

Locally Fair Partitioning

Pankaj K. Agarwal, Shao-Heng Ko, Kamesh Munagala, Erin Taylor

Department of Computer Science, Duke University
{pankaj, sk684, kamesh, ect15}@cs.duke.edu

Abstract

We model the societal task of redistricting political districts as a partitioning problem: Given a set of n points in the plane, each belonging to one of two parties, and a parameter k , our goal is to compute a partition Π of the plane into regions so that each region contains roughly $\sigma = n/k$ points. Π should satisfy a notion of “local” fairness, which is related to the notion of core, a well-studied concept in cooperative game theory. A region is associated with the majority party in that region, and a point is *unhappy* in Π if it belongs to the minority party. A group D of roughly σ contiguous points is called a *deviating* group with respect to Π if majority of points in D are unhappy in Π . The partition Π is *locally fair* if there is no deviating group with respect to Π .

This paper focuses on a restricted case when points lie in 1D. The problem is non-trivial even in this case. We consider both adversarial and “beyond worst-case” settings for this problem. For the former, we characterize the input parameters for which a locally fair partition always exists; we also show that a locally fair partition may not exist for certain parameters. We then consider input models where there are “runs” of red and blue points. For such clustered inputs, we show that a locally fair partition may not exist for certain values of σ , but an approximate locally fair partition exists if we allow some regions to have smaller sizes. We finally present a polynomial-time algorithm for computing a locally fair partition if one exists.

1 Introduction

Redistricting is a common societal decision making problem. In its basic form, there are two parties, say red and blue, and a parliament with some k representatives. Each individual (or voter) in the geographic region is aligned with one of the two parties. The goal is to divide the region into k parts – called districts – so that each part elects one representative to the parliament. It is typically assumed that each district does majority voting, so that if a district has more red voters than blue voters, then the chosen representative will be red.

Copyright © 2022, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

The societal question then is *how should these districts be drawn?* One natural constraint is that the district is a connected region and, more preferably, has a compact shape. Another consideration is that each district is population *balanced*, i.e., has roughly the same number of individuals.¹ A final consideration, and one that will be the focus of this paper, is *fairness*. If society has a large fraction of blue voters, the districts should not be drawn so that most representatives end up being red.

In this paper, we consider a *local* and strong notion of fairness. Let us say that a voter is unhappy if she is in a majority blue (resp. red) district, but her party is red (resp. blue). We say that a given set of districts is *locally fair* if no subset of unhappy voters of the same party can deviate and form a feasible district (nicely shaped and balanced) so that they are the majority in that district. In other words, these deviating voters have a *justified complaint* – there was a different hypothetical district where they could have been happy. This notion is akin to that of the *core* from cooperative game theory (Scarf 1967). As such, if a partition is locally fair, then it is *as fair as possible* to the relevant parties – there are no groups could potentially form a region and do better.

There are examples in which a group of voters have argued they have a justified complaint regarding the redistricting. In the 2012 election in North Carolina, 13 House seats were allocated, 4 to Democrats and 9 to Republicans. In contrast, the percentage of voters who voted for a Democrat candidate was 50.60%. The U.S. Court of Appeals ruled that two of the districts’ boundaries in this map were unconstitutional due to gerrymandering and required new maps to be drawn. Considering this map, it is clear there exist compact potential districts which could be considered a deviating group with respect to the districting. This case (Cooper v. Harris (2017)) is just one in a long line of judgments on the *fairness* of districting plans in the U.S.² The exhibition of deviating groups may help a political group or group of voters justify their complaint that a

¹US courts have ruled that districts be population balanced, compact, and contiguous; see, e.g., https://en.wikisource.org/wiki/Reynolds_v._Sims.

²See also, Benisek v. Lamone (2018), Gill v. Whitford (2018), Rucho v. Common Cause (2019), etc.

redistricting is unfair, and it may be effectively used in auditing proposed plans.

In contrast, some input instances may exhibit “natural gerrymandering”, when the distribution of the population prevents redistricting plans from being representative to all groups (Borodin et al. 2018). For example, if the minority party had 40% of the vote in total but the voters are uniformly distributed, it is unlikely that any deviating groups with a justified complaint would exist. In this case, one could argue no reasonable redistricting could ensure the minority group elects its fair share of the representatives. Thus, the notion of local fairness introduced in the paper allows us to distinguish between natural and artificial gerrymandering. In contrast, when a redistricting plan is not *globally* fair (proportionally representative), it is not clear whether any group has a justified complaint regarding the redistricting, or if it is an unavoidable consequence of the geometry of the map.

Our Results

In this paper, we study the existence and computation of a locally fair partition in the *one-dimensional case*, where we assume the n voters lie on a line or a circle. A *feasible* district or region is now an interval containing $\sigma = n/k$ voters. The “niceness” aspect is captured by the region being an interval, and the “balance” aspect is captured by the number of points in each interval being n/k . Even in this setting, we show that locally fair partitioning is surprisingly non-trivial, and leads to a rich space of algorithmic questions.

Relaxed local fairness notions. In the 1D case, regulating each interval to be containing exactly σ voters is extremely restrictive, and it is relatively easy to show that even a balanced (not necessarily fair) partition need not exist. We will therefore allow ourselves to relax the interval size. We parameterize this by ε , so that the number of voters in any allowable interval lies in $[(1 - \varepsilon)\sigma, (1 + \varepsilon)\sigma]$. This also relaxes the number of intervals to be some number in $[\frac{k}{1+\varepsilon}, \frac{k}{1-\varepsilon}]$.³ Further, we also relax the notion of *deviation*, so that if a subset of voters deviate, they need to become “really happy” – they need to be a strict majority in the interval to which they deviate. We call this parameter $\beta \in [1/2, 1]$, so that unhappy points only deviate to a new allowable interval if their population size is at least $\beta\sigma$. If a fair partition exists under such relaxations, we term it (ε, β) -fair.

Under these relaxations, our first set of results in Section 3 characterizes the (ε, β) values for which a fair partition exists, and those where it may not. For $\varepsilon \leq 1/5$, we show a *sharp threshold* at $\beta = 1 - \varepsilon$: When $\beta < 1 - \varepsilon$, for large enough values of n , there is an instance with no (ε, β) -fair partition, while when $\beta \geq 1 - \varepsilon$, the simple strategy of creating uniform intervals is (ε, β) -fair. If we restrict points to deviate only when the interval

³Many of our results extend to the setting in which the number of intervals must be exactly k .

they create has exactly σ points, this sharp threshold holds for all $\varepsilon \leq 1/3$. To interpret this result, when $\varepsilon = 1/3$, this means there is a fair partition where all intervals have size in the range $[\frac{2}{3}\sigma, \frac{4}{3}\sigma]$, and no subset of unhappy points can create an interval with σ points, where they form 2/3-majority. Furthermore, there is an instance where the bound of 2/3 on the majority cannot be reduced any further.

Beyond worst-case. The negative results above are adversarial: they need careful constructions of sequences of runs of red and blue points with precise lengths, so that any partitioning scheme that needs intervals of certain size to eventually straddle both red and blue points in a way that allows a deviating interval to take shape. However, this is an artifact of the intervals needing almost precise balance, i.e., their lengths being approximately σ . The next question we ask is: suppose we are allowed to place a small fraction α of points in intervals whose sizes can be smaller than $(1 - \varepsilon)\sigma$. In particular, we could construct intervals for these points so that they are all happy, preventing them from deviating; or we could think of it as eliminating these points. Then *is it possible to circumvent these lower bounds?*

In Section 4, we show that the above is indeed the case when the input sequences are reasonably benign. By “benign”, we mean that the input is clustered, i.e., composed of runs of red and blue points of arbitrary lengths, as long as these lengths are lower bounded by some value ℓ . This models phenomena like *Schelling segregation* (Schelling 1971; Zhang 2011; Immorlica et al. 2017), where individuals have a slight preference for like-minded neighbors and relocate to meet this constraint, which leads to “runs” of like-minded individuals.

For such input sequences, we show that as long as all runs are of length at least $\ell = 2\sigma$, once we allow a small fraction $\alpha = O(\frac{1}{k})$ of the points to be placed in unbalanced regions, there is a locally fair partition even for the strictest setting $(\varepsilon, \beta) = (0, 1/2)$: the remaining points are placed in intervals of size exactly σ , and no deviating interval of size σ has a simple majority of unhappy points.

Efficient partitioning. In Section 5, we finally study the algorithmic question: given parameters (ε, β) , decide whether a given input of length n admits a (ε, β) -fair partition. Note that the results so far have been worst-case existential results, and it is possible that even when $\beta < 1 - \varepsilon$, many inputs would have an (ε, β) -fair partition. The challenge in designing an algorithm is that a deviating interval could involve points from more than one interval in the partition. We resolve this via a dynamic programming algorithm whose running time is polynomial in n for any $\varepsilon \in [0, 1/2]$.

Due to length constraints of the paper, many proofs and discussion can be found in the full version (Agarwal et al. 2021).

Related Work

Fairness notions. Proportionality is a classic approach to achieving fairness in social choice. In a proportional solution, different demographic slices of voters feel they have been represented fairly. This general idea dates back more than a century (Droop 1881), and has recently received significant attention (Chamberlin and Courant 1983; Monroe 1995; Brams, Kilgour, and Sanver 2007; Aziz et al. 2017; Sánchez-Fernández et al. 2017; Aziz et al. 2018). In fact, there are several elections, both at a group level and a national level, that attempt to find committees (or parliaments) that provide approximately proportional representation.

The notion of core from cooperative game theory (Scarf 1967) represents the ultimate form of proportionality: every demographic slice of voters feel that they have been fairly represented and do not have incentive to deviate and choose their own solution which gives all of them higher utility. In the typical setting where these demographic slices are not known upfront, the notion of core attempts to be fair to all subsets of voters. Though the core has been traditionally considered in the context of resource allocation problems (Lindahl 1958; Gale and Shapley 1962; Foley 1970; Shapley and Scarf 1974), one of our main contributions is to adapt this notion in a non-trivial way to the redistricting problem.

Redistricting vs. clustering. The redistricting problem is closely related to the clustering problem. In a line of recent work, various models of fairness have been proposed for center-based clustering. One popular approach to fairness ensures that each cluster contains groups in (roughly) the same proportion in which they exist in the population (Chierichetti et al. 2017; Zafar et al. 2017). The redistricting problem we consider may take the opposite view – we effectively want the regions or clusters to be as close to monochromatic as possible to minimize the number of unhappy points in each region.

Chen et al. studied a variant of fair clustering problem where any large enough group of points with respect to the number of clusters are entitled to their own cluster center, if it is closer in distance to all of them (Chen et al. 2019). This extends the notion of the core in a natural way to clustering. However, this work defines happiness of a point in terms of its distance, while in the redistricting problem, the happiness is in terms of the color of the majority within that region. The latter leads to fundamentally different algorithmic questions.

Redistricting Algorithms. There has been extensive work on redistricting algorithms, going back to 1960s (Hess et al. 1965), for constructing contiguous, compact, and balanced districts. Many different approaches, including integer programming (Goderbauer 2014), simulated annealing (Altman and McDonald 2010), evolutionary algorithms (Liu, Cho, and Wang 2016), Voronoi diagram based methods (Svec, Burden, and Dilley 2007; Cohen-Addad, Klein, and Young 2018), MCMC methods (Bangia et al. 2017; DeFord, Duchin, and Solomon 2021), have been proposed; see (Becker and Solomon

2020) for a recent survey. A line of work on redistricting algorithms focuses on combating manipulation such as gerrymandering: when district plans have been engineered to provide advantage to individual candidates or to parties (Borodin et al. 2018). For example, Cohen-Addad et al. propose a districting strategy with desirable geometric properties such as each district being a convex polygon with a small number of sides on average (Cohen-Addad, Klein, and Young 2018). Using similar methods, Wheeler and Klein argue that the political advantage of urban or rural voters tends to be dramatically less than that afforded by district plans used in the real world (Wheeler and Klein 2020). In fact, Chen et al. show that district plans can also have unintentional bias arising from differences in geographic distribution of two parties (Chen and Rodden 2013).

Auditing. Another line of work in redistricting focuses on developing statistical tools to detect gerrymandering given a districting plan (Herschlag et al. 2020). In many redistricting algorithms, existing methods generate maps without explicitly incorporating notions of fairness, but instead focusing on compactness. Popular methods generate an ensemble of plans and compare the number of representatives each party gets in the generated maps with the number received under the actual proposed maps (Becker and Solomon 2020). In practice, political groups use many justifications for whether a plan is *fair*, and our paper offers a new formal model which may be used for auditing – arguing that various plans satisfy properties of fairness (Procaccia and Tucker-Foltz 2021; DeFord, Duchin, and Solomon 2021). Our work in contrast takes a more algorithmic approach – given a natural definition of what a fair redistricting should look like, we show existence and computational results.

2 Preliminaries

Let X be a set of n points in \mathbb{R}^1 , each colored red or blue, and let $\sigma \in [n]$ be a parameter called ideal population size.⁴ We wish to construct a locally fair partition of X into intervals so that all intervals have roughly σ points. Only ordering of points in X really matters, so we describe the input as a (binary) sequence $X = x_1, \dots, x_n$, where $x_i \in \{\text{R}, \text{B}\}$ represents the color of the i -th point on the line.

Define $R := \{i \in [n] \mid x_i = \text{R}\}$ and $B := \{i \in [n] \mid x_i = \text{B}\}$ to be the subset of all red points and blue points, respectively. An *interval* is a contiguous sequence, defined by a pair of integers $i, j \in [n]$ and denoted as either $[i, j]$ (where both points i and j are included) or $(i, j]$ (where only j is included). For an interval $I \subset [n]$, let $|I|$ denote its size, i.e., $|(i, j)| = j - i + 1$, $|(i, j]| = j - i$.

An alternative way to describe the input. Sometimes it is useful to re-describe the input as a series of alternating *maximal* monochromatic intervals. When appropriate, we denote the input as $X =$

⁴We assume points lie on a line for simplicity; our results extend to the case of a ring in a straightforward manner.

$R_1, B_1, R_2, \dots, R_\eta, B_\eta$, where each $R_j \subseteq R$ (resp. $B_j \subseteq B$) is a maximal sequence of red (resp. blue) points, $R_j \neq \emptyset$ for $j > 1$, and $B \neq \emptyset$ for $j < \eta$. For each R_j, B_j , it suffices to specify its size.

Balanced Partition. We are interested in partitioning $[n]$ into pairwise-disjoint intervals, i.e., computing a partition $\Pi = \langle \pi_1, \dots, \pi_T \rangle$, where $T = |\Pi|$, $\pi_t = (i_{t-1}, i_t]$ for all $t \in [T]$, and $0 = i_0 < i_1 < \dots < i_T = n$.

We parameterize the population deviation in an interval using an input parameter ε : For $\varepsilon \in [0, 1/2]$, an interval $\pi_t \in \Pi$ is called ε -allowable (or allowable for brevity) if it satisfies $(1 - \varepsilon)\sigma \leq |\pi_t| \leq (1 + \varepsilon)\sigma$. The partition Π is **balanced** if each of its interval is allowable. Note that a balanced partition may not always exist (take $n = 100$, $\varepsilon = .01$, and $\sigma = 40$). In the remainder of the paper, we assume that σ is chosen such that a balanced partition exists.

For $\varepsilon = 0$, each interval has exactly σ points; and for $\varepsilon = 1/2$, each interval contains between $\frac{\sigma}{2}$ and $\frac{3\sigma}{2}$ points. In principle, we can choose ε to be any value in $[0, 1]$; but as the value of ε increases, the sizes of intervals in the partition become increasingly unbalanced. At an extreme, when $\varepsilon = 1$, every point could form its allowable interval. Thus, we only consider the setting of $\varepsilon \in [0, 1/2]$, though most of our results extend to settings of larger ε .

For an interval $I \subseteq [n]$, it is sometimes more convenient to work with the *ratio* $|I|/\sigma$, which is 1 for an interval of σ points. We define the *measure* of I , denoted by $\|I\|$, as $\|I\| = |I|/\sigma$.

Locally fair partition. Next, we turn our focus to defining the notion of local fairness. For an interval I , let $\chi(I)$ represent its majority color. Formally, $\chi(I) = R$ if the interval I has red majority, i.e. $|R \cap I| > |B \cap I|$. Similarly, set $\chi(I) = B$ if $|R \cap I| < |B \cap I|$. Without loss of generality, if $|R \cap I| = |B \cap I|$, i.e., there is no majority in I , we define $\chi(I) = B$.

A point i is *happy* in Π if it is assigned to an interval matching its color, i.e., $\chi(\pi_t) = x_i$, where $\pi_t \in \Pi$ is the interval containing i ; otherwise i is *unhappy* in Π . For an interval $I \subseteq [n]$, let $\text{uhp}(I, \Pi) := \{i \in I \mid i \text{ is unhappy in } \Pi\}$ denote the subset of unhappy points in I in partition Π . Note that here I can be any interval and is not necessarily a part in Π . If Π is fixed or clear from the context, we write $\text{uhp}(I) := \text{uhp}(I, \Pi)$.

Intuitively, a partition Π is *locally fair* if there is no large set of unhappy points which could form an allowable interval in which they would be the majority. We make this concept more precise below:

Definition 1. Given an input instance (X, σ) and a parameter $\beta \in [\frac{1}{2}, 1]$, a **β -deviating group** with respect to a partition Π is an allowable interval D with more than $\max\left\{\frac{|D|}{2}, \beta\sigma\right\}$ unhappy monochromatic points, or equivalently, $\max\{\|\text{uhp}(D) \cap R\|, \|\text{uhp}(D) \cap B\|\} > \max\left\{\frac{\|D\|}{2}, \beta\right\}$. Note that we do not require the existence of a balanced partition with D being one of its

intervals.

D is called a β -deviating group because D may deviate from Π such that at least $\beta\sigma$ points that were unhappy in Π become happy if D was made a standalone part in another partition. We sometimes omit β and use the term “deviating group” when the context is clear. Intuitively, β controls how difficult it is for a set of unhappy points to deviate and form an interval in which they are happy. As β grows, the set of unhappy points that may form a deviating group must grow larger with respect to the desired interval size σ ; to deviate when $\beta = 1/2$, a set of more than $\sigma/2$ unhappy points must lie in an allowable interval in which they are the majority. At $\beta = 1$, the number of unhappy points required to form a deviating group increases to σ . See Figure 1 for an example.

Definition 2. Given (X, σ) , $\varepsilon \in [0, 1/2]$, and $\beta \in [1/2, 1]$, we call a balanced partition Π (ε, β) -locally fair if there is no β -deviating group with respect to Π .

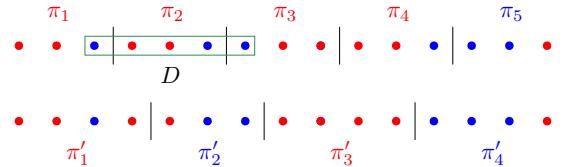


Figure 1: An with $n = 15$ points, $\sigma = 4$, $\varepsilon = 1/2$, and $\beta = 2/3$. The top partition $\{\pi_1, \dots, \pi_5\}$ admits a blue deviating group D , whereas the bottom partition $\{\pi'_1, \dots, \pi'_4\}$ is $(1/2, 2/3)$ -locally fair.

Remark. While we focus on $\beta \in [1/2, 1]$, the largest possible range to consider is $\beta \in [(1 - \varepsilon)/2, (1 + \varepsilon)]$. For applications of our model, we would expect the requirement on β to be stronger than a simple majority, so that we bias towards solutions (partitions) which are presumably fair to the rest of the points.

3 Existence of Locally Fair Partitions

In this section, we present our results on the existence of locally fair partitions. We first give characterizations of parameters ε and β for which a locally fair solution is guaranteed to exist. We show that for every $\varepsilon \in [0, 1/2]$, there is a threshold $\bar{\beta}(\varepsilon)$ such that if β is above this threshold, a simple partitioning strategy into small intervals results in a (ε, β) -fair partition.

Theorem 3. Let (X, σ) be an input instance as defined above. For any $\varepsilon \in [0, 1/2]$, there is a value $\bar{\beta}(\varepsilon)$ such that for any $\beta > \bar{\beta}(\varepsilon)$, there exists an (ε, β) -fair partition, where

$$\bar{\beta}(\varepsilon) = \begin{cases} \max\left\{1 - \varepsilon, \frac{1+3\varepsilon}{2}\right\} + O(\delta) & \text{for } \varepsilon \in [0, 1/3], \\ \max\left\{\frac{3(1-\varepsilon)}{2}, 2\varepsilon\right\} + O(\delta) & \text{for } \varepsilon \in [1/3, 1/2], \end{cases}$$

$$\text{and } \delta \leq \frac{1}{\frac{n}{\sigma} - 1} + \frac{1}{\sigma}.$$

Proof sketch. (See the full version for a detailed proof.) We construct a partition Π that uses as small as possible intervals (the length of each interval takes the form $(1 - \varepsilon + \delta)\sigma$). For $\varepsilon \leq 1/3$, an allowable deviating group D could intersect at most 3 intervals. We show if D intersects only 2 intervals, a simple calculation shows there are less than $\beta\sigma$ unhappy points and such a D cannot exist. On the other hand, if D intersects 3 intervals $\pi_i, \pi_{i+1}, \pi_{i+2}$, it must completely contain π_{i+1} . We can bound the number of points D can pull from $|\pi_i \cup \pi_{i+2}|$ as $(1 + \varepsilon)\sigma - |\pi_{i+1}|$, all of which could be unhappy. Additionally, it must be $|\text{uhp}(\pi_{i+1})| \leq |\pi_{i+1}/2|$. Combining the two above observations shows the total number of unhappy points is less than $\beta\sigma$, and so no deviating group exists. A similar proof holds for $\varepsilon \in [1/3, 1/2]$ by increasing the number of intervals the deviating group could intersect to 4.

Theorem 3 can be extended for any $\varepsilon \in [0, 1]$. Following the same proof approach gives the general form of $\bar{\beta}(\varepsilon) = \max \left\{ \frac{t(1-\varepsilon)}{2}, \frac{(3-t)+(t+1)\varepsilon}{2} \right\} + O(t\sigma)$, where t is an integer such that $\varepsilon \in \left[\frac{t-2}{t}, \frac{t-1}{t+1} \right]$, or in other words, $(t+1)$ is the largest number of intervals a deviating group can intersect.

Next, we show that for smaller values of β , a locally fair partition may not exist.

Theorem 4. Let $\varepsilon \in [0, 1/2)$ and $\beta \in [1/2, 1 - \varepsilon)$. For any $\sigma \geq 1$, there exists an input instance (X, σ) with $|X| = O\left(\frac{\beta\sigma}{1-\varepsilon-\beta}\right)$ for which no (ε, β) -locally fair partition exists.

Proof. We construct an instance for which there is always a deviating group. For simplicity, assume $\beta\sigma$ is an integer; we will relax this assumption later. Specifying the input using the runs of monochromatic intervals, let $X = R_1, B_1, R_2, B_2, \dots, R_\eta, B_\eta$ for $\eta = \left\lceil \frac{n}{2\beta\sigma} \right\rceil$. Set $|R_j| = \beta\sigma$ for $j = 1, \dots, \eta$ and $|B_j| = \beta\sigma$ for $j = 1, \dots, \eta - 1$, and $|B_\eta| = n - (2\eta - 1)\beta\sigma \leq \beta\sigma$.

In any fair partition Π , each R_j (resp. B_j) must intersect a red (resp. blue) interval of Π ; if there exists an entire R_j (resp. B_j) contained in a blue (resp. red) interval of Π , R_j (resp. B_j) could form a deviating group with its own $\beta\sigma$ unhappy points, and $(1 - \varepsilon - \beta)\sigma$ points from a neighboring B_j (resp. R_j). Since $\beta \geq 1/2$, these unhappy red (resp. blue) points are the majority in the deviating group. As a consequence, in any fair partition Π , there exists no red (resp. blue) interval π_t that intersects multiple monochromatic red intervals $R_j, R_{j'}$ (resp. blue intervals $B_j, B_{j'}$) in X .

Suppose there exists a fair partition Π for (X, σ) , and let π_1 be a red interval of Π that intersects R_1 . Since $\beta < 1 - \varepsilon$, π_1 must include points from a neighboring blue monochromatic interval; without loss

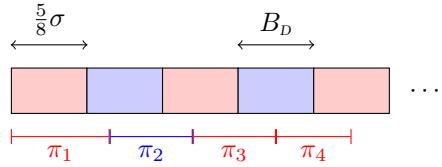


Figure 2: An instance so that each monochromatic interval has length $5\sigma/8$, which does not admit a $(\frac{1}{4}, \frac{5}{8})$ -fair partition. For example, partition Π is made of intervals of length $(1 - \varepsilon)\sigma = 3\sigma/4$; however, B_D forms a deviating group by pulling in points from neighboring intervals.

of generality, let it intersect with B_1 .⁵ Then π_1 cannot intersect R_2 ; otherwise B_1 would deviate. Therefore, π_1 includes at most $\beta\sigma$ points from R_1 and some points from B_1 . Now, consider interval π_2 of Π which is blue and intersects B_1 (but not B_2). By size constraints, $\|\pi_2\| \geq (1 - \varepsilon)$, and $\|\pi_2 \cap B_1\| < \beta$, since $\pi_1 \cap B_1 \neq \emptyset$. This implies $\|\pi_2 \cap R_2\| > (1 - \varepsilon - \beta)$. Next, interval π_3 is red-majority and intersects R_2 (but not R_3); moreover, we have $\|\pi_3 \cap R_2\| \leq \|R_2\| - \|\pi_2 \cap R_2\| < \beta - (1 - \varepsilon - \beta)$. This then implies $\|\pi_3 \cap B_2\| \geq \|\pi_3\| - \|\pi_3 \cap R_2\| > (1 - \varepsilon) - (\beta - (1 - \varepsilon - \beta)) = 2(1 - \varepsilon - \beta)$.

Continuing this argument, it can be shown that (i) every R_j must intersect a red interval in Π that also intersects with B_j ; and (ii) every B_j must intersect a blue interval in Π that also intersects with R_{j+1} . Combined with the fact that each red- (resp. blue-) majority π_t cannot intersect multiple monochromatic red (resp. blue) intervals of X , for every j it must hold that:

- π_{2j-1} is red-majority and intersects R_j ;
- π_{2j} is blue-majority and intersects B_j ;
- $\|\pi_{2j-1} \cap R_j\| < \beta - (2j - 3) \cdot (1 - \varepsilon - \beta)$;
- $\|\pi_{2j-1} \cap B_j\| > (2j - 2) \cdot (1 - \varepsilon - \beta)$;
- $\|\pi_{2j} \cap B_j\| < \beta - (2j - 2) \cdot (1 - \varepsilon - \beta)$;
- $\|\pi_{2j} \cap R_{j+1}\| > (2j - 1) \cdot (1 - \varepsilon - \beta)$.

Denote $j' = \left\lceil \frac{3-3\varepsilon-2\beta}{4(1-\varepsilon-\beta)} \right\rceil$, and assume $\eta \geq j'$. By the above, $\pi_{2j'}$ is blue-majority, and we have $\|\pi_{2j'} \cap B_{j'}\| < \beta - (2j' - 2) \cdot (1 - \varepsilon - \beta) < \frac{1-\varepsilon}{2}$, which implies $\pi_{2j'}$ cannot be blue-majority, a contradiction. In other words, there are not enough points in $\pi_{2j'}$ to create a majority matching its color. Since $\eta = \left\lceil \frac{n}{2\beta\sigma} \right\rceil$, the above holds for $\frac{n}{2\beta\sigma} > \frac{3-3\varepsilon-2\beta}{4(1-\varepsilon-\beta)}$, or $n > \frac{(3-3\varepsilon-2\beta)\beta\sigma}{2(1-\varepsilon-\beta)} = O\left(\frac{\beta\sigma}{1-\varepsilon-\beta}\right)$. Finally, if $\beta\sigma$ is not an integer, let $\beta' = \frac{\lfloor \beta\sigma \rfloor}{\sigma}$. Then the above argument still holds for $n > \frac{(3-3\varepsilon-2\beta')\beta'\sigma}{2(1-\varepsilon-\beta')}$. \square

See Figure 2 for an example of the construction. For $\varepsilon \in [0, 1/5]$, Theorem 3 and Theorem 4 provide an almost

⁵We assume the input lies on a circle. Since π_1 must intersect at least one of the two blue monochromatic intervals neighboring R_1 , we can order the input so that the intersected interval is B_1 .

sharp threshold: if $\beta > 1 - \varepsilon + O(\delta)$ (for δ defined in Theorem 3), a locally fair partition always exists, but for $\beta < 1 - \varepsilon$ there are instances that do not admit fair partitions. In fact, if we enforce the deviating group to have exactly σ points, Theorems 3 and 4 become almost tight for all $\varepsilon \in [0, 1/3]$. We next extend Theorem 4 so that a single instance X has no locally fair partition for a wide range of σ values.

Corollary 5. *Let $\varepsilon \in [0, 1/2)$, $\beta \in [1/2, 1 - \varepsilon)$, and let $\mathcal{S} = \{\sigma_1, \sigma_2, \dots, \sigma_M\}$ be the set of desired interval sizes. If $\frac{n}{M\sigma_m} > \left\lceil \frac{1}{1-\varepsilon-\beta} \right\rceil$ holds for all $m \in [M]$, there exists an input X such that (i) $|X| = n$; (ii) for all $m = 1, \dots, M$, the instance (X, σ_m) has no (ε, β) -locally fair partition.*

4 Clustered Instances

As manifested in the previous section, under many specific parameters ε, β , there exist adversarial input instances (X, σ) that rule out the existence of any locally fair partition. However, such negative instances often seem artificial, and are not robust to perturbation. In this section, we turn our attention to a category of interesting inputs, when points are “clustered” into large monochromatic intervals. Such instances arise in applications in which we expect points of the same color to gather together. We show that fair partitions exist when the input instance is comprised of large monochromatic intervals, while incurring a small approximation on the balancedness of the fair partition.

For a constant $\alpha \in [0, 1)$, we say a partition Π of X is **α -balanced** if the union of all its allowable intervals make up at least a $(1 - \alpha)$ -fraction of the total input. Formally, let $\tilde{\Pi} := \{\pi_t \in \Pi \mid |\pi_t| \in [(1 - \varepsilon)\sigma, (1 + \varepsilon)\sigma]\}$ be the set of allowable intervals in Π . Then Π is α -balanced if $|\bigcup_{\pi_t \in \tilde{\Pi}} \pi_t| \geq (1 - \alpha)n$.

In fact, in this section our results hold for any $\beta \in [1/2, 1]$, so we omit β as a parameter, and instead refer to a fair partition as ε -locally fair.

First, we show that if the size of each monochromatic interval in X is at least 2σ , we can compute a fair partition by letting allowable intervals not contain a small fraction of the population.

Theorem 6. *Given instance $(X = R_1, B_1, R_2, \dots, \sigma)$ with $\|R_j\|, \|B_j\| \geq 2$, for all j and parameter $\varepsilon \in [0, 1/2]$, there is a $\left(\frac{(1-\varepsilon)\sigma}{n}\right)$ -balanced, ε -locally fair partition Π .*

Proof. We assume that the first monochromatic interval R_1 is red and the last monochromatic interval B_η is blue. The other cases follow analogously. Consider an arbitrary maximal monochromatic red interval R_j with measure $\|R_j\| \geq 2$. We divide R_j into allowable intervals such that the residual interval R'_j (points of R_j not assigned to an allowable interval) is as small as possible. Note that $\|R'_j\| \in [0, 1 - \varepsilon]$. We assign the points of R'_j to a size $\lceil(1 - \varepsilon)\sigma\rceil$ interval using $\lceil((1 - \varepsilon) - \|R'_j\|)\sigma\rceil$ points

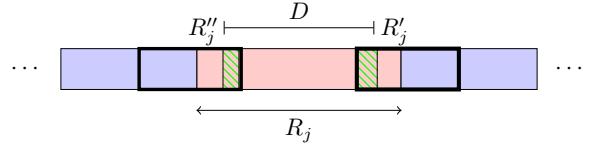


Figure 3: No deviating group D intersects R''_j and R'_j .

from the next interval, B_j . Partition the entire instance in this manner and let Π be the resulting partition. All intervals of Π are allowable except for the residual interval B'_η from the last monochromatic interval B_η . Since $|B'_\eta| < (1 - \varepsilon)\sigma$, Π is $\left(\frac{(1-\varepsilon)\sigma}{n}\right)$ -balanced.

We next show Π is locally fair regardless of the value of ε . Suppose there exists a deviating group D . Without loss of generality, assume D is red-majority and D intersects two consecutive red intervals R_j and R_{j+1} of X . Then we have $B_{j+1} \subset D$. Since $\|B_{j+1}\| \geq 2$, we have $\|D\| > 2 > (1 + \varepsilon)$, a violation of the size constraint. Hence, D can only intersect one monochromatic red interval R_j for some j .

Define $R''_j \subseteq R_j$ (resp. $R'_j \subseteq R_j$) to be the (not necessarily non-empty) set of red points assigned to the same interval, denoted by π'' (resp. π') as a subset B'_j of B_j (resp. B'_{j+1} of B_{j+1}) (See Figure 3). If π' is blue majority, then $\|R''_j\| < \frac{1-\varepsilon}{2}$, by construction. Similarly, if π' is blue majority, then $\|R'_j\| < \frac{1-\varepsilon}{2}$. Assume D intersects both R'_j and R''_j . Then both π' and π'' need to be blue-majority, and we have $R \setminus (R'_j \cup R''_j) \subset D$. But this implies $\|R \setminus (R'_j \cup R''_j)\| > 2 - \frac{1-\varepsilon}{2} - \frac{1-\varepsilon}{2} = 1 + \varepsilon$ and thus $\|D\| > 1 + \varepsilon$, a contradiction.

Hence, D can only intersect either R'_j or R''_j . In both cases, $\|\text{uhp}(D)\| < \frac{1-\varepsilon}{2}$, implying D does not have sufficient points to form a deviating group. Therefore, no deviating group exists in Π . \square

In fact, we can use the same partitioning strategy to find balanced fair partitions (i.e., every point belongs to an allowable interval π_t) if each monochromatic interval of an instance has size at least $\left\lceil \frac{(1-\varepsilon)^2}{2\varepsilon} \right\rceil$.

Corollary 7. *For an instance (X, σ) and parameter ε , such that for all j : $\|R_j\|, \|B_j\| \geq \left\lceil \frac{(1-\varepsilon)^2}{2\varepsilon} \right\rceil$, there is always a balanced ε -locally fair partition Π of X .*

Next, we relax the requirement that *every* monochromatic interval is long and consider “mostly” clustered instances. Specifically, assume that at most γn points of X lie in monochromatic intervals smaller than size 2σ . Applying Theorem 6 to this setting, we can construct an α -balanced fair partition with α depending on γ .

Theorem 8. *Given (X, σ) and parameter $\varepsilon \in [0, 1/2]$, where $X = R_1, B_1, \dots, R_\eta, B_\eta$, let Y denote the set of monochromatic intervals of X of length smaller than 2σ : $Y := \{I \in \{R_1, \dots, R_\eta, B_1, \dots, B_\eta\} \mid \|I\| < 2\}$.*

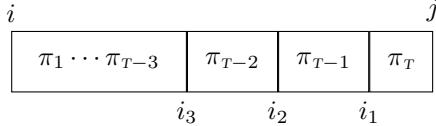


Figure 4: Π is a partition of interval $(i, j]$, where i_1, i_2 , and i_3 are the last three boundaries.

If $|\bigcup_{I \in Y} I| \leq \gamma n$, then there is a $\left(\frac{(1-\varepsilon)\sigma}{n} + \gamma\right)$ -balanced, ε -locally fair partition.

For all ε, α improves (i.e., decreases) as γ decreases, as more of the input lies in larger monochromatic intervals. Similar to Corollary 7, if $\varepsilon \geq 3 - 2\sqrt{2}$, we can prove the above process gives a γ -balanced, ε -locally fair partition.

5 Partitioning Algorithm

Finally, we shift our focus to following algorithmic question: *Given an instance (X, σ) and parameters $\varepsilon \in [0, 1/2], \beta \in [1/2, 1]$, does a (ε, β) -locally fair solution exist for (X, σ) ?*

In this section, we focus on a fixed input instance, so throughout this section, treat X, σ, ε , and β as fixed, and describe an algorithm that determines whether an (ε, β) -locally fair balanced partition, possibly with additional constraints, exists, for an interval $I \subseteq [n]$ of the input, where $|X| = n$.

We now define the recursive subproblems. For any interval $I = (i, j] \in [n]$ and for $i \leq i_3 \leq i_2 \leq i_1 < j$, define $\text{LF}(I, i_1, i_2, i_3) = \text{True}$ if and only if there exists at least one fair partition $\Pi = \{\pi_1, \dots, \pi_T\}$ of I that satisfies the following conditions:

- $i_1 = i_2 = i_3 = i$, $\Pi = \{(i, j]\}$, for $T = 1$;
- $i_1 > i_2 = i_3 = i$, $\Pi = \{(i, i_1], (i_1, j]\}$, for $T = 2$;
- $i_1 > i_2 > i_3 = i$, $\Pi = \{(i, i_2], (i_2, i_1], (i_1, j]\}$, for $T = 3$;
- $i_1 > i_2 > i_3 > i$, $\Pi = \{\pi_1, \dots, (i_3, i_2], (i_2, i_1], (i_1, j]\}$, o/w.

In other words, i_1, i_2 , and i_3 are the last three “interval boundaries” in Π ; see Figure 4.

We first define the base cases of our algorithm. Consider every interval $I = (i, j]$ that is allowable, i.e., $(1 - \varepsilon)\sigma \leq |I| \leq (1 + \varepsilon)\sigma$. Without loss of generality, let $\chi(I) = \mathbb{B}$. We consider letting $\Pi = \{(i, j]\}$ be the trivial partition for I . This is locally fair if and only if no deviating group can form within I , i.e., there exists no $(i', j'] \subset (i, j]$ such that (i) $(1 - \varepsilon)\sigma \leq (j' - i')$ (so that $(i', j']$ is allowable), and (ii) $|\text{uhp}((i', j'] \cap R)| \geq \max\{\frac{\beta\sigma}{2}, \frac{j' - i'}{2}\}$ (so that $(i', j']$ is deviating). If the above holds, we have $\text{LF}(I, i, i, i) = \text{True}$, and $\text{LF}(I, i_1, i_2, i_3) = \text{False}$ for any other values of i_1, i_2 , and i_3 .

Intuitively, if an interval $I = (i, j] \subseteq [n]$ has a fair partition, either I is a standalone fair allowable interval (i.e., $\text{LF}((i, j], i, i, i) = \text{True}$), or there must exist a point $j' \in [j - (1 + \varepsilon)\sigma, j - (1 - \varepsilon)\sigma]$ such that

- (i) there is a fair partition for $(i, j']$;
- (ii) $(j', j]$ is a fair unpartitioned standalone interval;

(iii) no deviating group is formed by the last 3 intervals of fair partition of $(i, j']$ and the interval $(j', j]$.

Accordingly, we have the following lemma:

Lemma 9. $\text{LF}((i, j], i_1, i_2, i_3) = \text{True}$ if and only if there exists a point $i_4 \in [\max\{i, i_3 - (1 + \varepsilon)\sigma\}, i_1]$ such that

- $\text{LF}((i_1, j], i_1, i_1, i_1) = \text{True}$, and
- $\text{LF}((i, i_1], i_2, i_3, i_4) = \text{True}$, and
- There is no deviating group forming within $(i_4, j]$ with the partition $\{(i_4, i_3], (i_3, i_2], (i_2, i_1], (i_1, j]\}$.

Hence, for a general interval $(i, j]$, to compute $\text{LF}((i, j], i_1, i_2, i_3)$, it suffices to check for all possible values of i_4 , each incurring two lookups of previously computed subquery results and one check of fairness in a partition for an interval of length at most $4(1 + \varepsilon)\sigma$.

Algorithm. We use Lemma 9 and dynamic programming to compute $\bigvee_{1 \leq i_1 \leq i_2 \leq i_3 \leq n} \text{LF}([0, n], i_1, i_2, i_3)$, as follows. Our algorithm first enumerates all possible allowable intervals $(i, j]$ as base cases (i.e., $(1 - \varepsilon)\sigma \leq j - i \leq (1 + \varepsilon)\sigma$), and computes $\text{LF}((i, j], i, i, i)$ for each such $(i, j]$. Then, for general $(i, j]$, the algorithm computes $\text{LF}((i, j], i_1, i_2, i_3)$ for all possible values of i_1, i_2 , and i_3 given i and j , in increasing order of $(i + j)$; this ensures all the intermediate subqueries are already computed before $\text{LF}((i, j], i_1, i_2, i_3)$ is evaluated. After it computes $\text{LF}((i, j], i_1, i_2, i_3)$ for all possible values of i, j, i_1, i_2, i_3 , it examines whether $\text{LF}((0, n], i_1, i_2, i_3) = \text{True}$ for some i_1, i_2, i_3 ; any such true entry implies a fair partition of $[n]$, i.e., the original input. Note that $(0, n] = [1, n]$ is the complete set of points.

Running time and Enhancements. Refer to the full version for a detailed analysis of the algorithm, as well as two enhancement strategies that exploit precomputation to avoid redundant computations. We present the final result:

Theorem 10. Given an instance (X, σ) with $|X| = n$ and $\sigma \in [n]$, and parameters $\varepsilon \in [0, 1/2]$ and $\beta \in [1/2, 1]$, a (ε, β) -locally fair partition of $[n]$ can be computed, or report that none exists, in time $O(n\sigma^3)$ for $\varepsilon \in [0, 1/3]$ and $O(n\sigma^4)$ for $\varepsilon \in [1/3, 1/2]$.

6 Conclusion

The main open question is extending the model to two dimensions, which poses several challenges: how to define feasible regions, how these regions tile the plane, how a deviating group/region is defined. Part of the difficulty stems from the fact that points are no longer linearly ordered. Despite the challenges, the lower bounds presented in the paper directly extend to two dimensions. Additionally, the algorithm described in Section 5 may extend to 2D when the partitions considered have sufficient structure, such as when we restrict to a hierarchical partition of simple shapes, and the deviating region spans $O(1)$ regions of the partition.

Acknowledgments

The authors would like to thank Hsien-Chih Chang, Brandon Fain, and anonymous reviewers for helpful discussions and feedback. This work is supported by NSF grants CCF-2113798, IIS-18-14493, and CCF-20-07556 and ONR award N00014-19-1-2268.

References

- Agarwal, P. K.; Ko, S.-H.; Munagala, K.; and Taylor, E. 2021. Locally Fair Partitioning. <https://arxiv.org/abs/2112.06899>. arXiv:2112.06899.
- Altman, M.; and McDonald, M. 2010. The promise and perils of computers in redistricting. *Duke Journal of Constitutional Law & Public Policy*, 5: 69.
- Aziz, H.; Brill, M.; Conitzer, V.; Elkind, E.; Freeman, R.; and Walsh, T. 2017. Justified representation in approval-based committee voting. *Social Choice and Welfare*, 48(2): 461–485.
- Aziz, H.; Elkind, E.; Huang, S.; Lackner, M.; Fernández, L. S.; and Skowron, P. 2018. On the complexity of extended and proportional justified representation. In *AAAI*, 902–909.
- Bangia, S.; Graves, C. V.; Herschlag, G.; Kang, H. S.; Luo, J.; Mattingly, J. C.; and Ravier, R. 2017. Redistricting: Drawing the line. arXiv:1704.03360.
- Becker, A.; and Solomon, J. 2020. Redistricting algorithms. arXiv:2011.09504.
- Borodin, A.; Lev, O.; Shah, N.; and Strangway, T. 2018. Big city vs. the great outdoors: voter distribution and how it affects gerrymandering. In *IJCAI*, 98–104.
- Brams, S. J.; Kilgour, D. M.; and Sanver, M. R. 2007. A minimax procedure for electing committees. *Public Choice*, 132(3): 401–420.
- Chamberlin, J. R.; and Courant, P. N. 1983. Representative deliberations and representative decisions: Proportional representation and the Borda rule. *The American Political Science Review*, 77(3): 718–733.
- Chen, J.; and Rodden, J. 2013. Unintentional gerrymandering: Political geography and electoral bias in legislatures. *Quarterly Journal of Political Science*, 8(3): 239–269.
- Chen, X.; Fain, B.; Lyu, L.; and Munagala, K. 2019. Proportionally fair clustering. In *ICML*, 1032–1041.
- Chierichetti, F.; Kumar, R.; Lattanzi, S.; and Vassilvitskii, S. 2017. Fair clustering through fairlets. In *NeurIPS*, 5029–5037.
- Cohen-Addad, V.; Klein, P. N.; and Young, N. E. 2018. Balanced centroidal power diagrams for redistricting. In *ACM SIGSPATIAL*, 389–396.
- DeFord, D.; Duchin, M.; and Solomon, J. 2021. Recombination: a family of Markov chains for redistricting. *Harvard Data Science Review*.
- Droop, H. R. 1881. On methods of electing representatives. *Journal of the Statistical Society of London*, 44(2): 141–202.
- Foley, D. K. 1970. Lindahl’s solution and the core of an economy with public goods. *Econometrica*, 66–72.
- Gale, D.; and Shapley, L. S. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1): 9–15.
- Goderbauer, S. 2014. Political districting for elections to the German Bundestag: an optimization-based multi-stage heuristic respecting administrative boundaries. In *Operations Research*, 181–187.
- Herschlag, G.; Kang, H. S.; Luo, J.; Graves, C. V.; Bangia, S.; Ravier, R.; and Mattingly, J. C. 2020. Quantifying gerrymandering in North Carolina. *Statistics and Public Policy*, 7(1): 30–38.
- Hess, S. W.; Weaver, J.; Siegfeldt, H.; Whelan, J.; and Zitzlau, P. 1965. Nonpartisan political redistricting by computer. *Operations Research*, 13(6): 998–1006.
- Immorlica, N.; Kleinberg, R.; Lucier, B.; and Zadomaghaddam, M. 2017. Exponential segregation in a two-dimensional Schelling model with tolerant individuals. In *ACM-SIAM SODA*, 984–993.
- Lindahl, E. 1958. Just taxation—a positive solution. In *Classics in the Theory of Public Finance*, 168–176. Springer.
- Liu, Y. Y.; Cho, W. K. T.; and Wang, S. 2016. PEAR: a massively parallel evolutionary computation approach for political redistricting optimization and analysis. *Swarm and Evolutionary Computation*, 30: 78–92.
- Monroe, B. L. 1995. Fully proportional representation. *The American Political Science Review*, 89(4): 925–940.
- Procaccia, A. D.; and Tucker-Foltz, J. 2021. Compact Redistricting Plans Have Many Spanning Trees. arXiv:2109.13394.
- Sánchez-Fernández, L.; Elkind, E.; Lackner, M.; Fernández, N.; Fisteus, J. A.; Basanta Val, P.; and Skowron, P. 2017. Proportional justified representation. In *AAAI*, 670–676.
- Scarf, H. E. 1967. The core of an N person game. *Econometrica: Journal of the Econometric Society*, 50–69.
- Schelling, T. C. 1971. Dynamic models of segregation. *The Journal of Mathematical Sociology*, 1(2): 143–186.
- Shapley, L.; and Scarf, H. 1974. On cores and indivisibility. *Journal of Mathematical Economics*, 1(1): 23–37.
- Svec, L.; Burden, S.; and Dilley, A. 2007. Applying Voronoi diagrams to the redistricting problem. *The UMAP journal*, 28(3): 313–329.
- Wheeler, A.; and Klein, P. N. 2020. The impact of highly compact algorithmic redistricting on the rural-versus-urban balance. In *ACM SIGSPATIAL*, 397–400.
- Zafar, M. B.; Valera, I.; Gomez Rodriguez, M.; and Gummadi, K. P. 2017. Fairness beyond disparate treatment & disparate impact: Learning classification without disparate mistreatment. In *ACM WWW*, 1171–1180.
- Zhang, J. 2011. Tipping and residential segregation: a unified Schelling model. *Journal of Regional Science*, 51(1): 167–193.