Reconfiguration of Polygonal Subdivisions via Recombination

Hugo A. Akitaya
University of Massachusetts Lowell, MA, USA

Andrei Gonczi
Tufts University, Medford, MA, USA

Diane L. Souvaine
Tufts University, Medford, MA, USA

Csaba D. Tóth
California State University Northridge, Los Angeles, CA, USA
Tufts University, Medford, MA, USA

Thomas Weighill
University of North Carolina Greensboro, NC, USA

Abstract

Motivated by the problem of redistricting, we study area-preserving reconfigurations of connected subdivisions of a simple polygon. A connected subdivision of a polygon $\mathcal{R}$, called a *district map*, is a set of interior disjoint connected polygons called *districts* whose union equals $\mathcal{R}$. We consider the *recombination* as the reconfiguration move which takes a subdivision and produces another by merging two adjacent districts, and by splitting them into two connected polygons of the same area as the original districts. The *complexity* of a map is the number of vertices in the boundaries of its districts. Given two maps with k districts, with complexity $O(n)$, and a perfect matching between districts of the same area in the two maps, we show constructively that $(\log n)^{O(\log k)}$ recombination moves are sufficient to reconfigure one into the other. We also show that $\Omega(\log n)$ recombination moves are sometimes necessary even when $k = 3$, thus providing a tight bound when $k = 3$.

2012 ACM Subject Classification Social and professional topics → Social engineering attacks; Theory of computation → Computational geometry; Theory of computation → Nonconvex optimization

Keywords and phrases configuration space, gerrymandering, polygonal subdivision, recombination

Digital Object Identifier 10.4230/LIPIcs.ESA.2023.6

Related Version *Full Version*: https://arxiv.org/abs/2307.00704

Funding *Csaba D. Tóth*: Research funded in part by NSF awards DMS-1800734 and DMS-2154347.
Thomas Weighill: Research funded in part by NSF award OIA-1937095.

1 Introduction

We consider the problem of *redistricting* – the partition of a geographic domain into disjoint districts. In particular, we consider the case when these districts are required to be connected and of roughly equal population. These criteria are typically enforced in political redistricting, wherein each district elects one or more representatives to serve on a governing body, a canonical example being Congressional districts in the United States. Even under these restrictions, the space of possible redistricting plans for a typical domain is intractably vast, making it difficult to sample from this space. Recently, algorithms for generating large samples of plans have made it possible to find the neutral baseline for a particular state, which in turn can be used to detect and describe gerrymanders (i.e., unfair maps) [9, 8, 10, 13, 14, 17, 18].

The most common and successful sampling algorithms for redistricting are Markov chains that perform a sequence of reconfiguration moves on an initial map. The most prominent reconfiguration move is the *recombination* or *ReCom* move (see Figure 1), which is a move that modifies two adjacent districts while maintaining population balance and connectivity [13, 14]. In order to properly sample from the space of redistricting plans, we should require that any feasible redistricting plan can be reached from the initial map by a finite sequence of ReCom moves. That is, we want to positively answer the *reachability* question for this reconfiguration move; in the language of Markov chains this would be to prove that any chain built on the ReCom move is *irreducibile*.

Historically, most redistricting algorithms have operated on a discretized version of the geographic domain. In this framework, a district map is modeled as a vertex partition of an adjacency graph [14, 24]. This is natural since population data is only available at the level of fixed geographic units, such as Census blocks in the case of the United States. The ReCom algorithm fits within this framework, and current versions all use a spanning tree method on the adjacency graph to perform the ReCom move. Unfortunately, it is easy to construct small pathological examples of graphs for which ReCom reachability fails. Moreover, even determining whether two plans can be connected via a sequence of ReCom moves is PSPACE-complete [3] for general (planar) graphs.

A reasonable but unproven hypothesis is that for real-world adjacency graphs representing sufficiently fine discretizations of the geographic domain, we will indeed have reachability. A general theorem covering all adjacency graphs of interest seems beyond reach, which has led to a search for intermediate results. One direction of investigation is to allow a large class of graphs but relax the population balance constraint considerably; in such cases theoretical results are possible [2, 3] (see Related Work below). Reachability on grid graphs or triangular lattices is an active area of research but as of yet without concrete results.

In this paper, we return to the original hypothesis – that sufficient discretization leads to reachability – to motivate our result. Instead of modeling redistricting plans as graph partitions, we adopt a continuous model where the districts are connected polygons of equal population which partition a polygonal domain. Note that sampling algorithms based on this model do exist in the literature, most notably the power diagram method in [11], but these algorithms are not Markov chains and require an extra refinement step to go from polygonal partitions to partitions that respect the geographic units.

In our continuous model, we are able to establish reachability for the ReCom move – that is, any two polygonal partitions can be connected by finitely many ReCom steps that merge and resplit adjacent polygonal districts. The implication is that given two real-world redistricting plans, a sufficiently fine discretization of the geographic domain allows a finite sequence of ReCom moves (on the adjacency graph) to connect them. In practice this could mean that a particular map is not reachable from the initial map when considering voting precincts as geographic units, but could become reachable when working with Census blocks.

Related Work. In the discrete setting, the context for the reachability problem consists of a graph G with n nodes, a number of districts k and a *slack* $\varepsilon \geq 0$. Valid partitions are defined as partitions of $V(G)$ into k non-empty subsets (called *districts*) that each induce connected subgraphs such that the number of vertices in each district lies in the interval $[(1-\varepsilon)\frac{n}{k}, (1+\varepsilon)\frac{n}{k}]$. Two common reconfiguration moves on the space of valid partitions are the switch move and ReCom move. A *switch move* [15, 21] consists of reassigning a single node to a new district. Using the switch move allows one to construct a Markov chain on the space of valid partitions with easily computable transition probabilities. A

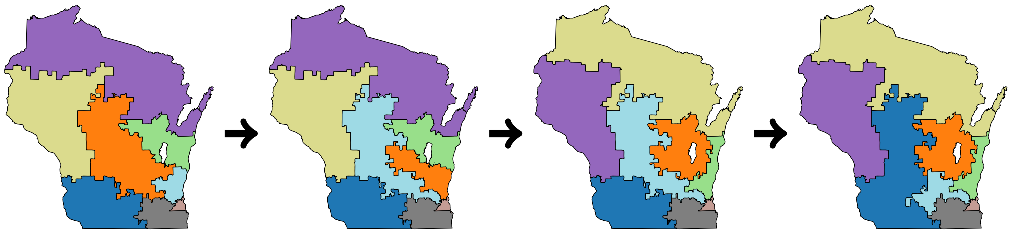


Figure 1 A sequence of three recombination moves on the state of Wisconsin. At each step, two districts are merged and split again. The reachability problem is to determine whether any map can be reached from any other by a finite sequence of such steps.

Metropolis-Hastings weighting can then be used to ensure that the chain samples (in the limit) from any desired distribution on the space of valid partitions. Crucially, however, this relies on the assumption that the state space is connected, i.e., that any two partitions can be connected by switch moves. It is not hard to design concrete examples of graphs for which this is not true with $\varepsilon = 0$. It is known that for $\varepsilon = \infty$, the state space is connected under the switch move when G is biconnected; furthermore, that deciding whether two partitions can be connected by switch moves is PSPACE-complete even when G is planar [2].

The usefulness of the switch move is hampered by the fact that Markov chains built on it tend to mix slowly [23]. As a result, larger reconfiguration moves, that are often more effective on real-world instances, were introduced. The *ReCom move* [13, 14] consists of merging and resplitting two adjacent districts (note that the switch move is a special case of a ReCom move). When designing a Markov chain based on this move, the most common method for resplitting is to draw a random spanning tree of the merged districts and cut an edge such that the resulting connected components form a valid partition. The disadvantage to such a process is that the transition probabilities between partitions appear to be intractable, so that the resulting Markov chain has an unknown stationary distribution. Recently, modifications of the original ReCom Markov chain have been proposed which have computable transition probabilities [4, 6]; however, an accurate description of the stationary distribution still requires the state space to be connected. It is easy to construct a graph G for which the space of valid partitions is not ReCom-connected for $\varepsilon = 0$ (even for a 6×6 grid graph [6]). It is known [3] that the state space is connected whenever G is connected and $\varepsilon = \infty$, and also when G is Hamiltonian and $\varepsilon \geq 2$; deciding whether two partitions can be connected by ReCom moves is PSPACE-complete even when G is a triangulation.

Contributions. In this paper, we introduce a continuous model for redistricting and ReCom moves, where the districts can be arbitrary connected polygons (with real coordinates) in a polygonal domain (Section 2). While the configuration space in this setting contains infinitely many maps, we prove that it is always connected under ReCom moves. Our proof is constructive, and provides an upper bound on the minimum number of ReCom moves between any two maps in terms of the number of districts k and the complexity of the district maps n (i.e., the number of vertices of all polygons in the initial and target maps). We start with the first nontrivial case, $k = 3$ districts in a unit square domain, and show that between any two maps of complexity $O(n)$, there is a reconfiguration path consisting of $O(\log n)$ ReCom moves (Theorem 9 in Section 3). Importantly, the complexity of the map remains $O(n)$ in all intermediate steps. Our reconfiguration algorithm generalizes to k districts in an arbitrary polygonal domain, using a recursion of depth $O(\log k)$. It yields an $\exp(O(\log k \log\log n))$

bound on the number of ReCom moves between two maps; however, for the complexity of intermediate maps we obtain only a weaker bound of $n^{k^{O(1)}}$ (Theorem 10 in Section 4). On the other hand, we show that (even for $k = 3$) the diameter of the configuration space is infinite by constructing pairs of maps which require arbitrarily large numbers of ReCom moves to connect (Theorem 12 in Section 5). The number of moves for these examples grows logarithmically with the complexity of the maps, thereby providing a lower bound which perfectly matches our upper bound.

2 Preliminaries

A *region* is a connected set in $\mathbb{R}^2$ bounded by one or more pairwise disjoint Jordan curves. A *k-district map* $M(\mathcal{R}) = \{D_1, \ldots, D_k\}$ is a decomposition of a region $\mathcal{R}$ into k interior-disjoint regions (that is, $\mathcal{R} = \bigcup_{i=1}^{k} D_i$ and $\mathrm{int}(D_i) \cap \mathrm{int}(D_j) = \emptyset$ for $i \neq j$), where $\mathcal{R}$ is the *domain*, and $D_1, \ldots, D_k$ are the *districts* of the map. We may refer to $M(\mathcal{R})$ simply as M if $\mathcal{R}$ is clear from the context. A *recombination* move (for short, *ReCom*) takes a map $M(\mathcal{R})$ and two districts $D_i, D_j \in M(\mathcal{R})$ and returns a new district map of the same domain $M'(\mathcal{R}) = M(\mathcal{R}) \setminus \{D_i, D_j\} \cup \{D'_i, D'_j\}$. A recombination is *area-preserving* if $\mathrm{area}(D_i) = \mathrm{area}(D'_i)$ and $\mathrm{area}(D_j) = \mathrm{area}(D'_j)$. Two k-district maps, $M(\mathcal{R}) = \{D_1, \ldots, D_k\}$ and $M'(\mathcal{R}) = \{D'_1, \ldots, D'_k\}$, on a domain $\mathcal{R}$ are *area-compatible* if there is a permutation $\pi : \{1, \ldots, k\} \to \{1, \ldots, k\}$ such that $\mathrm{area}(D_i) = \mathrm{area}(D'_{\pi(i)})$ for all $i = 1, \ldots, k$.

We assume that the domain $\mathcal{R}$ is a simple polygon, and each district is a connected polygon (possibly with holes). The *configuration space* of a map $M(\mathcal{R})$ is the set of all polygonal district maps on $\mathcal{R}$ that are area-compatible with $M(\mathcal{R})$. We define the *complexity* of a map M as the total number of vertices on the boundaries of all districts in $M(\mathcal{R})$. We show (in Section 4) that w.l.o.g. we may assume a unit square domain $\mathcal{R} = [0, 1]^2$. The *area* of a polygon P, denoted $\mathrm{area}(P)$, is either the Euclidean area of P or the integral $\int_P \delta$ of some nonnegative integrable density function $\delta : \mathcal{R} \to \mathbb{R}_{\geq 0}$. We show (Theorem 10) that between any pair of area-compatible district maps there is a sequence of area-preserving recombinations (i.e., the configuration space of area-compatible district maps is connected).

Weak Representation. In intermediate maps of a ReCom sequence, we use infinitesimally narrow *corridors* to keep the districts connected. In order to handle narrow corridors efficiently, we rely on a compressed representation of district maps using *weak embeddings* (defined below), where each corridor is represented by a polygonal path; see Fig. 2. The compressed representation has two key advantages: (1) We may assume that corridors have zero area; and (2) we may reduce the total number of vertices by representing several parallel corridors by overlapping polygonal paths (with shared vertices). In Sections 3–4, we construct a sequence of ReCom moves on compressed maps. We show (in Proposition 1 below) that the polygonal paths can be thickened into narrow corridors in each stage of the ReCom sequence to produce a ReCom sequence in which the districts are simple polygons.

An *embedding* of a planar graph G is an injective function from G (seen as a 1-dimensional topological space) to $\mathbb{R}^2$; intuitively it is a drawing of G in which edges can intersect only at common endpoints. A *weak embedding* of G is a continuous function from G to $\mathbb{R}^2$ such that, for every $\varepsilon > 0$, each vertex can be moved by at most ε and each edge can be replaced by a curve within Fréchet distance ε to form an embedding of G (i.e., an *ε-perturbation* of a weak embedding is an embedding). In particular, a *simple polygon* is a piecewise linear embedding of a cycle and the region bound by it; and a *weakly simple polygon* is a piecewise linear weak embedding of a cycle and the region bounded by it. A *polygon* (with possible holes) is a simple polygon with pairwise disjoint simple polygons (holes) removed. Similarly, a *weak polygon* is a weakly simple polygon with pairwise disjoint weakly simple polygons removed.

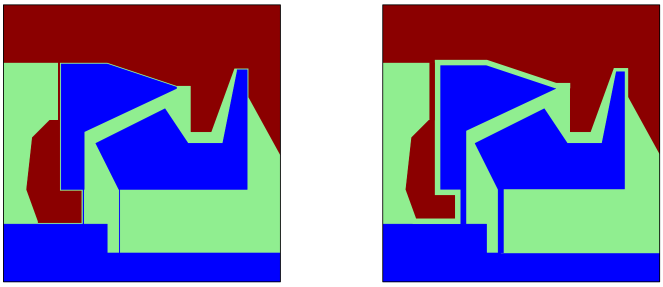


Figure 2 An example of a weak embedding of a map. Left: multiple corridors connect disconnected regions. Right: the corridors are thickened to create three simple polygons.

For a district map M, the boundaries of the districts jointly form a straight-line embedding of some abstract graph G. By identifying edges on opposite sides of narrow corridors, we obtain a weak embedding of G. In a weak embedding, two or more corridors may overlap, and we maintain a linear order among all overlapping corridors.

We use the machinery introduced by Akitaya et al. [1] (based on earlier work [12, 7]); see also [5, 16]. A weak embedding of G is a piecewise linear map of $\varphi : G \to \mathbb{R}^2$. The *image graph* H is a planar straight-line graph formed by the image $\varphi(G)$, where overlapping vertices (edges) of G are mapped to the same vertex (edge). A *weak representation* of G comprises of a weak embedding φ and a linear order of overlapping edges of $\varphi(G)$ along each edge of H. We define an ε*-thickening* of H so that G admits an embedding ψ into the ε-thickening of H so that the Fréchet distance between φ and ψ is at most ε. We call ψ an ε*-perturbation* of the weak representation if the order of the edges of G in the neighborhood of an edge of H agrees with the given linear order. It is known that if G has n vertices, then an ε-perturbation $\psi(G)$ with $O(n)$ vertices can be computed in $O(n \log n)$ time [1].

Weak Representation for ReCom Sequences. We construct a ReCom sequence in two passes: The first pass operates on a generic ε-perturbation, where the *area* of each district is given by the weak representation (hence the corridors have zero area). The second pass then expands the weak representations into an ε-perturbation, using Proposition 1 below (see the full version of the paper for omitted proofs), where each district is a simple polygon with the desired area. Note that the number of moves is determined in the first pass.

▶ **Proposition 1.** *Given two area-compatible k-district maps and a sequence of area-preserving ReCom moves where districts in intermediate maps are weak polygons with $O(n)$ vertices, we can compute a sequence of area-preserving ReCom moves of the same length where the districts in intermediate maps are all polygons with $O(n)$ vertices.*

We define the *compressed complexity* of a district map as the number of vertices in the image graph H of the weak representation (that is, repeated vertices are counted only once). The number of ReCom moves produced by our algorithm in Sections 3–4 depends on the compressed complexity. Using ε-perturbations would increase the complexity of maps. For this reason, it is also useful in our analysis to convert an ε-perturbations to a weak embedding which we do by applying the inverse of the operations described here. Throughout this paper we use set operations on weak polygons such as $D_1 \cup D_2$ where D_1 and D_2 are weak polygons. Let D_1' and D_2' be the polygons obtained by the ε-perturbation defined in Proposition 1. We define $D_1 \cup D_2$ to be the weak polygon obtained from $D_1' \cup D_2'$.

3 Reconfiguration for Three Districts

In this section, we consider maps with three districts with a total of n vertices in a unit square domain $\mathcal{R} = [0,1]^2$. We show that any 3-district map $M(\mathcal{R}) = \{D_1, D_2, D_3\}$ can be transformed by a finite sequence of ReCom moves into an area-compatible *canonical map* in which the districts are axis-aligned rectangles, Q_1, Q_2 and Q_3, of unit width such that $\text{area}(Q_i) = \text{area}(D_i)$ for $i = 1, 2, 3$.

3.1 Overview of the Algorithm

Our algorithm for transforming a map into the canonical map consists of three stages, each containing multiple ReCom moves:

- *Preprocessing* (Section 3.2). In this stage, we ensure that our three districts are ordered top to bottom in a well-defined way, and the middle district has the largest area. Moves needed: $O(1)$.
- *Gravity moves* (Section 3.3). We perform three ReCom moves to place the districts into their final positions, with the possible exception of corridors. Moves needed: 3.
- *Exchange sequences* (Section 3.5). Corridors maintaining connectivity are carefully removed, using a tree representation to determine a move that simultaneously removes a constant fraction of corridors. Moves needed: $O(\log n)$.

3.2 Preprocessing: Ordering Property

First we transform the three given districts into simple polygons if necessary. While there is a district D_i that is a polygon with holes, there is an adjacent district D_j contained within a hole. Recombine D_i and D_j to create a single-edge corridor between D_j and the outer boundary of D_i. Next, we create corridors, if necessary, such that each district touches both the left and right sides of $\mathcal{R}$. While there is a district that is not adjacent to the left (resp., right) side s of the $\mathcal{R}$, let D_i be such a district closest to s and let D_j be an adjacent district that already touches s; then we recombine D_i and D_j and append to D_i a shortest path to s along the boundary of D_j. Thus, both districts remain simply connected. As all corridors run along existing vertices of the three districts, the complexity of the map does not increase. This stage takes $O(1)$ ReCom moves.

After preprocessing, the intersection of each district with the left (resp., right) side of the square domain is connected; and the order of these intersections is the same on both sides, or else two districts would cross. Therefore, the districts can be ordered from top to bottom.

We also need to establish the property that the middle district has the largest area. This can be done trivially with a single ReCom move between the middle district and the largest district of the three. We call these properties combined the *ordering property*:

▶ **Definition 2.** *A three district map $M(\mathcal{R}) = \{D_1, \ldots, D_k\}$ satisfies the **ordering property** if the intersection of each district with the left (resp., right) side of the square domain is connected, and the middle district, as defined by the resulting order from top to bottom, has area greater than or equal to each other district.*

We assume that the districts are simple polygons in the unit square with a total of n vertices and describe the details of the recombination moves as we use them in the algorithm. To reconfigure the districts into their canonical positions, apart from possible corridors, we perform three gravity moves.

3.3 Gravity Move

Assume that M is a 3-district map satisfying the ordering property, with districts labeled D_1 (red), D_2 (green), and D_3 (blue) from top to bottom. We describe the move GRAVITY(D_1, D_2), which recombines the red and green districts; refer to Figures 3–4. Let $P = D_1 \cup D_2$, which is a weakly simple polygon by the ordering property. By continuity, there exists a horizontal line ℓ (that we call the *waterline*) that partitions the plane into upper and lower halfplanes ℓ^+ and ℓ^-, resp., such that $\text{area}(P \cap \ell^+) = \text{area}(D_1)$ and $\text{area}(P \cap \ell^-) = \text{area}(D_2)$. We shall define new districts D_1' and D_2', resp., that contain $P \cap \ell^+$ and $P \cap \ell^-$.

Note, however, that $P \cap \ell^+$ and $P \cap \ell^-$ may be disconnected. We then reconnect disjoint components of each district by corridors along the boundary of P; see Fig. 4. Note that, by the ordering property, there is a path π on the boundary of D_3 (blue) between the left and right side of the domain $\mathcal{R}$. If there are two or more components of $P \cap \ell^+$, they are separated by blue and, therefore they all touch the path π. Therefore $(P \cap \ell^+) \cup \pi$ is a connected region. Similarly, $(P \cap \ell^-) \cup \pi$ is also connected.

We define a *red graph* as follows: the vertices are the connected components of $P \cap \ell^+$ and edges are minimal arcs along $\pi \cap \ell^-$ that connect two distinct components of $P \cap \ell^+$. Since $(P \cap \ell^+) \cup \pi$ is connected, then the red graph is connected. Consider an arbitrary spanning tree of the red graph, and add its edges (as corridors) to the red district along the boundary of P. This completes the definition of D_1'.

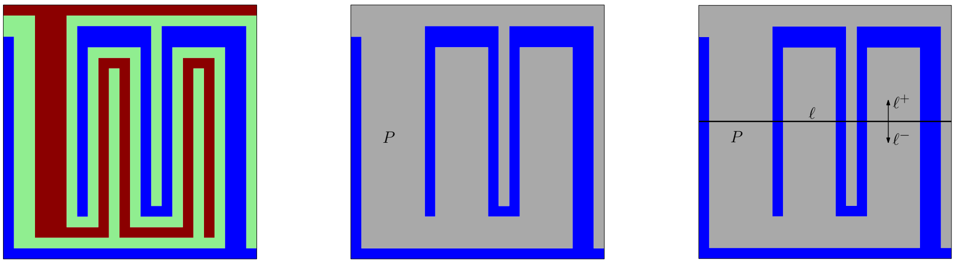


Figure 3 The setup for a gravity move between the red district D_1 and green district D_2. Left: a district map satisfying the ordering property. Middle: the union $P = D_1 \cup D_2$ is shown in gray. Right: the horizontal line ℓ equipartitions the gray polygon P.

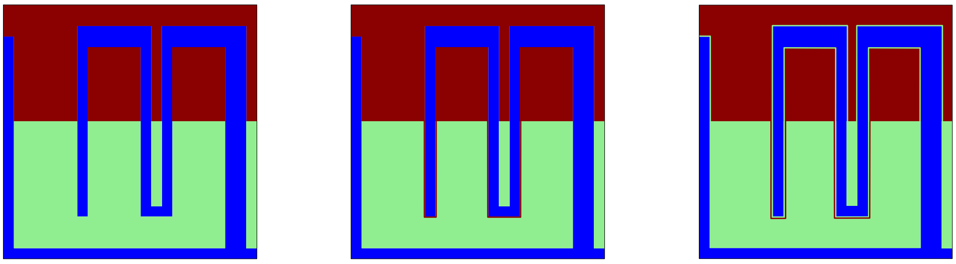


Figure 4 Constructing the result of a gravity move between red and green on the map in Figure 3. Left: the red region $P \cap \ell^+$ and the green region $P \cap \ell^-$ are each disconnected. Middle: red corridors create a connected red district D_1'. Right: green corridors create a connected green district D_2' and restore the ordering property.

Since the blue district is simply connected, each component of $P \cap \ell^-$ also intersects ℓ and therefore is adjacent to the red district D_1'. Intuitively, we add corridors along π "coating" the blue district with green and thus restoring the ordering property. Note that π may pass along the boundary of D_1', including all red corridors, and the boundaries of the components of $P \cap \ell^-$. Formally, we add green corridors at the intersection of π and $\partial D_1'$, if such a corridor is parallel to a red corridor, it runs between the blue district and the red corridor. That defines D_2' and concludes the description of the gravity move.

▶ **Lemma 3.** *Assume D_1 and D_2 are the top two districts on a map satisfying the ordering property. Then* Gravity(D_1, D_2) *is an area-preserving ReCom move that maintains the ordering property.*

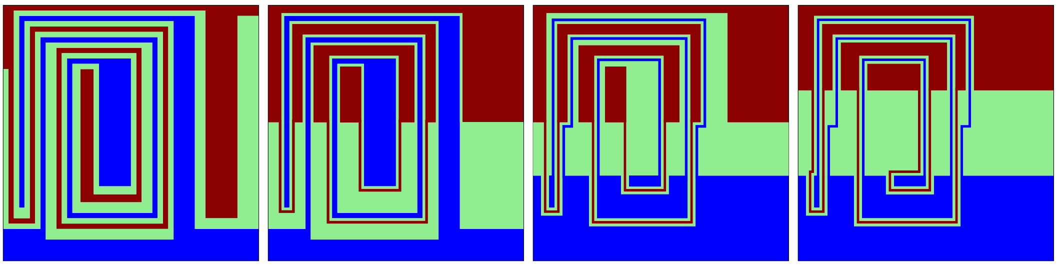
Since each waterline intersects an edge of a district at most once we have:

▶ **Lemma 4.** *Assume D_1, D_2 and D_3 each have at most m vertices. Then* Gravity(D_1,D_2) *produces districts D_1' and D_2', each with at most $O(m)$ vertices.*

The move Gravity(D_3, D_2) is defined analogously: a reflection in a horizontal line that reverses the order of the three districts, such that $D_1' = D_3$ becomes the top district and $D_2' = D_2$ is the middle district, then apply Gravity(D_1', D_2'), followed by another reflection.

▶ **Lemma 5.** *Let M be a map satisfying the ordering property, with districts D_1, D_2 and D_3 from top to bottom. Then* Gravity(D_1, D_2) *returns a map that satisfies the ordering property and D_1' is disjoint from Q_3 with the possible exception of corridors, where Q_3 is the axis-aligned rectangle of the blue district in the canonical map.*

Proof. It suffices to show that the horizontal line ℓ is above Q_3. Lemma 3 yields the rest. By definition, the area below ℓ is at least area$(D_2) \geq$ area(Q_3), since D_2 has the maximum area of the three districts. Thus, the line ℓ is above Q_3. ◀



(a) (b) (c) (d)

Figure 5 An example of a sequence of three gravity moves. (a) A starting configuration; (b) the result of Gravity(D_1, D_2); (c): the result of Gravity(D_3, D_2); (d) the result of the third gravity move Gravity(D_1, D_2).

After three gravity moves each district has positive area only in the regions in their corresponding districts in the canonical configuration (see Figure 5 for an example).

▶ **Lemma 6.** *Let M be a map satisfying the ordering property, with districts D_1, D_2 and D_3 from top to bottom. Then the sequence of three moves* Gravity(D_1, D_2), Gravity(D_2, D_3), *and* Gravity(D_1, D_2) *return a map M' where D_1, D_2, and D_3 are each contained in their canonical rectangles, with the possible exception of some corridors.*

3.4 Tree Representation of a Region

After prepocessing and the three gravity moves in Lemma 6, we want to eliminate corridors. We encode the topology of the region $P = D_1 \cup D_2$ in a graph that we use for the Exchange sequence, described below.

We define the *corridor graph* $T(R)$ of a weakly simple polygon $R \subset \mathcal{R}$. A weakly simple polygon has a natural decomposition into pairwise disjoint simple polygons and corridors (polygonal paths). The nodes of $T(R)$ are simple polygons in the decomposition of R, and the edges represent corridors between two polygons in R. Denote the set of edges by $E(T(R))$. At each node, the rotation of the incident edges represents the counterclockwise order of corridors along the corresponding polygon in R. The weight of each node is the area of the corresponding polygon. As corridors have zero thickness, the total weight of the nodes is $W = \text{area}(R)$.

In particular, we want to consider the corridor graph of $P = D_1 \cup D_2$. Assume that M is a 3-district map returned by the three gravity moves in Lemma 6. By the ordering property, we know that the intersection of D_1 and D_2 is a simple path - either from one side of the square to another or, if D_1 is contained in D_2, then it is a closed curve. Thus, P is a weakly simple polygon. Let Q_{12} be the union of the two axis-aligned rectangles that contain D_1 and D_2 in the canonical configuration. Then, the nodes of $T(P)$ are simple polygons in $P \cap Q_{12}$ (regions bounded by corridors of D_3) and the edges are corridors in $R \setminus Q_{12}$ that connect two such polygons (corridors of D_1 and D_2 running through Q_3). Note, however, that a corridor in P may be the union of three parallel corridors in D_2, D_1, and D_2, resp.; see Fig. 6. Since P is a weakly simple polygon, $T(P)$ is a tree; see Fig. 6. Note that the number of vertices in $T(P)$ is bounded above by the compressed complexity of the map and that many different maps can have the same corridor graph.

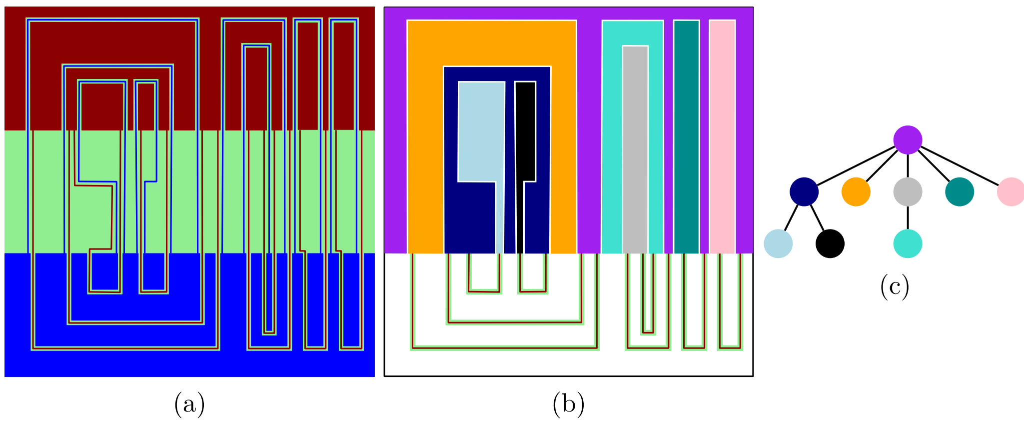


Figure 6 (a) A map M after 3 Gravity moves. (b) The nodes of the corridor graph $T(P)$ correspond to connected components of $P \cap Q_{12}$, indicated by distinct colors. (c) The corridor graph $T(P)$ encodes the topology of P.

We use the corridor graph $T(P)$ to eliminate corridors. Consider what happens if the tree has a leaf that is entirely part of the green district (see Fig. 7). This means that by doing a gravity move between green and blue we can eliminate the green and blue corridors adjacent to this leaf, removing the leaf from the tree altogether. Our goal is therefore to create a part of the tree which is entirely green.

The *centroid* of a vertex-weighted tree of total weight W is a vertex whose removal partitions the tree into subtrees of weight at most $\frac{W}{2}$ each. Jordan [19] proved that every tree (with unit weights) has a centroid; this was perhaps the oldest separator theorem [20, 22]. The result extends to weighted trees: a greedy algorithm finds the centroid in linear time.

Let c be a centroid of $T(P)$, and assume that $T(P)$ is rooted at c. A subtree of $T(P)$ is *contiguous* if it consists of the centroid c, some children of c that are consecutive in the rotation order of c, and all their descendants in $T(P)$.

▶ **Lemma 7.** *There exists a contiguous subtree T^* of $T(P)$ such that:* (i) *T^* contains at least $\frac{1}{3}$ of the vertices of $T(P)$, and* (ii) *the weight of T^* is at most $\frac{W}{2} + w(c)$, where $w(c)$ is the weight of the centroid c.*

Proof. By the definition of the centroid c, the removal of c produces a forest $T(P) - c$, where the weight of each component (tree) is at most $\frac{W}{2}$. Partition these $\deg(c)$ trees into up to three forests of consecutive subtrees such that each forest has weight at most $\frac{W}{2}$ as follows. Begin with a partition into $\deg(c)$ forests, each containing a single tree, and maintain their cyclic order around c. While there are two consecutive forests whose combined weight is at most $\frac{W}{2}$, merge them into a single forest. The while loop terminates with three or fewer forests: Indeed, for four or more forests, the combined weight of at least one of the consecutive pairs would be at most $\frac{W}{2}$ by the pigeonhole principle. Since we partition $T(P) - c$ into three forests, one of them contains at least $\frac{1}{3}$ of the vertices $T(P) - c$. Adding c to this forest, we obtain a contiguous subtree of $T(P)$ containing at least $\frac{1}{3}$ of the vertices of $T(P)$. ◀

3.5 Exchange Sequence

We now describe the *exchange sequence*, a sequence of three ReCom moves, which eliminates a fraction of the corridors and reduces the (compressed) complexity of the map. Assume we are given a 3-district map M satisfying the ordering property. As before, label its districts red, green, and blue from top to bottom. We further require that there exist two horizontal lines ℓ_1 and ℓ_2 such that red has positive area only above ℓ_2, blue has positive area only below ℓ_1, and green has positive area only between ℓ_1 and ℓ_2 (cf. Lemma 6). See Figure 7 for an example.

Let c be a centroid of $T(P)$, where $P = D_1 \cup D_2$ and let T^* be a contiguous subtree of $T(P)$ rooted at the centroid, as in Lemma 7. The exchange sequence consists of the following three ReCom moves:

1. ReCom green and red: Let Q denote the regions of T^* except for the region corresponding to node c. First make Q green. Then partition the remaining region $P \setminus Q$ with a gravity-like move as follows. Apply a Gravity move w.r.t. $P \setminus Q$ to subdivide it into two weak polygons of areas $\text{area}(D_1)$ for red and $\text{area}(D_2) - \text{area}(Q)$ for green; see Fig. 7. After this ReCom move, D_1 is weakly simple and D_2 is a weak polygon (in which D_1 is a hole if Q is a weak polygon with a hole).
2. ReCom green and blue removing unnecessary green and blue corridors simultaneously as follows. Remove any green and blue monochromatic corridors corresponding to all edges of T^*. Note that this merges some nodes of $T(D_3)$ (see Fig. 7), and creates cycles in $T(D_3)$. While there is a cycle in $T(D_3)$ remove a blue corridor in an edge of $T(D_3)$ in a cycle. As this process modifies only green and red, it requires a single ReCom move. After this ReCom move, D_3 is a weakly simple polygon and D_2 is a weakly simple or weak polygon.
3. ReCom green and red with a Gravity move, restoring the ordering property.

▶ **Lemma 8.** *Let $M = \{D_1, D_2, D_3\}$ be a 3-district map with the ordering property, and $M' = \{D'_1, D'_2, D'_3\}$ the map returned by an Exchange sequence on M. Let $P = D_1 \cup D_2$ and $P' = D'_1 \cup D'_2$. Then, M' satisfies the ordering property, and $|E(T(P'))| \leq \frac{2}{3}|E(T(P))|$.*

3.6 Full Reconfiguration Algorithm

Overall, the algorithm for a 3-district map $M([0,1]^2) = \{D_1, D_2, D_3\}$ works as follows: after a preprocessing phase of $O(1)$ ReCom moves, apply the sequence of three moves Gravity(D_1, D_2), Gravity(D_2, D_3), and Gravity(D_1, D_2); compute the corridor graph

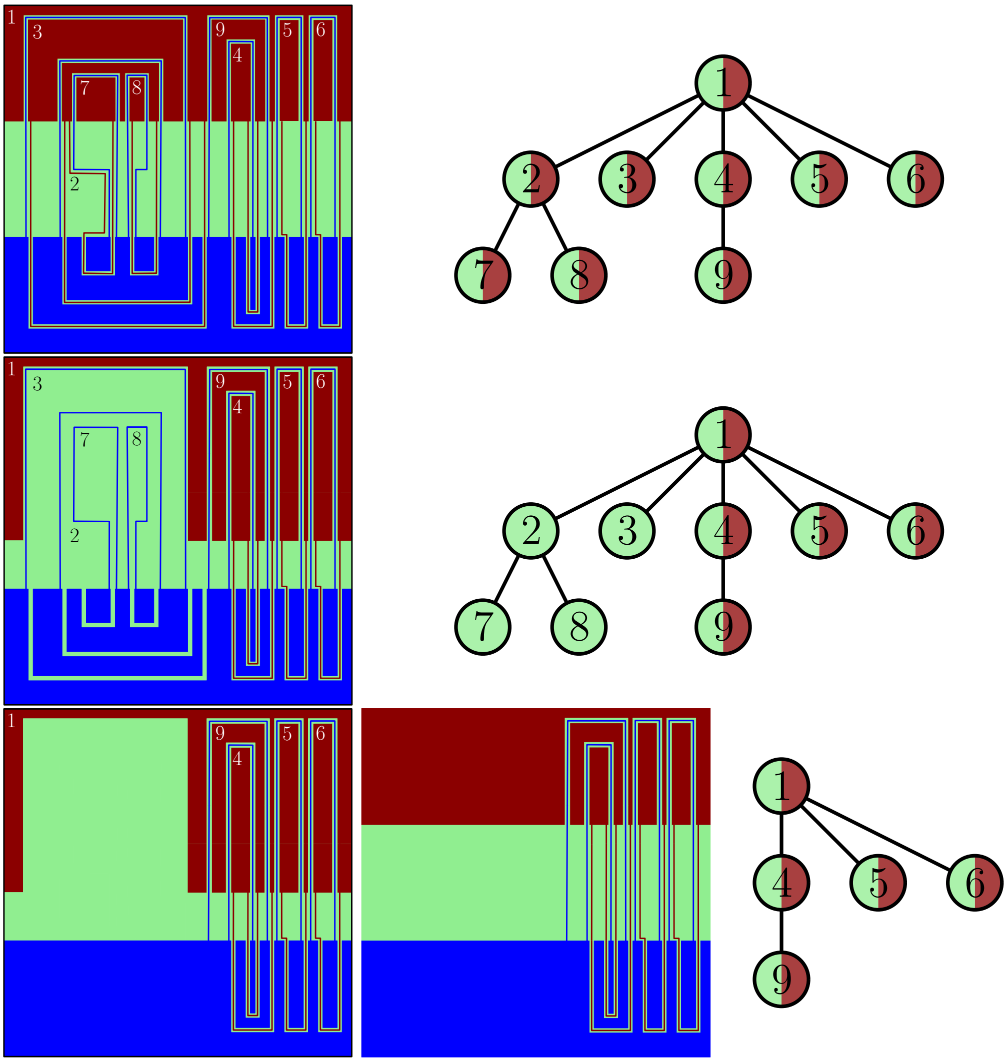


Figure 7 An exchange sequence, shown with maps (left) and corresponding tree representations (right). Top: a map returned by a sequence of three gravity moves. Middle: using node 1 as the centroid c and filling the subtree containing nodes 2, 3, 7 and 8 with green. Bottom: removing unnecessary corridors and performing a gravity move.

$T(P)$ for $P = D_1 \cup D_2$; while $T(P)$ has two or more nodes, apply an exchange sequence. Once $T(P)$ has one node, Gravity(D_1, D_2) yields the canonical configuration.

▶ **Theorem 9.** *Given a 3-district map $M([0,1]^2) = \{D_1, D_2, D_3\}$ of complexity n, there is a sequence of $O(\log n)$ ReCom moves that transforms it into a canonical map. Furthermore, the districts in each intermediate map are polygons with $O(n)$ vertices and at most one hole.*

Proof. After preprocessing, three Gravity moves bring the districts into canonical form with the possible exception of corridors. Each exchange sequence eliminates a constant fraction of corridors by Lemma 8. After $O(\log n)$ ReCom moves we then obtain the canonical configuration.

The algorithm described above produces a sequence of ReCom moves, where the districts in intermediate maps are weak polygons. By Proposition 1, these maps can successively be perturbed into polygons. This completes the proof of the first claim.

It remains to show that the districts in each intermediate map are polygons with $O(n)$ vertices and at most one hole. By construction, the only possible hole appears in the green district after the first step of the EXCHANGE sequence. Each of the $O(1)$ ReCom moves in preprocessing adds a corridor with $O(n)$ vertices, and so each district has $O(n)$ vertices at the end of preprocessing. By Lemma 4, each gravity move increases the number of vertices by a constant factor. After three gravity moves, each district still has $O(n)$ vertices.

The algorithm applies $O(\log n)$ exchange sequences. At the end of every exchange sequence, the districts are in canonical form with the exception of corridors. Each exchange sequence removes some of the corridors, and does not create new corridors. It should be clear that the complexity of the blue district only decreases since corridors are only eliminated and never created. Note that intermediate ReCom moves within an EXCHANGE sequence (step 1) may add $O(n)$ new vertices to the red district. In an exchange sequence, the 1st ReCom move is a GRAVITY move w.r.t. a sub-polygon, and creates only $O(n)$ new vertices by Lemma 4. The 2nd ReCom move eliminates corridors (and the corresponding vertices); and the 3rd ReCom move eliminates any other vertices created in the 1st move of the sequence. Thus, the complexities of the red and green districts decrease after one EXCHANGE sequence.

Finally, when we perturb all weak polygons into polygons in the entire ReCom sequence, the number of vertices remains $O(n)$ for each district by Proposition 1. ◀

4 Reconfiguration for k Districts

We generalize our algorithm to an arbitrary number of districts, using recursion. For any $3 \le k \le n$, an *instance* $I = (M(\mathcal{R}), M'(\mathcal{R}), \delta)$ of the problem consists of two area-compatible k-district maps $M(\mathcal{R}) = \{D_1, \dots, D_k\}$ and $M'(\mathcal{R}) = \{D'_1, \dots, D'_k\}$, where $\mathcal{R}$ is a weak polygon with at most one hole, and δ is a density function. We define the complexity of I (denoted $|I|$) as the pair (k, n), where n is the maximum over the compressed complexities of M and M', and the complexities of all districts D_i and D'_i ($i \in \{1, \dots, k\}$). The overall recursive strategy goes as follows (see the full paper for the details): First construct a piecewise linear retraction from a (possibly punctured) unit square $\mathcal{S}$ to $\mathcal{R}$, and extend M and M' to two maps on $\mathcal{S}$. If $k \ge 4$, then group the k districts into three *superdistricts*, each containing $\lfloor k/3 \rfloor$ or $\lceil k/3 \rceil$ districts; and run the algorithm in Section 3 on the superdistricts. Note that each ReCom move on a pair of superdistricts is an instance of our problem with fewer districts, which can be solved recursively. The retraction then transforms the ReCom sequence on $\mathcal{S}$ to a ReCom sequence on $\mathcal{R}$. We analyze the recursion to give a bound on the number of ReCom moves.

▶ **Theorem 10.** *Given any two area-compatible polygonal k-district maps of complexity at most n in a simply connected domain,* $\exp(O(\log k \log \log n)) = (\log n)^{O(\log k)} = k^{O(\log \log n)}$ *ReCom moves are sufficient to transform one into the other. Furthermore, the complexity of each map in intermediate steps is $n^{k^{O(1)}}$.*

Proof. For $3 \le k \le n$, let $T(k, n)$ denote the minimum number of ReCom moves that can transform any polygonal k-district map to any other with compatible areas, and the domain as well as each district is a polygon with at most n vertices. From an instance $I(k, n)$, our algorithm makes $O(\log n)$ recursive calls of the form $I(\frac{2k}{3}, c \cdot n)$, where c is a constant. Then,

$$T(k, n) \le O\left(T\left(\frac{2k}{3}, c \cdot n\right) \cdot \log n\right) + O(k).$$

The height of the recursion tree is $O(\log k)$ and the maximum branching factor is $O(\log(n \cdot c^{\log k})) = O(\log n + \log k) = O(\log n)$ since $k < n$. Then $T(k, n)$ solves to $\exp(O(\log k \log\log n)) = (\log n)^{O(\log k)} = k^{O(\log\log n)}$. By Proposition 1, we can convert the ReCom sequence on weak representation to a ReCom sequence of the same length in which all districts are simple polygons.

The analysis above prioritized the number of ReCom moves, rather than the complexity of the map at intermediate steps. For instance, consider the recursion that simulates a ReCom move of superdistricts transforming a k-district map $M(\mathcal{R})$ into $M'(\mathcal{R})$. Our algorithm recurses on a $\frac{2k}{3}$-district map of complexity $c \cdot n$ on a punctured square $\mathcal{S}$, which yields a sequence of $O(\log n)$ ReCom moves. However, to convert this into a sequence of ReCom moves on k-district maps, one must apply a retraction $\mathcal{H}^*$ (in the full paper) to every intermediate map, retracting a weakly simple polygon H to its boundary ∂H. Since the complexity of H could be $\Omega(n)$, H might cross the same district $\Omega(n)$ times, which causes $\mathcal{H}^*$ to push the district into $\Omega(n)$ narrow corridors along the boundary of H. This might cause the complexity of the district to increase to $\Omega(n^2)$ in intermediate steps. The retraction $\mathcal{H}^*$, described in the full paper, ensures that the complexity goes up from n to at most $O(n^2)$ after applying $\mathcal{H}^*$ in each recursive step. Since the depth of the recursion tree is $O(\log k)$, the maximum complexity of all intermediate maps is $n^{2^{O(\log k)}} = n^{k^{O(1)}}$. Note that this does not increase the number of ReCom moves since M and M' are determined in the parent level, and $\mathcal{H}^*$ is only applied to recover intermediate steps between M and M', which are obtained from lower complexity maps in the children level. ◀

5 Lower Bound Construction

This section shows that $\Omega(\log n)$ ReCom moves are sometimes necessary to transform a given map of complexity n into canonical form, even for three districts of equal areas in $[0, 1]^2$.

Overview. We describe an initial map with 3 districts in a unit square, and show that after k ReCom moves, each district contains an arc of a specific combinatorial pattern (defined below). These arcs are defined recursively, each iteration roughly tripling the complexity of the arcs. Thus the total complexity of the arcs in iteration ℓ is $O(3^\ell)$. The initial district map is a thickening of one of these arcs after $m \geq 6$ iterations. We show that if each district contains an arc from iteration ℓ, then after a recombination they each contain an arc of iteration $\ell - 4$. In the canonical configuration, each district can only contain arcs of iteration 1. Then, the number of recombinations from the initial district map to the canonical configuration is at least linear in the number of iterations.

Construction. We first describe the family of simple arcs F_ℓ, for all $\ell \in \mathbb{N}_0$, mentioned in the overview. All arcs in F_ℓ will start at the ε-neighborhood of the left side of the square and end at the ε-neighborhood of the right side, crossing the middle section 3^ℓ times. Each family F_ℓ can be described with a combinatorial pattern, namely, the order in which the arcs traverse the 3^ℓ segments in the middle section of the square. In the base case, F_0 is the set of arcs that cross the middle section only once. Given an arc $\gamma_\ell \in F_\ell$, we describe an arc $\gamma_{\ell+1} \in F_{\ell+1}$. We construct two arcs, $\gamma_\ell^+, \gamma_\ell^- \in F_\ell$, that closely follow γ_ℓ on the left and on the right, respectively, and are mutually noncrossing. Then $\gamma_{\ell+1}$ is the concatenation of γ_ℓ, the reverse of γ_ℓ^-, and γ_ℓ^+, where two consecutive arcs are connected by short arcs in the left and right ε-neighborhoods of the square; see Fig 8. Let $F_{\ell+1}$ be the family of all arcs with the same combinatorial pattern as $\gamma_{\ell+1}$. The following observation follows by construction.

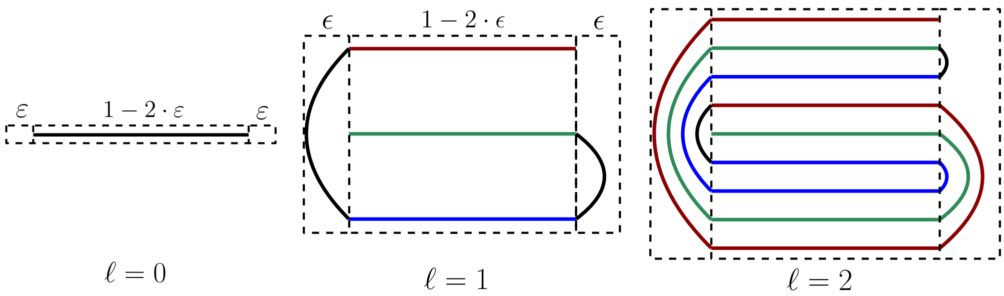


Figure 8 The first three levels of the recursive construction for arcs in F_ℓ, for $\ell \in \{0, 1, 2\}$. Note that the blue, green, and red arcs for $\ell = 2$ each resemble a copy of the entire stage for $\ell = 1$.

▶ **Observation 11.** *For $0 \leq j \leq \ell$, we can partition every arc $\gamma_\ell \in F_\ell$ into 3^j arcs in $F_{\ell-j}$.*

Initial Map. The initial map is drawn relative to an arc $\gamma_m \in F_m$ whose middle segments are equally spaced horizontal line segments in the unit square. The map is a "thickening" of γ_m where the middle section is partitioned into 3^m rectangles of equal area. Each of the three districts is created based on one of three rough copies of γ_{m-1}, i.e., the $(m-1)$-anchors of γ_m. We use the portions of the anchors of γ_m in the ε-neighborhoods of the vertical sides of the unit square to construct corridors that make each district connected.

In the full paper, we show that any ReCom move can only make constant progress (in the number of iterations) towards the canonical map.

▶ **Theorem 12.** *There exist two area-compatible 3-district maps, M and M', both with complexity $O(n)$, such that $\Omega(\log n)$ ReCom moves are necessary to reconfigure M into M', even when the districts in both maps are axis-aligned orthogonal polygons with vertices on an integer grid of size $O(n) \times O(n)$.*

6 Conclusions

We have shown that (in our continuous setting) any pair of area-compatible district maps can be reconfigured into each other by a sequence of area-preserving recombination moves. Though the discrete version of this result remains unsolved (see Related Work), our result suggests that for any two maps, with a discretization of the geographic domain which is granular enough, we can connect them by ReCom moves. However, establishing quantitative bounds on the necessary granularity is left for future work.

Between 3-district maps, the number of recombination moves is $O(\log n)$, where n is the combinatorial complexity of the maps, matching our worst-case lower bound of $\Omega(\log n)$. Between k-district maps, for $k \geq 4$, we construct a sequence of $\exp(O(\log k \log\log n)) = (\log n)^{O(\log k)}$ ReCom moves. It remains an open problem whether the number of moves can be reduced to be polynomial in both k and n. For $k \geq 4$ districts, our algorithm uses a recursion of depth $O(\log k)$. However, this approach increases the complexity of intermediate maps to $n^{k^{O(1)}}$. It is also an open problem whether there exists a sequence of ReCom moves where the complexity of intermediate maps remains polynomial in both k and n.

References

1 Hugo A Akitaya, Greg Aloupis, Jeff Erickson, and Csaba D Tóth. Recognizing weakly simple polygons. *Discrete & Computational Geometry*, 58(4):785–821, 2017.

2 Hugo A. Akitaya, Matthew D. Jones, Matias Korman, Oliver Korten, Christopher Meierfrankenfeld, Michael J. Munje, Diane L. Souvaine, Michael Thramann, and Csaba D. Tóth. Reconfiguration of connected graph partitions. *Journal of Graph Theory*, 102(1):35–66, 2023. doi:10.1002/jgt.22856.

3 Hugo A. Akitaya, Matias Korman, Oliver Korten, Diane L. Souvaine, and Csaba D. Tóth. Reconfiguration of connected graph partitions via recombination. *Theoretical Computer Science*, 923:13–26, 2022. doi:10.1016/j.tcs.2022.04.049.

4 Eric A. Autry, Daniel Carter, Gregory Herschlag, Zach Hunter, and Jonathan C. Mattingly. Multi-scale merge-split Markov chain Monte Carlo for redistricting, 2020. arXiv:2008.08054.

5 Thomas Bläsius, Simon D. Fink, and Ignaz Rutter. Synchronized planarity with applications to constrained planarity problems. In *Proc. 29th European Symposium on Algorithms (ESA)*, volume 204 of *LIPIcs*, pages 19:1–19:14. Schloss Dagstuhl, 2021. doi:10.4230/LIPIcs.ESA.2021.19.

6 Sarah Cannon, Moon Duchin, Dana Randall, and Parker Rule. Spanning tree methods for sampling graph partitions, 2022. arXiv:2210.01401.

7 Hsien-Chih Chang, Jeff Erickson, and Chao Xu. Detecting weakly simple polygons. In *Proc. 26th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1655–1670, 2014. doi:10.1137/1.9781611973730.110.

8 Jowei Chen and Jonathan Rodden. Unintentional gerrymandering: Political geography and electoral bias in legislatures. *Quarterly Journal of Political Science*, 8(3):239–269, 2013. doi:10.1561/100.00012033.

9 Jowei Chen and Jonathan Rodden. Cutting through the thicket: Redistricting simulations and the detection of partisan gerrymanders. *Election Law Journal*, 14(4):331–345, 2015. doi:10.1089/elj.2015.0317.

10 Maria Chikina, Alan Frieze, and Wesley Pegden. Assessing significance in a Markov chain without mixing. *Proceedings of the National Academy of Sciences*, 114(11):2860–2864, 2017. doi:10.1073/pnas.16175401.

11 Vincent Cohen-Addad, Philip N Klein, and Neal E Young. Balanced centroidal power diagrams for redistricting. In *Proc. 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 389–396, 2018.

12 Pier Francesco Cortese, Giuseppe Di Battista, Maurizio Patrignani, and Maurizio Pizzonia. On embedding a cycle in a plane graph. *Discrete Mathematics*, 309(7):1856–1869, 2009. doi:10.1016/j.disc.2007.12.090.

13 Daryl DeFord and Moon Duchin. Redistricting reform in Virginia: Districting criteria in context. *Virginia Policy Review*, 12(2):120–146, 2019.

14 Daryl DeFord, Moon Duchin, and Justin Solomon. Recombination: A family of Markov chains for redistricting. *Harvard Data Science Review*, March 2021. doi:10.1162/99608f92.eb30390f.

15 Benjamin Fifield, Michael Higgins, Kosuke Imai, and Alexander Tarr. Automated redistricting simulation using Markov chain Monte Carlo. *Journal of Computational and Graphical Statistics*, 29(4):715–728, 2020. doi:10.1080/10618600.2020.1739532.

16 Radoslav Fulek and Csaba D. Tóth. Atomic embeddability, clustered planarity, and thickenability. *J. ACM*, 69(2):13:1–13:34, 2022. doi:10.1145/3502264.

17 Gregory Herschlag, Han Sung Kang, Justin Luo, Christy Vaughn Graves, Sachet Bangia, Robert Ravier, and Jonathan C. Mattingly. Quantifying gerrymandering in north carolina. *Statistics and Public Policy*, 7(1):30–38, 2020. doi:10.1080/2330443X.2020.1796400.

18 Gregory Herschlag, Robert Ravier, and Jonathan C. Mattingly. Evaluating partisan gerrymandering in Wisconsin, 2017. arXiv:1709.01596.

19 Camille Jordan. Sur les assemblages des lignes. *Journal für die reine und angewandte Mathematik*, 70:185–190, 1869.

20 Richard J. Lipton and Robert Endre Tarjan. A separator theorem for planar graphs. *SIAM Journal on Applied Mathematics*, 36(2):177–189, 1979. doi:10.1137/0136016.

21 Jonathan C. Mattingly and Christy Vaughn. Redistricting and the will of the people, 2014. arXiv:1410.8796.

22 Gary L. Miller, Shang-Hua Teng, William P. Thurston, and Stephen A. Vavasis. Separators for sphere-packings and nearest neighbor graphs. *J. ACM*, 44(1):1–29, 1997. doi:10.1145/256292.256294.

23 Elle Najt, Daryl DeFord, and Justin Solomon. Complexity and geometry of sampling connected graph partitions, 2019. arXiv:1908.08881.

24 Patti B. Saris, Gary King, Jamal Greene, and Moon Duchin. Redistricting and representation. *Bulletin of the American Academy of Arts and Sciences*, 71(2):47–58, 2018.