell cannot have anymore blue below. Contradiction. If the red component ended in this cell, then we would not have removed the red in the top-right corner of this cell. Thus this situation can also not occur.
3 If either $p_1$ or $p_2$ is blue, we connect $r_1$ and $r_2$ along this respective side of the cell (see Fig. 13(a)). We connect either $b_1$ or $b_4$ to an adjacent blue region to break the red cycle and create a single blue component again.
4 When both $p_1$ and $p_2$ are red, we connect $r_1$ and $r_2$ along the side of the cell that is on the opposite side of the corner that the connection between both red regions encloses (we pass $p_2$ in Fig. 13(b)). This results in three blue connected components. Assume w.l.o.g. that the red corners are connected via a path leaving through the bottom and entering from the right. We join the blue components by connecting $b_1$ and $b_2$, and $b_3$ and $b_4$, separating both sides of the connecting path. We also recolor the border along the side of the cell that is not enclosed by the red connection between the top-left red region and the other red regions ($b_5$ in this case) to ensure connectivity of the red.

Repeated application of the above reduction rules, interlaced with the reduction of the number of red components in a cell, must result in a 2-tiling. Every cell still has a single blue component as no rules increases the number of blue components in a cell. Furthermore, as none of the above rules applies anymore, every cell has a single red polygon as well. Thus, the tiling is a 2-tiling. □