Reconfiguration of Connected Graph Partitions via Recombination*

Hugo A. Akitaya[1][0000−0002−6827−2200], Matias Korman[2], Oliver Korten[3], Diane L. Souvaine[4], and Csaba D. Tóth[4,5][0000−0002−8769−3190]

[1] Department of Computer Science, University of Massachusetts Lowell, Lowell, MA, USA. hugo_akitaya@uml.edu
[2] Siemens Electronic Design Automation, Wilsonville, OR, USA.
[3] Department of Computer Science, Columbia University, New York, NY, USA.
[4] Department of Computer Science, Tufts University, Medford, MA, USA.
[5] Department of Mathematics, Cal State Northridge, Los Angeles, CA, USA.

Abstract. Motivated by applications in gerrymandering detection, we study a reconfiguration problem on connected partitions of a connected graph G. A partition of $V(G)$ is **connected** if every part induces a connected subgraph. In many applications, it is desirable to obtain parts of roughly the same size, possibly with some slack s. A **Balanced Connected k-Partition with slack s**, denoted (k, s)**-BCP**, is a partition of $V(G)$ into k nonempty subsets, of sizes $n_1, \ldots, n_k$ with $|n_i - n/k| \leq s$, each of which induces a connected subgraph (when $s = 0$, the k parts are perfectly balanced, and we call it k**-BCP** for short).
A **recombination** is an operation that takes a (k, s)-BCP of a graph G and produces another by merging two adjacent subgraphs and repartitioning them. Given two k-BCPs, A and B, of G and a slack $s \geq 0$, we wish to determine whether there exists a sequence of recombinations that transform A into B via (k, s)-BCPs. We obtain four results related to this problem: (1) When s is unbounded, the transformation is always possible using at most $6(k-1)$ recombinations. (2) If G is Hamiltonian, the transformation is possible using $O(kn)$ recombinations for any $s \geq n/k$, and (3) we provide negative instances for $s \leq n/(3k)$. (4) We show that the problem is PSPACE-complete when $k \in O(n^{\varepsilon})$ and $s \in O(n^{1-\varepsilon})$, for any constant $0 < \varepsilon \leq 1$, even for restricted settings such as when G is an edge-maximal planar graph or when $k \geq 3$ and G is planar.

1 Introduction

Partitioning the vertex set of a graph $G = (V, E)$ into k nonempty subsets $V = \bigcup_{i=1}^{k} V_i$ that each induces a connected graph $G[V_i]$ is a classical problem, known as the **Connected Graph Partition** problem [9,16]. Motivated by fault-tolerant network design and facility location problems, it is part of a broader family of problems where each induced graph $G[V_i]$ must have a certain graph property (e.g., ℓ-connected or H-minor-free). In some instances, it is desirable that

* Supported by NSF CCF-1422311, CCF-1423615, DMS-1800734, and NSERC.

the parts $V_1, \dots, V_k$ have the approximately the same size (depending on some pre-established threshold). A **Balanced Connected k-Partition** (for short, k**-BCP**) is a connected partition requiring that $|V_i| = n/k$, for $i \in \{1, \dots, k\}$ where $n = |V(G)|$ is the total number of vertices. Dyer and Frieze [7] proved that finding a k-BCP is NP-hard for all $2 \leq k \leq n/3$. For $k = 2, 3$ the problem can be solved efficiently when G is bi- or triconnected, respectively [20,22], and is equivalent to the perfect matching problem for $k = n/2$. Later Chlebíková [4] and Chataigner et al. [3] obtained approximation and inapproximability results for maximizing the "balance" ratio $\max_i |V_i| / \min_j |V_j|$ over all connected k-partitions. See also [12,14,19,23] for variants under various other optimization criteria.

In this paper, our basic element is a connected k-partition of a graph $G = (V, E)$ that is balanced up to some additive threshold that we call a **slack** $s \geq 0$, denoted (k, s)**-BCP**. We explore the space of all (k, s)**-BCP**s of the graph $G = (V, E)$. Note that the total number of (k, s)**-BCP**s for all $s \geq 0$, is bounded above by the number k-partitions of V, which is the Stirling number of the second kind $S(n, k)$, and asymptotically equals $(1 + o(1))k^n/k!$ for constant k. This bound is the best possible for the complete graph $G = K_n$.

In a recent application [1,6,18], $G = (V, E)$ represents the adjacency graph of precincts in an electoral map, which should be partitioned into k districts $V_1, \dots, V_k$ where each district will elect one representative. Motivated by the design and evaluation of electoral maps under optimization criteria designed to promote electoral fairness, practitioners developed empirical methods to sample the configuration space of potentialdistrict maps by a random walk on the graph where each step corresponds to some elementary **reconfiguration move** [8]. From a theoretical perspective, the stochastic process converges to uniform sampling [13,15]. However, the move should be **local**, i.e., it must affect a constant number of districts, to allow efficient computation of each move, and it should support rapid mixing (i.e., the random walk should converge, in total variation distance, to its steady state distribution in time polynomial in n). Crucially, the space of (approximately balanced) k-partitions of G must be **connected** under the proposed move. Previous research considered the **single switch** move, in which a single vertex $v \in V$ switches from one set V_i to another set V_j (assuming that both $G[V_i]$ and $G[V_j]$ remain connected). Akitaya et al. [2] proved that the configuration space is connected under single switch moves if G is biconnected, but in general it is NP-hard both to decide whether the space is connected and to find a shortest path between two valid k-partitions. While the single switch is local, both worst-case constructions and empirical evidence [5,18] indicate that it does not support rapid mixing.

In this paper we consider a different move. Specifically, we consider the configuration space of k-partitions under the **recombination** move, proposed by DeFord et al. [5], in which the vertices in $V_i \cup V_j$, for some $i, j \in \{1, \dots, k\}$, are re-partitioned into $V_i' \cup V_j'$ such that both $G[V_i']$ and $G[V_j']$ are connected. We also study variants restricted to balanced or near-balanced partitions, that is, when $|V_i| = n/k$ for all $i \in \{1, \dots, k\}$, or when $\big||V_i| - n/k\big| \leq s$ for a given slack $s \geq 0$. In application domains mentioned above, the underlying graph G is often planar or near-planar, and in some cases it is a triangulation (i.e., an edge-maximal planar graph). Results pertaining to these special cases are of particular interest. Our results lay down theoretical foundations for this model in graph theory and computational tractability. Although our results imply lower bounds in the mixing time of worst-case instances, they have no direct implication for the average case analysis.

Definitions. Let $G = (V, E)$ be a graph with $n = |V(G)|$. For a positive integer k, a **connected k-partition** Π of G is a partition of $V(G)$ into disjoint nonempty subsets $\{V_1, \dots, V_k\}$ such that the induced subgraph $G[V_i]$ is connected for all $i \in \{1, \dots, k\}$. Each subgraph induced by V_i is called a **district**. We write $\Pi(v)$ for the subset in Π that contains vertex v.

Denote by $\mathrm{Part}(G, k)$ the set of connected k-partitions on G. We also consider subsets of $\mathrm{Part}(G, k)$ in which all districts have the same or almost the same number of vertices. A connected k-partition of G is **balanced** (k-BCP) if every district has precisely n/k vertices (which implies that n is a multiple of k); and it is **balanced with slack** $s \geq 0$ ((k, s)-BCP), if $\big||U| - n/k\big| \leq s$ for every district $U \subset V$. Let $\mathsf{Bal}_s(G, k)$ denote the set of connected k-partitions on G that are balanced with slack s, i.e., the set of all (k, s)-BCPs. The set of balanced k-partitions is denoted $\mathsf{Bal}(G, k) = \mathsf{Bal}_0(G, k)$; and $\mathrm{Part}(G, k) = \mathsf{Bal}_\infty(G, k)$.

We now formally define a **recombination move** as a binary relation on $\mathsf{Bal}_s(G, k)$. Two non-identical (k, s)-BCPs, $\Pi_1 = \{V_1, \dots, V_k\}$ and $\Pi_2 = \{W_1, \dots, W_k\}$,are related by a recombination move if there exist $i, j \in \{1, \dots, k\}$, and a permutation π on $\{1, \dots, k\}$ such that $V_i \cup V_j = W_{\pi(i)} \cup W_{\pi(j)}$ and $V_\ell = W_{\pi(\ell)}$ for all $\ell \in \{1, \dots, k\} \setminus \{i, j\}$. We say that Π_1 and Π_2 are a recombination of each other. This binary relation is symmetric and defines a graph on $\mathsf{Bal}_s(G, k)$ for all $s \geq 0$. This graph is the **configuration space** of $\mathsf{Bal}_s(G, k)$ under recombination, denoted by $\mathcal{R}_s(G, k)$.

Balanced Recombination Problem BR(G, k, s): Given a graph $G = (V, E)$ with $|V| = n$ vertices and two (k, s)-BCPs A and B, decide whether there exists a path between A and B in $\mathcal{R}_s(G, k)$, i.e. whether there is a sequence of recombination moves that carries A to B such that every intermediate partition is a (k, s)-BCP.

Our Results. We prove, in Section 2, that the configuration space $\mathcal{R}_\infty(G, k)$ is connected whenever the underlying graph G is connected and the size of the districts is unrestricted. It is easy, however, to construct a graph G where $\mathcal{R}_0(G, k)$ is disconnected. We study what is the minimum slack s, as a function of n and k, that guarantees that $\mathcal{R}_s(G, k)$ is connected for all connected (or possibly biconnected) graphs G with n vertices. We prove that $\mathcal{R}_s(G, k)$ is connected and its diameter is $O(nk)$ for $s = n/k$ when G is a Hamiltonian graph (Section 3). As a counterpart, we construct a family of Hamiltonian planar graphs G such that $\mathcal{R}_s(G, k)$ is disconnected for $s < n/(3k)$ (Section 4).

We prove in Section 5 that $\mathrm{BR}(G, k, s)$ is PSPACE-complete even for the special case when G is a triangulation (i.e., an edge-maximal planar graph), k

is $O(n^{\varepsilon})$ and s is $O(n^{1-\varepsilon})$ for constant $0 < \varepsilon \leq 1$. As a consequence we show that finding a (k, s)-BCP of G is NP-hard in the same setting. Note that the previously known hardness proofs for finding k-BCPs either require that G is weighted and nonplanar [3] or G contain cut vertices [7]. In contrast, if G is planar and 4-connected, then G admits a Hamilton cycle [21] and, therefore, a (k, s)-BCP is easily obtained by partitioning a Hamilton cycle into the desired pieces. Finally, we modify our construction to also show that $\mathrm{BR}(G, k, s)$ is PSPACE-complete even for the special case when G is planar, $k \geq 3$, and s is bounded above by $O(n^{1-\varepsilon})$ for constant $0 < \varepsilon \leq 1$.

2 Recombination with Unbounded Slack

In this section, we show that the configuration space $\mathcal{R}_\infty(G, k)$ is connected under recombination moves if G is connected (cf. Theorem 1). The proof proceeds by induction on k, where the induction step depends on Lemma 2 below.

We briefly review some standard graph terminology. A **block** of a graph G is a maximal biconnected component of G. A vertex $v \in V(G)$ is a **cut vertex** if it lies in two or more blocks of G, otherwise it is a **block vertex**. In particular, if v is a block vertex, then $G - v$ is connected. If G is a connected graph with two or more vertices, then every block has at least two vertices. A block is a **leaf-block** if it contains precisely one cut vertex of G. Every connected graph either is biconnected or has at least two leaf blocks. The **arboricity** of a graph G is the minimum number of forests that cover all edges in $E = (G)$. The **degeneracy** of G is the largest minimum vertex degree over all induced subgraphs of G. It is well known that if the arboricity of a graph is a, then its degeneracy is between a and $2a - 1$.

Lemma 1. *If the arboricity of a graph is a, then it contains a block vertex of degree at most $2a - 1$.*

The proof of Lemma 1 can be found in [17]. The heart of the induction step of our main result hinges on the following lemma.

Lemma 2. *Let G be a connected graph, $k \geq 2$ an integer, and $\Pi_1, \Pi_2 \in Part(G, k)$ be two k-partitions of G. Then there exists a block vertex $v \in V(G)$ such that up to three recombination moves can transform Π_1 and Π_2 each to two new k-partitions in which $\{v\}$ is a singleton distinct.*

Proof. Let $\Pi_1 = \{V_1, \ldots, V_k\}$ and $\Pi_2 = \{W_1, \ldots, W_k\}$. We construct two spanning trees, T_1 and T_2, for G that each contain $k-1$ edges between the districts of Π_1 and Π_2, respectively. Specifically, for $i \in \{1, \ldots, k\}$, let $T(V_i)$ be a spanning tree of $G[V_i]$, $T(W_i)$ a spanning tree of $G[W_i]$. As G is connected, we can augment the forest $\bigcup_{i=1}^{k} T(V_i)$ to a spanning tree T_1 of G, using $k - 1$ new edges, which connect vertices in distinct districts. Similarly, we can augment $\bigcup_{i=1}^{k} T(W_i)$ to a spanning tree T_2 of G. Now, let $G' = T_1 \cup T_2$. By definition, the arboricity of G' is at most 2. By Lemma 1, G' contains a block vertex v with $\deg_{G'}(v) \leq 3$.

We show that we can modify Π_1 (resp., Π_2) to create a singleton district $\{v\}$ in at most three moves. Assume without loss of generality that $v \in V_1$ and $v \in W_1$. Since $\deg_{G'}(v) \leq 3$, we have $\deg_{T(V_1)}(v) \leq 3$ and $\deg_{T(W_1)}(v) \leq 3$. Consequently, $T(V_1) - v$ (resp., $T(W_1) - v$) has at most three components, each of which is adjacent to some other district, since $G' - v$ is connected. Up to three successive recombinations can decrease the district V_1 with the components of $T(V_1) - v$, and reduce V_1 to $\{v\}$. Similarly, at most three successive recombinations can reduce W_1 to $\{v\}$. □

Theorem 1. *Let G be a connected graph and $k \geq 1$ a positive integer. For all $\Pi_1, \Pi_2 \in Part(G, k)$, there exists a sequence of at most $6(k-1)$ recombination moves that transforms Π_1 to Π_2.*

Proof. We proceed by induction on k. In the base case, $k = 1$, and $\Pi_1 = \Pi_2$. Assume that $k > 1$ and claim holds for $k - 1$. By Lemma 2, we can find a block vertex $v \in V(G)$ and up to six recombination moves transform Π_1 and Π_2 into Π_1' and Π_2' such that both contain $\{v\}$ as a singleton district. Since v is a block vertex, $G - v$ is connected; and since $\{v\}$ is a singleton district in both Π_1' and Π_2', we have $\Pi_1 - \{v\}, \Pi_2 - \{v\} \in \mathrm{Part}(G - v, k - 1)$. By induction, a sequence of up to $6(k-2)$ recombination moves in $G - v$ can transform $\Pi_1 - \{v\}$ into $\Pi_2 - \{v\}$. These moves remain valid recombination moves in G if we add singleton district $\{v\}$. Overall, the combination of these sequences yields a sequence of up to $6 + 6(k-2) = 6(k-1)$ recombination moves that transforms Π_1 to Π_2. This completes the induction step. □

3 Recombination with Slack

In this section, we prove that the configuration space $\mathcal{R}_s(G, k)$ is connected if the slack is greater or equal to the average district size, that is, $s \geq n/k$, and the underlying graph G is Hamiltonian (Theorem 2).

Let G be a graph with n vertices that contains a Hamilton cycle C. Assume that n is a multiple of k. A k-partition in $\mathsf{Bal}_s(G, k)$ is **canonical** if each district consists of consecutive vertices along C. Using a slack of $s \geq n/k$, we can transform any canonical k-partition to any other using $O(k^2)$ reconfigurations.

Lemma 3. *Let G be a graph with n vertices and a Hamilton cycle C, $k \geq 1$ is a divisor of n, and $s \geq n/k$. Then the subgraph of $\mathcal{R}_s(G, k)$ induced by canonical k-partitions is connected and its diameter is at most $k^2 + 1$.*

Proof Sketch. We proceed by induction: We assume that the first $\ell \in \{0, \ldots, k\}$ districts each have size $\frac{n}{k}$, and we change the size of the $(\ell + 1)$st district to $\frac{n}{k}$ using at most $k - \ell - 1$ recombinations. Since the average size of the remaining $k - \ell$ districts is n/k, there are two consecutive districts of size at most $\frac{n}{k}$ and at least $\frac{n}{k}$, respectively. We recombine the first such pair of districts, and propagate the changes to the $(\ell + 1)$st district, completing the induction step. □

In the remainder of this section, we show that every k-partition in $\mathsf{Bal}_s(G, k)$ can be brought into canonical form by a sequence of $O(nk)$ recombinations.

Preliminaries. We introduce some terminology. Let $\Pi = \{V_1, \ldots, V_k\} \in \mathsf{Bal}_s(G,k)$ with a slack of $s \geq n/k$. For every $i \in \{1,\ldots,k\}$, a **fragment** of $G[V_i]$ is a maximum set $F \subset V_i$ of vertices that are contiguous along C. Every set V_i is the disjoint union of one or more fragments. The k-partition Π is canonical if and only if every district has precisely one fragment. Our strategy is to "defragment" Π if it is not canonical, that is, we reduce the number of fragments using recombination moves.

We distinguish between two types of districts in Π: A district V_i is **small** if $|V_i| \leq n/k$, otherwise it is **large**. Every edge in $E(G)$ is either an edge or a **chord** of the cycle C. For every $i \in \{1,\ldots,k\}$, let f_i be the number of fragments of V_i. Let T_i be a spanning tree of $G[V_i]$ that contains the minimum number of chords. The edges of $G[V_i]$ along C form a forest of f_i paths; we can construct T_i by augmenting this forest to a spanning tree of $G[V_i]$ using $f_i - 1$ chords.

The **center** of a tree T is a vertex $v \in V(T)$ such that each component of $T - v$ has up to $|V(T)|/2$ vertices. It is well known that every tree has a center. For $i \in \{1,\ldots,k\}$, let c_i be a center of the spanning tree T_i of $G[V_i]$. Let the fragment of V_i be **heavy** if it contains c_i; and **light** otherwise. We also define a parent-child relation between the fragments of V_i. Fragments A and B are in a parent-child relation if they are adjacent in T_i and if c_i is closer to A than to B in T_i. Note that a light fragment and its descendants jointly contain less than $|V_i|/2 \leq (n/k + s)/2$ vertices; see Fig. 1.

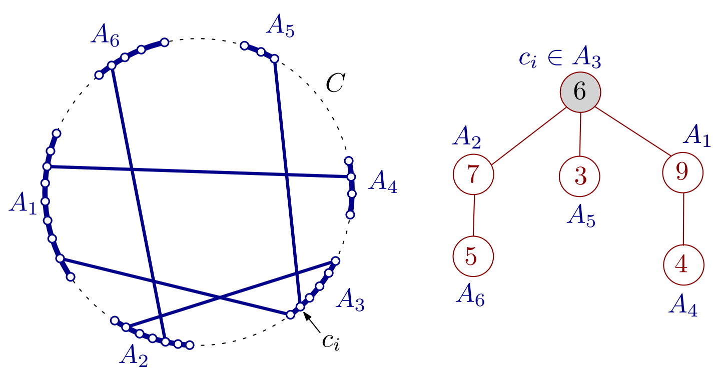

Fig. 1: Left: A distinct V_i with 26 vertices (hollow dots) in six fragments (bold arcs) along C. The spanning tree T_i of $G[V_i]$ contains five edges, with a center at c_i. Right: Parent-child relationship between fragments is defined by the tree rooted at the fragment containing c_i.

The following four lemmas show that we can decrease the number of fragments under some conditions. In all four lemmas, we assume that G is a graph with a Hamiltonian cycle C, and Π is a noncanonical (k,s)-BCP with $s \geq n/k$.

Lemma 4. *If a light fragment of a large district is adjacent to a small district along C, then a recombination move can decrease the number of fragments.*

Proof. Assume without loss of generality that v_1v_2 is an edge of C, where $v_1 \in F_1 \subset V_1$, $v_2 \in F_2 \subseteq V_2$, F_1 is a light fragment of a large district V_1, and F_2 is some fragment of a small district V_2. Let $\overline{F}_1$ be the union of fragment F_1 and all its descendants. By the definition of the center c_1, we have $|\overline{F}_1| < |V_1|/2$. Apply a recombination replacing V_1 and V_2 with $W_1 = V_1 \setminus \overline{F}_1$ and $W_2 = V_2 \cup \overline{F}_1$.

We show that the resulting partition is a (k, s)-BCP. Note also that both $G[\overline{F}_1]$ and $G[V_1 \setminus \overline{F}_1] = G[W_1]$ are connected. Since $v_1v_2 \in E(G)$, then $G[V_2 \cup \overline{F}_1] = G[W_2]$ is also connected. As W_1 contains the center of V_1, we have $|W_1| \geq 1$ and $|W_1| < |V_1| \leq n/k+s$. As V_2 is small, have $|W_2| = |V_2|+|\overline{F}_1| < n/k+n/k \leq 2n/k \leq n/k + s$. Finally, note that $F_1 \cup F_2$ is a single fragment in the resulting k-partition, hence the number of fragments decreased by at least 1. □

Lemma 5. *If no light fragment of a large district is adjacent to any small district along C, then there exists two adjacent districts along C whose combined size is at most $2n/k$.*

Proof. Suppose, to the contrary, that every small district is adjacent only to heavy fragments along C, and the combined size of every pair of adjacent districts along C is greater than $\frac{2n}{k}$, meaning that at least one district is large. We assign every small district to an adjacent large district as follows. For every small district V_i, let F_i be one of its arbitrary fragments. We assign V_i to the large district whose heavy fragment is adjacent to F_i in the clockwise direction along C. Since every large district has a unique heavy fragment, and at most one district precedes it in clockwise order along C, the assignment is a matching of the small districts to large districts. Denote this matching by M. Every district that is not part of a pair in M must be large. By assumption, every pair in M has combined size greater than $\frac{2n}{k}$, so the average district size over the districts in M is greater than $\frac{n}{k}$. The districts not in M are large so their average size also exceeds $\frac{n}{k}$. Overall the average district size exceeds $\frac{n}{k}$. But Π is a k-partition of n vertices, hence the average district size is exactly $\frac{n}{k}$, a contradiction. □

Lemma 6. *If districts V_1 and V_2 are adjacent along C and $|V_1 \cup V_2| \leq n/k+s$, then there is a recombination move that either decreases the number of fragments, or maintains the same number of fragments and creates a singleton district.*

Proof. Assume, w.l.o.g., that $v_1 \in F_1 \subseteq V_1$, $v_2 \in F_2 \subseteq V_2$, where v_1v_2 is an edge of C, and F_1 and F_2 are fragments of V_1 and V_2, respectively. The induced graph $G[V_1 \cup V_2]$ is connected, and $T_1 \cup T_2 \cup v_1v_2$ is one of its spanning trees. If T_1 or T_2 contains a chord, say e, then $(T_1 \cup T_2 \cup v_1v_2) - e$ has two components, T_3 and T_4, each of size at most $n/k + s - 1$. A recombination move can replace V_1 and V_2 with $V(T_3)$ and $V(T_4)$. Since fragments F_1 and F_2 merge into one fragment, the number of fragments decreases by at least one. Otherwise, neither T_1 nor T_2 contains a chord. Then V_1 and V_2 each has a single fragment, so $V_1 \cup V_2$ is a chain of vertices along C. Let v be the first vertex in this chain. A recombination move can replace V_1 and V_2 with $W_1 = \{v\}$ and $W_2 = (V_1 \cup V_2) \setminus \{v\}$. By construction both $G[W_1]$ and $G[W_2]$ are connected, $|W_1| = 1$, $|W_2| = |V_1 \cup V_2| - 1 \leq n/k+s-1$, and the number of fragments does not change. □

Lemma 7. *If there exists a singleton district, then there exists a sequence of at most $k-1$ recombination moves that decreases the number of fragments.*

Proof. Let $C=(v_1,\ldots,v_n)$. Assume without loss of generality that $V_1=\{v_1\}$ is a singleton district, and $v_2\in F_2\subseteq V_2$, where F_2 is a fragment of district V_2. Since not all districts are singletons, we may further assume that $|V_2|\geq 2$. We distinguish between two cases.

Case 1: $F_2\neq V_2$ (i.e., V_2 has two or more fragments). Let e be an arbitrary chord in T_2, and denote the two subtrees of T_2-e by T_2^- and T_2^+ such that v_2 is T_2^-. Since $|V_2|\leq n/k+s$, the subtrees T_2^- and T_2^+ each have at most $n/k+s-1$ vertices. We can recombine V_1 and V_2 into $W_1=V_1\cup V(T_2^-)$ and $W_2=V(T_2^+)$. Then $|W_1|\leq 1+(n/k+s-1)=n/k+s$ and $|W_2|\leq n/k+s-1$; they both induce a connected subgraph of G. As the singleton fragment V_1 and F_2 merge into one fragment of W_1, the number of fragments decreases by at least one.

Case 2: $F_2=V_2$ (i.e., district V_2 has only one fragment). Let $t>2$ be the smallest index such that v_t is in a district that has two or more fragments (such district exists since Π is not canonical). Then the chain $(v_1,\ldots,v_{t-1})$ is covered by single-fragment districts that we denote by $V_1,\ldots,V_\ell$ along C. By recombining V_i and V_{i+1} for $i=1,\ldots,\ell-1$, we create new single-fragment districts $W_1,\ldots,W_\ell$ such that $|W_i|=|V_{i+1}|$ for $i=1,\ldots,\ell+1$ and $|W_\ell|=|V_1|=1$. Now we can apply Case 1 for the singleton district W_ℓ. ☐

We are now ready to prove the main result of this section.

Theorem 2. *If G is a Hamiltonian graph on n vertices and $s\geq n/k$, then $\mathcal{R}_s(G,k)$ is connected and its diameter is $O(nk)$.*

Proof. Based on Lemmas 4–7, the following algorithm successively reduces the number of fragments to k, thereby transforming any balanced k-partition to a canonical partition. While the number of fragments is more than k, do:

1. If a fragment of a small district is adjacent to a light fragment of a large district along C, then apply the recombination move in Lemma 4, which decreases the number of fragments.
2. Else, by Lemma 5, there are two adjacent districts along C whose combined size is at most $2n/k$. Apply a recombination move in Lemma 6. If this move does not decrease the number of fragments, it creates a singleton district, and then up to $k-1$ recombination moves in Lemma 7 decrease the number of fragments by at least one.

There can be at most n different fragments in a k-partition of a set of n vertices. We can reduce the number of fragments using up to k recombination moves. Overall, $O(nk)$ recombination moves can bring any two (k,s)-BCPs to canonical form, which are within k^2+1 moves apart by Lemma 3. ☐

4 Disconnected Configuration Space

In this section we show that the configuration space is not always connected, even in Hamiltonian graphs. Specifically, we show the following result:

Theorem 3. *For any $k \geq 4$ and $s > 0$ there exists a Hamiltonian planar graph G of $n = k(3s+2)$ vertices such that $\mathcal{R}_s(G,k)$ is disconnected.*

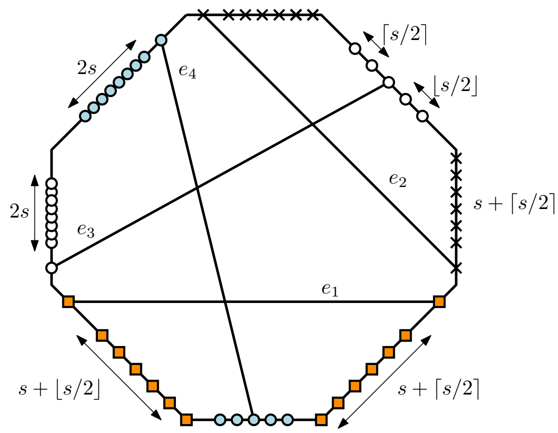

Fig. 2: Problem instance showing that $\mathcal{R}_s(G,k)$ is not always connected (for $k = 4$, $n = 56$ and $s = 4 = \frac{n}{3k} - O(1)$).

Proof Sketch. The proof is constructive and can be found in [17]. We construct an instance that consists of a cycle and 4 chords (shown in Fig. 2). Each district consists of two contiguous arcs along the cycle, which are connected by a unique chord. We prove that no sequence of recombinations can change this fact. Indeed, the chords are sufficiently limiting that a district can only gain/lose vertices in a very restricted fashion (e.g., the district of represented by orange squares can gain up to s vertices from the district of blue circles). $\square$

5 Hardness Results

This section presents our hardness results. Only a sketch of our reductions are included in this extended abstract; see [17] for full details. Our reductions are from Nondeterministic Constrained Logic (NCL) reconfiguration which is PSPACE-complete [10,11]. An instance of NCL is given by a planar cubic undirected graph G_{NCL} where each edge is colored either red or blue. Each vertex is either incident to three blue edges or incident to two red and one blue edges. We respectively call such vertices **OR** and **AND** vertices. An orientation of G_{NCL} must satisfy the constraint that at every vertex $v \in V(G_{NCL})$, at least one blue edge or at least two red edges are oriented towards v. A **move** is an operation that transforms a satisfying orientation to another by reversing the orientation of a single edge. The problem gives two satisfying orientations A and B of G_{NCL} and asks for a sequence of moves to transform A into B. As in [2], we subdivide each edge in G_{NCL} obtaining a bipartite graph G'_{NCL} with one part formed by original vertices in $V(G_{NCL})$ and another part formed by degree-2 vertices. We require that an orientation must additionally satisfy the constraint that each degree-2 vertex v must have an edge oriented towards v. The question of whether there exists a sequence of moves transforming orientation A' into B' of G'_{NCL} remains PSPACE-complete. We follow the framework in [2] with a few crucial differences. The main technical challenge is dealing with the slack constraints while maintaining the desired behavior for the gadgets. We first describe the reduction to instances with slack equals zero, and We then generalize the proofs.

Zero Slack. In the following reduction, we are given a bipartite instance of NCL given by (G'_{NCL}, A', B'), and we produce an instance of $\mathrm{BR}(G, k, s)$ of the balanced recombination problem consisting of two (k, s)-BCP of a planar graph G, Π_A and Π_B, with $k = O(|V(G_{NCL})|)$ districts, and slack $s = 0$. Here we give a brief overview of the reduction. Details can be found in [17].

The AND, OR and degree-2 gadgets are shown in Figures 3 (a), (b) and (c) respectively. The green (black) dots are called **heavy** (**light**) vertices and are considered to be weighed with integer weight more than one (equal to one). We can implement weights by attaching an appropriate number of degree-1 vertices to a heavy vertex so that, in order for a k-BCP to be connected, whichever district contains the heavy vertex must also contain all degree-1 vertices attached to it. Every edge $e \in E(G'_{NCL})$ is represented by two light vertices of G, e^+ and e^-, that belong to two neighboring gadgets as shown in Figures 3 (d).

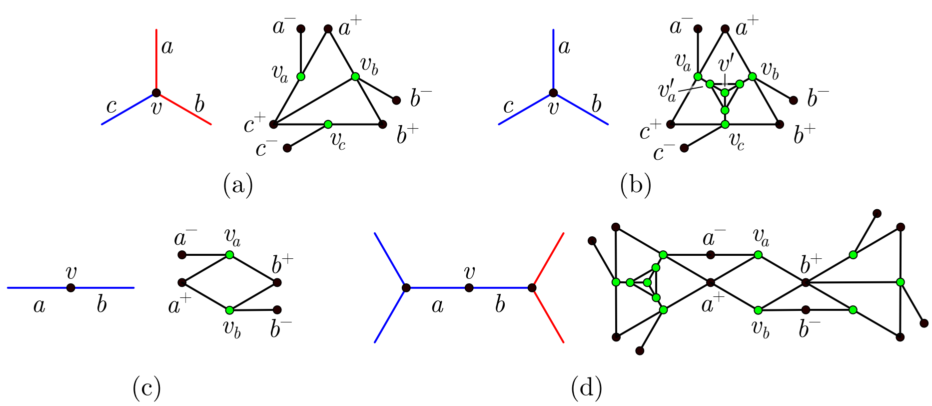

Fig. 3: Gadgets for the reduction from NCL to $\mathrm{BR}(G, k, 0)$.

The weights of heavy vertices are set up so that, for each AND or OR gadget, a district must contain its heavy vertices v_a, v_b and v_c and no heavy vertices of neighbor degree-2 gadgets. Then, for all degree-2 gadget, there is a district that contains v_a and v_b and no other heavy vertex. Additionally, for all OR gadgets, a district must contain v' and exactly one vertex in $\{v'_a, v'_b, v'_c\}$. We encode whether an edge e points toward a vertex by whether a district of the corresponding gadget contains e^+. Then, the connectivity constrains of the districts simulate the NCL constraints. See Figure 4.

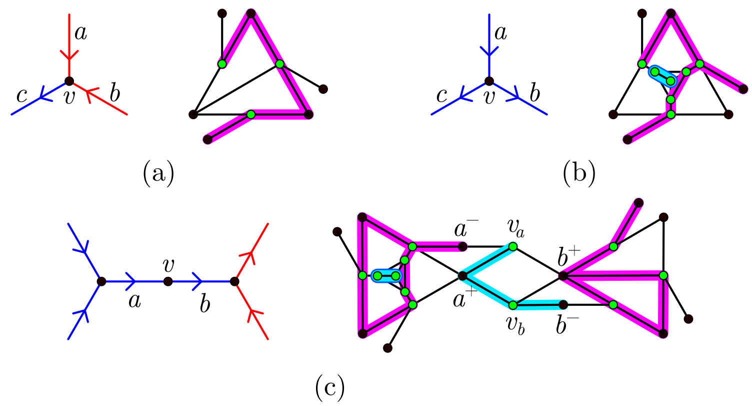

Fig. 4: Equivalence between a satisfying orientation of G'_{NCL} and a k-BCP of G.

Lemma 8. *$BR(G, k, 0)$ is PSPACE-complete even for a planar graph G with constant maximum degree.*

Generalizations. We generalize the reduction of Lemma 8.

Bounded-degree triangulation G. The main new technical tool presented in this section is the **filler gadget** shown in Figure 5 (b). Each face marked with a dot is called a **heavy face** associated with an integer weight, and whose recursive construction is shown in Figure 5 (a). Figure 5 (c) shows how to use copies of the filler gadget to transform G in a triangulation. The main property of the filler gadget is that we set the weights of heavy faces so that each red vertex must belong to a different district and the gadget only intersects 5 districts. Then, such districts are "trapped" in the filler gadget and don't interfere with the other gadgets.

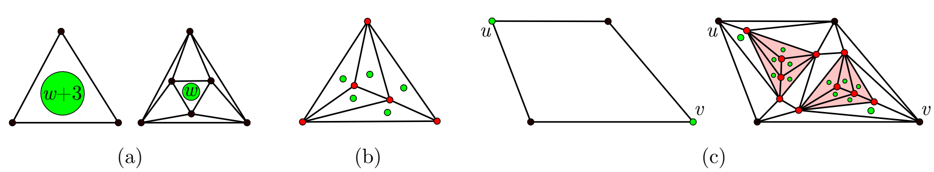

Fig. 5: Construction of the filler gadget.

Theorem 4. *$BR(G, k, s)$ is PSPACE-complete even if G is maximal planar of constant maximum degree, $k \in O(n^{\varepsilon})$, and $s \in O(n^{1-\varepsilon})$ for $0 < \varepsilon \leq 1$.*

Finding Balanced Connected Partitions. The **NCL orientation problem** is defined by an input undirected graph G_{NCL} edge colored as before, and asks whether there exist an orientation of G_{NCL} that satisfies the NCL constraints. This problem is NP-complete [11]. We remark that our construction implies the following theorem.

Theorem 5. *It is NP-complete to decide whether there exist a (k, s)-BCP of a graph G, even if G is maximal planar of constant maximum degree, $k \in O(n^{\varepsilon})$, and $s \in O(n^{1-\varepsilon})$ for $0 < \varepsilon \leq 1$.*

Constant number of districts. The drawback of the previous construction is that it requires 5 new districts for each filler gadget. We obtain PSPACE-hardness with $k = 3$, but we lose the restriction that G is a triangulation, and instead we only require that G is a bounded-degree planar graph. The main technical difficulty is to guarantee that the same subset of heavy vertices is always contained in the same district. This property is obtained by a careful setting of the weights so that there is a unique partition of weights of the heavy vertices that allow for the 3 districts to be balanced within s slack. This allow us to label the districts according to the heavy vertices that is contains. One of the districts then locally acts like the districts that previously contained v_a, v_b and v_3 in each AND and OR gadgets, maintaining the equivalency between the conectedness of this district and the NCL constraints. Our proof can be adapted for any $k \geq 3$.

Theorem 6. *$BR(G, 3, s)$ is PSPACE-complete even if G is planar with constant maximum degree, and $s \in O(n^{1-\varepsilon})$ for $0 < \varepsilon \leq 1$.*

6 Conclusion and Open Problems

We have shown that the configuration space $\mathcal{R}_s(G, k)$ of (k, s)-BCPs is connected when G is connected and $s = \infty$, or when G is Hamiltonian and $s \geq n/k$. We hope that our results inform future research on the properties of G, k, and s that are sufficient to obtain an efficient sampling of $\mathsf{Bal}_s(G, k)$. We also leave it as an open problem whether our results in Section 3 generalize to other classes of graphs. We conjecture that the configuration space $\mathcal{R}_s(n, k)$ is connected for every **biconnected** graph G on n vertices when $s \geq n/k$. However, our techniques do not directly generalize; it is unclear how to extend the notion of canonical k-partitions in the absence of a Hamilton cycle.

We have shown that $\mathrm{BR}(G, k, s)$ is PSPACE-complete even in specific settings that are of interest in applications such as sampling electoral maps. Our results imply that the configuration space $\mathcal{R}_s(G, k)$ has diameter exponential in n, establishing as well an exponential lower bound on the mixing time of a Markov chain on $\mathcal{R}_s(G, k)$ for these settings. We note that Theorems 4 and 6 do not include other settings of interest such as when G is maximal planar (or even 3-connected) and k is a constant. We leave these as open problems.

References

1. Abrishami, T., Guillen, N., Rule, P., Schutzman, Z., Solomon, J., Weighill, T., Wu, S.: Geometry of graph partitions via optimal transport. SIAM J. Sci. Comput. **42**(5), A3340–A3366 (2020). https://doi.org/10.1137/19M1295258
2. Akitaya, H.A., Jones, M.D., Korman, M., Meierfrankenfeld, C., Munje, M.J., Souvaine, D.L., Thramann, M., Tóth, C.D.: Reconfiguration of connected graph partitions. Preprint (2019), http://arxiv.org/abs/1902.10765
3. Chataigner, F., Salgado, L.B., Wakabayashi, Y.: Approximation and inapproximability results on balanced connected partitions of graphs. Discret. Math. Theor. Comput. Sci. **9**(1) (2007), http://dmtcs.episciences.org/384
4. Chlebíková, J.: Approximating the maximally balanced connected partition problem in graphs. Inf. Process. Lett. **60**(5), 223–230 (1996). https://doi.org/10.1016/S0020-0190(96)00175-5
5. DeFord, D.R., Duchin, M., Solomon, J.: Recombination: A family of Markov chains for redistricting. Preprint (2019), http://arxiv.org/abs/1911.05725
6. Duchin, M.: Gerrymandering metrics: How to measure? What's the baseline? Preprint (2018), https://arxiv.org/abs/1801.02064
7. Dyer, M.E., Frieze, A.M.: On the complexity of partitioning graphs into connected subgraphs. Discret. Appl. Math. **10**(2), 139–153 (1985). https://doi.org/10.1016/0166-218X(85)90008-3
8. Fifield, B., Higgins, M., Imai, K., Tarr, A.: Automated redistricting simulation using Markov chain Monte Carlo. Journal of Computational and Graphical Statistics pp. 1–14 (2020). https://doi.org/10.1080/10618600.2020.1739532
9. Győri, E.: On division of graphs to connected subgraphs. In: Combinatorics (Proc. Fifth Hungarian Combinatorial Coll., 1976., Keszthely), p. 485–494. Bolyai — North-Holland (1978)
10. Hearn, R.A., Demaine, E.D.: PSPACE-completeness of sliding-block puzzles and other problems through the nondeterministic constraint logic model of computation. Theoretical Computer Science **343**(1), 72–96 (2005). https://doi.org/10.1016/j.tcs.2005.05.008
11. Hearn, R.A., Demaine, E.D.: Games, Puzzles, and Computation. AK Peters/CRC Press (2009)
12. Ito, T., Zhou, X., Nishizeki, T.: Partitioning a graph of bounded tree-width to connected subgraphs of almost uniform size. J. Discrete Algorithms **4**(1), 142–154 (2006). https://doi.org/10.1016/j.jda.2005.01.005
13. Jerrum, M., Valiant, L.G., Vazirani, V.V.: Random generation of combinatorial structures from a uniform distribution. Theor. Comput. Sci. **43**, 169–188 (1986). https://doi.org/10.1016/0304-3975(86)90174-X
14. Lari, I., Ricca, F., Puerto, J., Scozzari, A.: Partitioning a graph into connected components with fixed centers and optimizing cost-based objective functions or equipartition criteria. Networks **67**(1), 69–81 (2016). https://doi.org/10.1002/net.21661
15. Levin, D.A., Peres, Y.: Markov Chains and Mixing Times. AMS, Providence, RI, 2nd edn. (2017)
16. Lovász, L.: A homology theory for spanning tress of a graph. Acta Mathematica Academiae Scientiarum Hungarica **30**(3), 241–251 (1977). https://doi.org/10.1007/BF01896190
17. andhugo (dot) alves (underscore) akitaya (at) tufts (dot) edu Matias Korman, H.A.A., Korten, O., Souvaine, D.L., Tóth, C.D.: Reconfiguration of connected

graph partitions via recombination. Preprint (2020), https://arxiv.org/abs/2011.07378
18. Najt, L., DeFord, D.R., Solomon, J.: Complexity and geometry of sampling connected graph partitions. Preprint (2019)
19. Soltan, S., Yannakakis, M., Zussman, G.: Doubly balanced connected graph partitioning. ACM Trans. Algorithms **16**(2), 20:1–20:24 (2020). https://doi.org/10.1145/3381419
20. Suzuki, H., Takahashi, N., Nishizeki, T.: A linear algorithm for bipartition of biconnected graphs. Inf. Process. Lett. **33**(5), 227–231 (1990). https://doi.org/10.1016/0020-0190(90)90189-5
21. Tutte, W.T.: A theorem on planar graphs. Transactions of the American Mathematical Society **82**(1), 99–116 (1956)
22. Wada, K., Kawaguchi, K.: Efficient algorithms for tripartitioning triconnected graphs and 3-edge-connected graphs. In: Proc. 19th Workshop on Graph-Theoretic Concepts in Computer Science (WG). LNCS, vol. 790, pp. 132–143. Springer (1993). https://doi.org/10.1007/3-540-57899-4_47
23. Wu, D., Zhang, Z., Wu, W.: Approximation algorithm for the balanced 2-connected k-partition problem. Theor. Comput. Sci. **609**, 627–638 (2016). https://doi.org/10.1016/j.tcs.2015.02.001