How Does Fairness Affect the Complexity of Gerrymandering?

Sandip Banerjee [1] Rajesh Chitnis [2] Abhiruk Lahiri [3]

[1]Dalle Molle Institute for Artificial Intelligence, Switzerland [2]University of Birmingham, United Kingdom [3]Heinrich Heine University Düsseldorf, Germany

Gerrymandering

The term "Gerrymandering" was coined after Elbridge Gerry, the Governor of Massachusetts in 1812. It has been used on several occasions to manipulate elections.

The process "Gerrymandering" is generally done in the following two ways:

- **Packing:** Drawing lines to include the maximum number of opposing party voters in the minimum number of districts to minimise the strength in most of the districts.
- **Cracking:** Split up the "influencing voters" of the opposition into several districts to not consolidate the supporters of the influencing voters from the opposition.

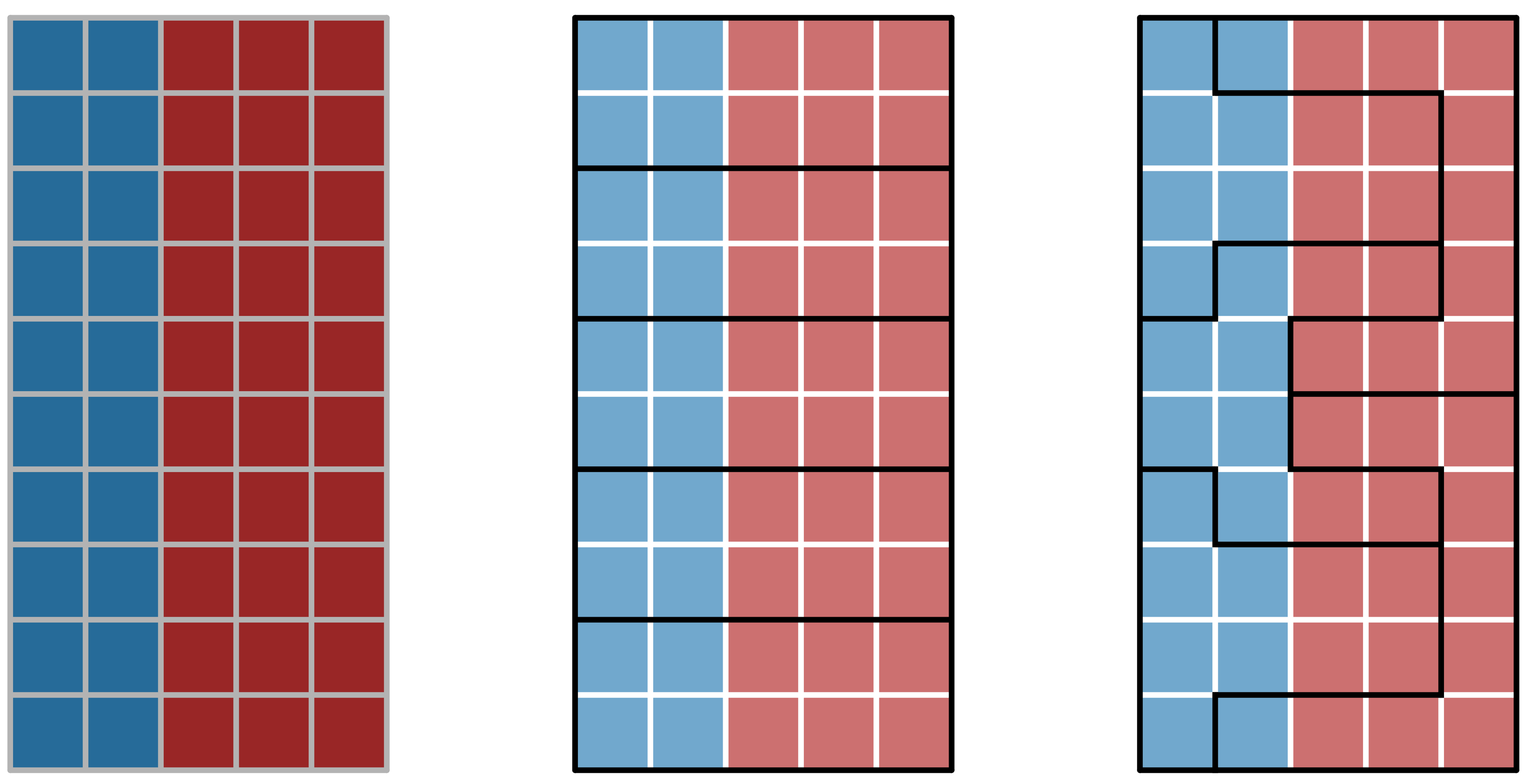


Figure 1. How to steal an election. Left: 50 precincts with 60% red party supporters and 40% blue party supporters. Middle: 5 districts where the red party wins all. Right: A redistricting where the blue party win 3 out of 5 districts, thus winning by majority. The diagram is influenced by a blog post on fairvote.org.

Fairness to Combat Gerrymandering

We impose the following two fairness rules:

- The number of voters at each of the ballot boxes is bounded i.e. $\leq$ `upper` and $\geq$ `lower`.
- The margin on victory at each of the ballot boxes is also bounded i.e. $\leq \texttt{margin}_{\texttt{up}}$ and $\geq \texttt{margin}_{\texttt{low}}$.

Fair-Gerrymandering-$(\mathcal{X}, \rho)$

Input: A set of candidates $\mathcal{C}$, a set $\mathcal{V}$ of n voters located at points in $\mathcal{X}$ whose preferences are known, a set $\mathcal{B}$ of m possible ballot box locations in $\mathcal{X}$ and a specific candidate "OUR" $\in \mathcal{C}$

Parameters: $k, \ell, m, n, \texttt{upper}, \texttt{lower}, \texttt{margin}_{\texttt{low}}, \texttt{margin}_{\texttt{up}}$

Assumptions:

- Each voter votes at the ballot box nearest to them, where distances are calculated using the metric ρ.
- The *plurality* rule is used: each voter votes for their top-ranked candidate, and a ballot box is won by the candidate who secures the most votes. There are no ties.
- The number of voters voting for a candidate at every ballot box is bounded i.e., $\leq$ `upper` and $\geq$ `lower`.
- The margin of victory at every ballot box is $\leq \texttt{margin}_{\texttt{up}}$ and $\geq \texttt{margin}_{\texttt{low}}$.

Question: Is there a set $\mathcal{P} \subseteq \mathcal{B}$ such that $k = |\mathcal{P}|$ such that opening ballot boxes at locations in $\mathcal{P}$ then "OUR" candidate wins at least ℓ of the ballot boxes for some $\ell \leq k \leq m$.

Our Contributions

Theorem (Algorithm)

If $\mathcal{C}$ is the set of candidates in an election, n be the number of voters and m be the possible ballot box locations in the plane, then Fair-Gerrymandering-$(\mathbb{R}^2, \ell_2)$ *is solvable in time* $(m+n)^{\mathcal{O}(\sqrt{k})} \cdot |\mathcal{C}|^{(\texttt{upper}+\texttt{lower}+\texttt{margin}_{\texttt{up}}+\texttt{margin}_{\texttt{low}})}$ *where k is the number of ballot boxes for the election.*

A brute-force search for a solution will run in $m^k n^{O(1)}$. Our algorithm is more efficient than an exhaustive search.

We complement this algorithm with an almost-matching lower bound:

Theorem (Lowerbound)

For any $d \geq 2$, under the Exponential Time Hypothesis (ETH), the Fair-Gerrymandering-$(\mathbb{R}^d, \rho)$ *problem cannot be solved in $f(k, n, \texttt{upper}, \texttt{lower}) \cdot m^{o(k^{1-1/d})}$ time where f is any computable function, n is the number of voters, and k is the number of the ballot boxes opened, m is the total number of possible locations of ballot boxes and ρ is either the ℓ_∞-metric or the ℓ_q-metric for some $q \geq 1$. This lower bound holds even when there are only 2 candidates, $k = \ell$ and $\texttt{margin}_{\texttt{low}} = 1 = \texttt{margin}_{\texttt{up}}$.*

Exponential Time Hypothesis (ETH)

Satisfiability of 3-CNF Boolean formulas cannot be solved in subexponential time, i.e., $2^{\varepsilon n}$ for all constant $\varepsilon > 0$, where n is the number of variables in the formula. [Impagliazzo & Paturi, CCC '99]

Definition (Grids)

The d-dimensional grid $\mathsf{R}[\kappa, d]$ is an undirected graph with vertex set $[\kappa]^d$ and two vertices $\mathbf{a}, \mathbf{b}$ are adjacent if and only if $\sum_{i=1}^{d} |\mathbf{a}[i] - \mathbf{b}[i]| = 1$.

Definition (Binary CSPs)

A binary constraint satisfaction problem (CSP) is a triple $\mathcal{I} = (V, D, C)$ where V is a set of variables, D is a domain of values and C is a set of constraints. It has two types of constraints:

- *Unary constraints*: For some $v \in V$ there is a unary constraint $\langle v, R_v \rangle$ where $R_v \subseteq D$.
- *Binary constraints*: For some $u, v \in V$ there is a binary constraint $\langle (u, v), R_{u,v} \rangle$ where $R_{u,v} \subseteq D \times D$.

Definition (Geometric CSPs)

A d-dimensional geometric $\geq$-CSP $\mathcal{I} = (V, D, C)$ is a binary CSP whose

- set of variables V is a subset of $\mathsf{R}[\kappa, d]$ for some $\kappa \geq 1$
- domain is $[N]^d$ for some integer $N \geq 1$
- constraint graph $G_{\mathcal{I}}$ is an induced subgraph of $\mathsf{R}[\kappa, d]$
- binary constraints are of the following type: if $\mathbf{a}, \mathbf{a}' \in V$ such that $\mathbf{a}' = \mathbf{a} \oplus \mathbf{e}_i$ for some $i \in [d]$ then there is a binary constraint $\langle (\mathbf{a}, \mathbf{a}'), R_{\mathbf{a},\mathbf{a}'} \rangle$ with $R_{\mathbf{a},\mathbf{a}'} = \Big\{ \big((x_1, x_2, \ldots, x_d), (y_1, y_2, \ldots, y_d)\big) \mid x_i \geq y_i \Big\}$

Sketch of the Reduction

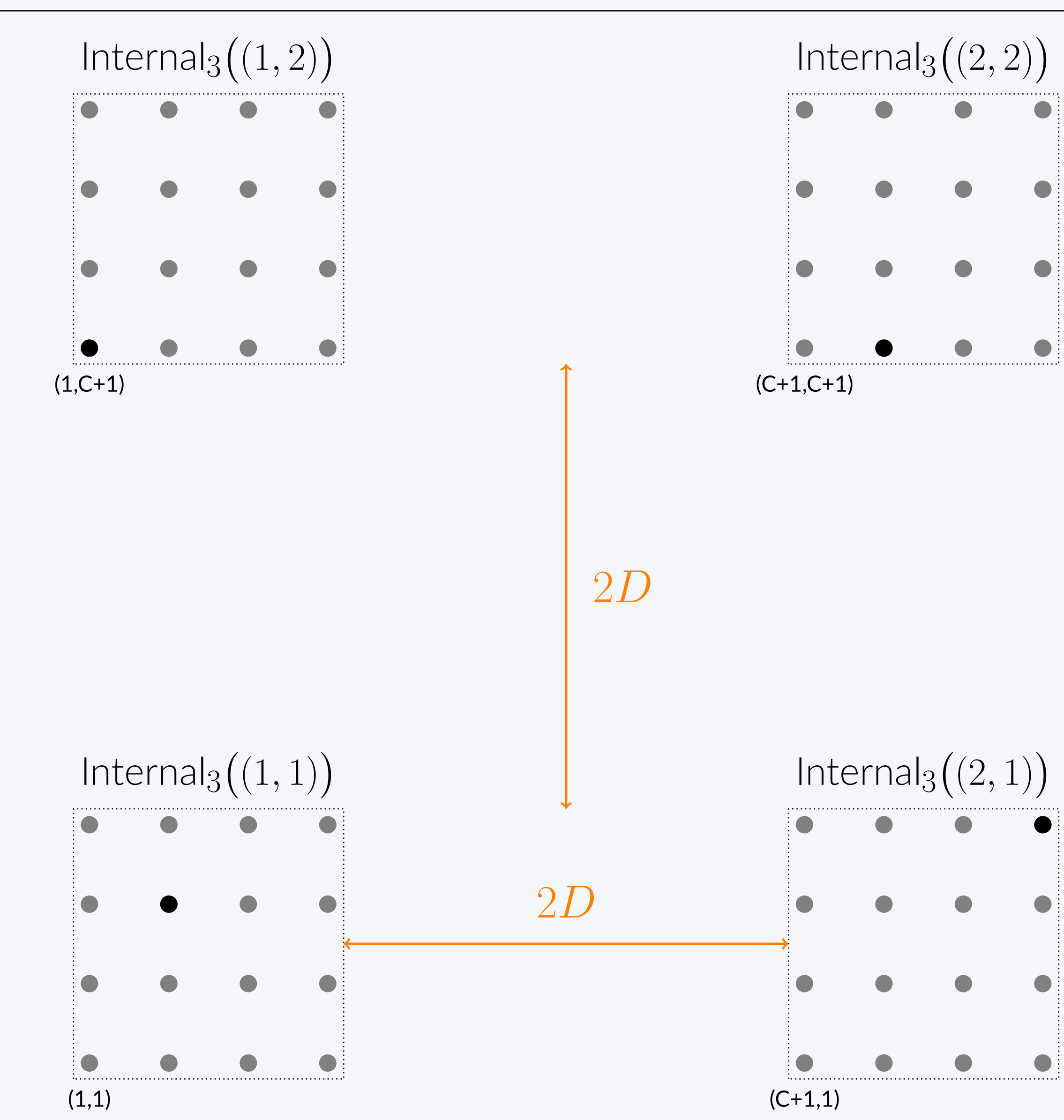
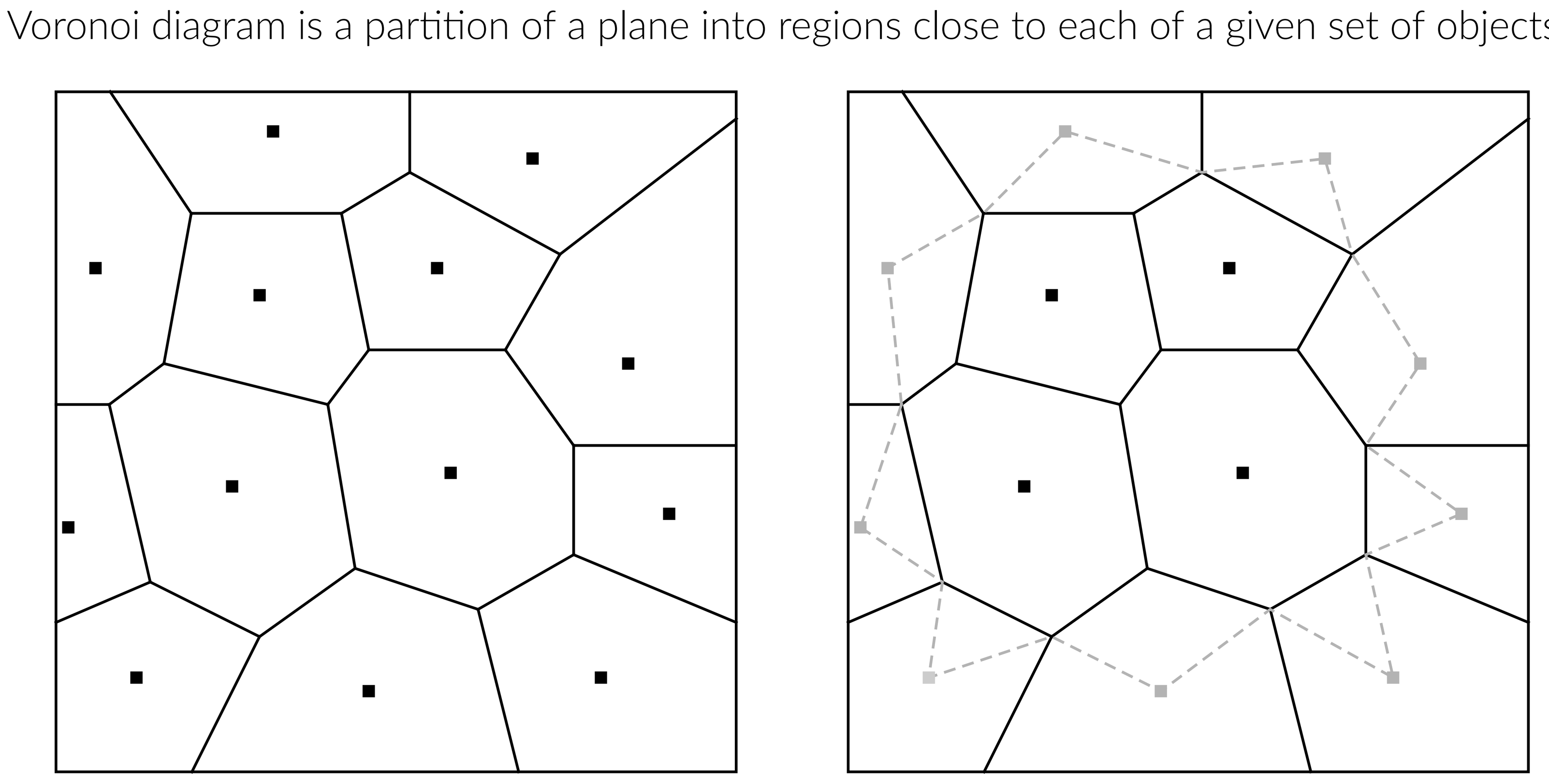


Figure 2. An illustration of the construction of the Gerrymandering-$(\mathbb{R}^d, \ell_2)$ instance when $\kappa = 2 = d$, $N = 4$ and $V = \mathsf{R}[\kappa, 2] = \big\{(1,1), (1,2), (2,1), (2,2)\big\}$.

Voronoi Diagram

A Voronoi diagram is a partition of a plane into regions close to each of a given set of objects.



In our case, these objects are potential locations of ballot boxes.

Lemma (Planar Separator)

Let S be a set of k points on the plane. Consider the Voronoi diagram of S and the planar graph G associated with it. There is a polygon Γ which has length $\mathcal{O}(\sqrt{k})$ and its vertices alternate between elements of S and vertices of G and each segment of Γ lies inside a face of G. At most $\frac{2}{3}k$ faces lie strictly inside Γ, and at most $\frac{2}{3}k$ strictly outside. [Marx & Pilipczuk, ESA '15]

Overview of the Algorithm

- Our algorithm guesses Γ, shown with the dotted line.
- It takes the boxes in the solution and then solves the inside and outside separately.

Open Question

The interesting open question is whether one can design an algorithm that works also in higher dimensions.

Funded by Deutsche Forschungsgemeinschaft – project number 416767905 and Polish National Science Centre (NCN) Grant 2020/39/B/ST6/01641.