Political districting to optimize the Polsby-Popper compactness score with application to voting rights

Pietro Belotti
Electronics, Information and Bioengineering, Politecnico di Milano, pietro.belotti@polimi.it

Austin Buchanan, Soraya Ezazipour
Industrial Engineering & Management, Oklahoma State University, {buchanan,sezazip}@okstate.edu

In the academic literature and in expert testimony, the Polsby-Popper score is the most popular way to measure the compactness of a political district. Given a district with area A and perimeter P, its Polsby-Popper score is given by $(4\pi A)/P^2$. This score takes values between zero and one, with circular districts achieving a perfect score of one. In this paper, we propose the first mathematical optimization models to draw districts (or districting plans) with optimum Polsby-Popper score. Specifically, we propose new mixed-integer second-order cone programs (MISOCPs), which can be solved with existing optimization software. Experiments show that they can identify the most compact *single districts* at the precinct level and the most compact *plans* at the county level. Then, we turn to the problem of drawing compact plans with a large number of majority-minority districts. This is the task faced by plaintiffs in Voting Rights Act cases who must show that an alternative plan exists in which the minority group could achieve better representation, a legal hurdle known as the first *Gingles* precondition. For this task, we propose new MISOCP-based heuristics that often outperform enacted maps on standard criteria, sometimes by substantial margins. They also perform well against state-of-the-art heuristics like short bursts and can be used to polish maps with hundreds of thousands of census blocks. Our techniques could assist plaintiffs when seeking to overturn maps that dilute the voting strength of minority groups. Our code is available on GitHub.

1. Introduction

In the academic literature and in expert testimony, the Polsby-Popper score is the most popular way to measure the compactness of a political district. This score was first proposed for districting by Polsby and Popper (1991) in a law journal, although the same score had been proposed for measuring the roundness of sand grains in a paleontology journal 64 years prior (Cox 1927), and the ideas behind it date back 3,000 years to Dido's problem and ancient Carthage (Osserman 1978, Bandle 2017). The Polsby-Popper score of a district is a function of its area A and perimeter P:

$$\text{(Polsby-Popper score)} \quad \frac{4\pi A}{P^2}.$$

The normalizing factor 4π ensures that the score takes values between 0 and 1, with circular districts achieving a perfect score of 1—a consequence of the well-known isoperimetric inequality.

Compactness—whether measured via the Polsby-Popper score or otherwise—is one of many criteria that are used when partitioning a state into political districts. Other *traditional redistricting principles* state that districts should have nearly equal populations ("population balance"), that they should be contiguous on the map, and that they should not unnecessarily divide political subdivisions such as counties, cities, and towns. In response to recent partisan gerrymandering, reform groups have also sought new laws to promote competitiveness, partisan fairness, or proportionality, or to promote the use of independent redistricting commissions to draw district lines (instead of state legislatures who have an obvious conflict of interest).

Federal law in the USA also requires states to abide by the Voting Rights Act (VRA) of 1965, which prohibits racial discrimination in voting (Hebert et al. 2010, Davis et al. 2019, Gordon and Spencer 2022). Section 2 of the VRA has been used to overturn maps that dilute the voting power of minority groups. Under the Supreme Court's ruling in *Thornburg v. Gingles* (1986), a successful vote dilution claim must first show the "compactness" precondition, i.e., that the minority group is "sufficiently large and geographically compact to constitute a majority in a single-member district". That is, the court requires plaintiffs to provide an alternative map ("demonstration districts") with more majority-minority districts. The Supreme Court also ruled that, when drawing minority-opportunity districts, plans must abide by traditional districting principles "including but not limited to compactness, contiguity, respect for political subdivisions or communities defined by actual shared interests" and not let race be the "predominant factor" lest they be ruled a racial gerrymander in violation of the Equal Protection Clause of the 14th Amendment, see *Shaw v. Reno* (1993).

Following the Supreme Court case *Allen v. Milligan* (2023), Alabama was required to draw a second Black-opportunity district to satisfy the VRA. Much of the expert testimony in this case concerned the ability to draw suitable demonstration districts, with particular attention being paid to the tradeoffs between political subdivision preservation, district compactness, and minority representation. Pertinently for us, an amicus brief was filed in this case by a team of computational redistricting experts who wrote that "optimization algorithms are well-suited to the task of generating [remedial plans in Section 2 litigation] ... as they can identify innovative combinations of geography that better comply with multiple traditional redistricting principles than any individual mapmaker is likely to find manually through trial and error" (DeFord et al. 2022).

1.1. Our Contributions

In light of this discussion, it is clear that political districting involves tradeoffs between numerous competing criteria and that generating maps that perform well on them can have huge impacts in the legal arena. To do this in a transparent manner, we propose to use mathematical optimization. To have the biggest impact, we would like to use the Polsby-Popper score to measure compactness as it is the most popular score, but its nonlinear nature makes it mathematically distinct from the (integer) linear optimization models found in the operations research literature on districting. Further, districting is already an NP-hard combinatorial problem without this nonlinearity, which may scare off the disinclined. Nevertheless, we show that the Polsby-Popper score can be captured in a mixed-integer second-order cone program (MISOCP) which is a type of optimization problem that some solvers have only recently begun to support. Additionally, we show how to encode the equivalent inverse or L^{-1} average Polsby-Popper score (Chikina et al. 2017, Duchin 2018) and the Schwartzberg score (Schwartzberg 1965) as MISOCPs. Despite the challenges posed by nonlinearity and NP-hardness, we find that our optimization models can identify the most compact *single districts* at the precinct level and identify the most compact *plans* at the county level.

Then, we turn to the problem of drawing compact plans with a large number of majority-minority districts. This is the first legal hurdle—the first *Gingles* precondition—that plaintiffs must clear to bring a vote dilution claim under Section 2 of the VRA. This task can be difficult, and heuristic optimization approaches have provided key assistance in several recent court cases, see Buchanan (2023), with "short bursts" being one of the primary tools (Cannon et al. 2023, Becker et al. 2021). In this paper, we provide another powerful tool that uses the newly proposed MISOCPs as subroutines. Our approach has three distinct phases, as summarized below.

1. *Carve*: Carve a compact majority-minority district from the state; repeat as able.
2. *Complete*: District the rest of the state in a recursive bipartition fashion.
3. *Cleanup*: Make the plan more compact, while retaining its other properties.

We find that the approach works quite well, often outperforming enacted maps, both in terms of the number of majority-minority districts and in terms of other traditional districting criteria like compactness and splitting, sometimes by substantial margins. This "cushion" provides practitioners with much-needed flexibility to fine-tune the computer-generated plans based on their domain expertise and local knowledge, without worrying that such tweaks will cause the plan's measurables to dip below those of the enacted plan. Our implementation is also flexible in the sense that it has several "knobs" that can be turned to prioritize some criteria over others (e.g., Polsby-Popper score, population deviation, minority percentage, county splitting). The approach also fares well against state-of-the-art heuristics such as short bursts, both in terms of the number of majority-minority districts and the performance on other districting criteria. We will also see that *Cleanup* can polish *Gingles* demonstration plans built from nearly 200,000 census blocks. Our python code and computational results are publicly available at: `https://github.com/AustinLBuchanan/Polsby_Popper_optimization`

1.2. Outline

Section 2 provides the necessary background and literature review. Section 3 proposes MISOCPs for drawing compact districts and plans. We run a limited set of experiments on precinct-level and county-level instances (where each vertex represents a precinct and county, respectively) to show their computational limits. Section 4 proposes to use the MISOCPs to draw compact plans with many majority-minority districts, motivated by the first *Gingles* hurdle to bringing a Section 2 VRA case. The proposed approach has three phases: *Carve*, *Complete*, and *Cleanup*. Although naive implementations of these phases lead to subpar results, small and intuitive tweaks lead to a powerful tool that can assist plaintiffs in meeting the first *Gingles* precondition. We experiment with the approach on states subject to recent and ongoing VRA litigation in Section 5 and compare against state-of-the-art heuristics such as short bursts. We conclude in Section 6.

2. Background and Literature Review

We have already seen that traditional redistricting principles include population balance, contiguity, and compactness. Current practice in the USA is that *congressional* districts often differ in population by just one person, e.g., each of Alabama's seven congressional districts has a population of either 717,754 or 717,755 according to the 2020 Census. However, larger deviations of up to 1% have been justified. For example, in *Tennant v. Jefferson County* (2012) the US Supreme Court upheld West Virginia's three congressional districts that had populations of 615,991 and 616,141 and 620,862 that, when compared to the ideal district population $617,664.67$, had a total deviation of $(620,862-615,991)/617,664.67 = 0.79\%$. Still, a majority of states enact districts with 1-person deviation to avoid litigation. Meanwhile, when drawing state senate and state house districts, larger deviations approaching 10% are typical in practice and permitted by the courts (Hebert et al. 2010, Davis et al. 2019). For example, California's State Senate districts vary in population from 938,834 at the low end, to 1,036,376 at the high end, for a total deviation of nearly 100,000 people, or 9.87%. Contiguity is often required by state law and is usually not a point of contention. Even when not required to do so, states typically enact contiguous districts.

Meanwhile, compactness is harder to define and open to interpretation. Although dozens of alternative compactness scores have been proposed, few states specify which ones to use, and many people simply rely on the "eyeball test". Indeed, the mathematician Peyton Young opined in 1988 that, due to the flaws of most compactness scores, "compactness should either be abandoned as a standard altogether, or left in the domain of the dictionary definition, to be interpreted by the courts in the light of the circumstances". We refer the reader to Young (1988) for an overview of various compactness scores (e.g., Reock, Schwartzberg, Taylor, Boyce-Clark, length-width, perimeter, moment-of-inertia) and their strengths and weaknesses, see also Niemi et al. (1990), Bar-Natan et al. (2020), Duchin (2022).

One criticism of perimeter-related scores, like the Polsby-Popper score, is that they suffer from the *Coastline Paradox* in which boundary lengths are not well-defined and depend on the choice of map projection and the "size of your ruler" (Bar-Natan et al. 2020, Barnes and Solomon 2021). Another criticism can be summarized with the slogan "land does not vote; people do". In 2010, 47% of all census blocks were uninhabited (Freeman 2014); reassigning these blocks to different districts can significantly change the Polsby-Popper score, but the districts would function the same. Or, imagine two counties separated by a river; if the river winds back and forth in a fractal-like manner, then the Polsby-Popper score would almost require these two counties to be in the same district, but if the river flowed straight then it may be convenient for them to be in different districts. As mathematician and expert witness Moon Duchin has put it, "it is not a great state of affairs when your metrics are heavily impacted by irrelevant factors but not impacted at all by important features of the problem you are studying" (Duchin 2022).

Nevertheless, compactness scores feature prominently in expert testimony and legal briefs. As law professor and mapmaker Nate Persily observed, "judges tend to like compactness measures because they have the feel of objective criteria against which you can evaluate whether one plan is better than another. . . Generally speaking, judges are also struck by the aesthetics, but they lean on the measures whenever possible" (Duchin and Walch 2022, Persily 2004). In particular, Duchin (2018) has declared the Polsby-Popper score to be "the most-cited compactness score in the redistricting literature and in expert testimony".

2.1. Minority Representation

Another key criterion in the USA is the Voting Rights Act (VRA) of 1965 which prohibits racial discrimination in voting. The VRA was originally passed in response to widespread disenfranchisement of Black citizens, particularly in southern states, and has been reauthorized and amended multiple times (Laney 2008, Gordon and Spencer 2022). For example, in 1975 protections were extended to "language minorities", which Congress defined as "persons who are American Indian, Asian American, Alaskan Natives or of Spanish heritage" (Ancheta 2007). In the 1980 case *Mobile v. Bolden*, the Supreme Court interpreted the VRA as forbidding discriminatory *intent*. Congress, not happy with this interpretation, clarified in 1982 that the VRA forbids discriminatory *effects*. Since 1965, the number of Black elected officials has increased 30-fold, with similar increases in Hispanic representation (Gordon and Spencer 2022). Without the VRA, this would have been unlikely (Grofman 1982, Chen and Stephanopoulos 2021, Duchin and Spencer 2021).

Section 2 of the VRA has been used to overturn maps that dilute the voting power of minority groups. To bring a Section 2 claim, one must first pass the three *Gingles* prongs (or "preconditions") established by the Supreme Court in the 1986 case *Thornburg v. Gingles*: "First, the minority group must be able to demonstrate that it is sufficiently large and geographically compact to constitute a majority in a single-member district... Second, the minority group must be able show that it is politically cohesive... Third, the minority must be able to demonstrate that the white majority votes sufficiently as a bloc to enable it... [to] usually defeat the minority's preferred candidate". The Supreme Court elaborated in *Bartlett v. Strickland* (2009) that the minority group must constitute a numerical majority ($> 50\%$) among the *voting age population* in the demonstration districts. After the *Gingles* hurdles are cleared, courts consider the "totality of the circumstances" and may require new maps (Persily 2004). In the remedial map, it is not required for the minority group to constitute a numerical majority; *minority opportunity districts* are also permitted in which the minority group constitutes less than 50% of the district yet is able to elect preferred candidates with crossover support from other groups (Grofman et al. 2001, Lublin et al. 2020, Becker et al. 2021).

Meanwhile, the Supreme Court has found racial gerrymandering to be unconstitutional, violating the Equal Protection Clause of the 14th Amendment. In *Shaw v. Reno* (1993), the court overturned North Carolina's 12th district, finding it so "bizarre" that it only could have been drawn on the basis of race. This district was reported to have a Polsby-Popper score of 0.01 (Hofeller 2015). Likewise, in *Miller v. Johnson* (1995), the court overturned Georgia's 11th district, finding it improper for map drawers to use race as the "predominant factor" to the subordination of "traditional race-neutral redistricting principles, including but not limited to compactness, contiguity, [and] respect for political subdivisions". In 2010, it was observed that every district overturned based on *Shaw* was built using an overreliance on census blocks, rather than from larger units like census tracts or precincts (Hebert et al. 2010). Figure 1 shows the overturned districts from *Shaw* and *Miller*.

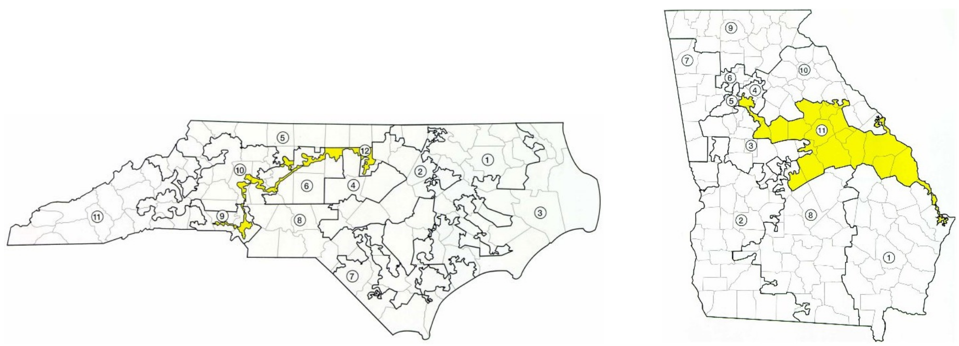


Figure 1 The US Supreme Court overturned North Carolina's 12th district in *Shaw v. Reno* (1993) and Georgia's 11th district in *Miller v. Johnson* (1995).

2.2. Variants of the Polsby-Popper score

As we have seen, the Polsby-Popper score of district D is defined as follows.

$$\text{(Polsby-Popper score)} \quad \mathrm{PP}(D) := \frac{4\pi A_D}{P_D^2},$$

where A_D is the area of the district and P_D its perimeter. The normalizing factor 4π ensures that the score takes values between 0 and 1, with circular districts achieving a perfect score of 1.

Example. Calculations below show that the Polsby-Popper scores of disks (●), squares (■), and regular hexagons (⬣) are approximately 1, 0.7854, and 0.9069, respectively.

$$\text{(disk, radius } r) \quad \mathrm{PP}(\bullet) = \frac{4\pi A_{\bullet}}{P_{\bullet}^2} = \frac{4\pi(\pi r^2)}{(2\pi r)^2} = 1.$$

$$\text{(square, side length } t) \quad \mathrm{PP}(\blacksquare) = \frac{4\pi A_{\blacksquare}}{P_{\blacksquare}^2} = \frac{4\pi(t^2)}{(4t)^2} = \frac{\pi}{4} \approx 0.7854.$$

$$\text{(hexagon, side length } t\text{)} \quad \mathrm{PP}(\hexagon) = \frac{4\pi A_{\hexagon}}{P_{\hexagon}^2} = \frac{4\pi(3\sqrt{3}t^2/2)}{(6t)^2} = \frac{\pi\sqrt{3}}{6} \approx 0.9069. \quad \square$$

If we take the multiplicative inverse (or reciprocal) of the Polsby-Popper score, we get an equivalent score in the sense that it ranks districts by compactness in the same way, but with smallest-score districts being most compact:

$$\text{(Inverse Polsby-Popper score)} \quad \mathrm{PP}(D)^{-1} := \frac{P_D^2}{4\pi A_D}.$$

For example, the inverse Polsby-Popper scores of disks, squares, and regular hexagons are 1, $4/\pi \approx 1.2732$, and $6/(\pi\sqrt{3}) \approx 1.1027$, respectively.

Taking the square root of the inverse Polsby-Popper score gives the Schwartzberg score (Schwartzberg 1965, Duchin 2022):

$$\text{(Schwartzberg score)} \quad \mathrm{PP}(D)^{-1/2} := \frac{P_D}{\sqrt{4\pi A_D}},$$

which also ranks districts in the same way.

Since these scores are equivalent, one could find an optimally compact *single* district by optimizing any one of the scores (making sure to choose min/max appropriately). Similarly, we can impose equivalent compactness *constraints* in terms of any one score. However, when optimizing or constraining the compactness of an *entire districting plan*, this is less clear. Should one aggregate the individual district compactness scores into a single score for the entire plan? If so, how? Here we review some existing approaches.

Years ago, the popular mapping software Maptitude began reporting the average Polsby-Popper score (along with the minimum, maximum, and standard deviation), so it may come as no surprise that the "straight average" has proliferated throughout districting circles (pun intended). This includes websites like Dave's Redistricting App (DRA 2023), Harvard's ALARM project (McCartan et al. 2022a), and the Princeton Gerrymandering Project (Princeton Gerrymandering Project 2023) (which also reports the minimum score). The straight average is also the usual number reported in court. For example, Duchin (2021), in her expert report in the case *Allen v. Milligan*, reported the average Polsby-Popper score for Alabama's enacted plan as 0.222, as well as the average scores of four demonstration plans A, B, C, and D, as 0.256, 0.282, 0.255, and 0.249, respectively. Other experts in this case also report the average, including William S. Cooper (for the plaintiffs) and Thomas M. Bryan (for the defendants).

Meanwhile, in her expert report in *League of Women Voters of Pennsylvania v. Commonwealth of Pennsylvania*, Duchin (2018) uses:

> "*what mathematicians would call an L^{-1} average Polsby-Popper score: we average the reciprocals of the PP scores of the 18 districts. The reason to average reciprocals instead of the straight scores is to attach a heavier penalty to plans with one extremely low score among the districts.*"

The equivalent *sum* of inverse Polsby-Popper scores was proposed earlier by Chikina et al. (2017) for the same reason. We will see that the average (or sum) inverse Polsby-Popper score is convenient for mathematical optimization modeling because it can be reformulated in a convex way.

2.3. Computational Approaches for Districting

Because of the population balance and contiguity constraints, districting problems are generally NP-hard (Dyer and Frieze 1985, Altman 1997) and are not expected to admit polynomial-time algorithms. Consequently, many researchers have proposed heuristics, beginning as early as 1961 with flood fill (Vickrey 1961). New heuristics are being proposed all the time, and we invite readers to consult surveys on districting heuristics by Ricca et al. (2013) and Becker and Solomon (2022).

Another legally impactful line of computational work is called *ensemble analysis* (Chen and Stephanopoulos 2021, Duchin and Spencer 2021, DeFord and Duchin 2022). Here the idea is to generate a large collection of districting plans, ideally drawn randomly from an explicit target distribution (McCartan and Imai 2023, Clelland et al. 2021, Cannon et al. 2022, Procaccia and Tucker-Foltz 2022) considering the "rules of the game" like traditional redistricting principles and the Voting Rights Act. If an enacted plan is an outlier in this (empirical) distribution of plans, say, with respect to the number of seats won by a particular party, then this may suggest that it was intentionally drawn to be a partisan gerrymander. Many ensemble approaches are based on a Markov chain Monte Carlo (MCMC) framework that moves from one districting plan to a neighboring or similar districting plan in a random walk (Fifield et al. 2015, Cho and Liu 2018, Adler and Wang 2019, Autry et al. 2021). To mitigate or avoid the dependence on initial conditions, more recent approaches use larger search neighborhoods like recombination (DeFord et al. 2021) or avoid the Markov chain approach altogether (McCartan and Imai 2023, McCartan et al. 2022b). Recent approaches also do a better job of preserving political subdivisions like counties (Autry et al. 2021, McCartan and Imai 2023, Clelland et al. 2022) and drawing majority-minority or minority-opportunity districts (Cannon et al. 2023, Becker et al. 2021).

In particular, the short bursts approach of Cannon et al. (2023) is arguably the most prominent and successful approach for drawing large numbers of majority-minority districts. It works by repeatedly taking a random walk for a small number of steps (called the "burst length") and restarting the random walk from the best-performing plan within this burst. Surprisingly, this simple approach empirically outperforms the "biased random walks" that are common in local search as well as popular metaheuristics like simulated annealing (MGGG 2024c).

Many optimization-based approaches have also been proposed, beginning as early as 1963 using facility location integer programming models and transportation techniques (Weaver and Hess 1963, Hess et al. 1965). Other integer programming models have been proposed based on exponentially many set partitioning variables (Garfinkel and Nemhauser 1970, Mehrotra et al. 1998) or polynomially many assignment or labeling variables, see Validi and Buchanan (2022). Many of these optimization models seek compactness as their objective (Hess et al. 1965, Hojati 1996, Validi et al. 2022, Validi and Buchanan 2022), although more recent models also consider partisan fairness (Swamy et al. 2023, Gurnee and Shmoys 2021), minority representation (Önal and Patrick 2016, Arredondo et al. 2021, Fravel et al. 2024), and political subdivision preservation (Birge 1983, Önal and Patrick 2016, Shahmizad and Buchanan 2023).

Challenges faced by these optimization models include the contiguity constraints and the large size of districting instances. A popular way to impose contiguity uses the flow-based constraints of

Shirabe (2005, 2009), see also Oehrlein and Haunert (2017), Validi et al. (2022). Another approach is to use cut-based or separator-based constraints. The number of these constraints grows exponentially in the number of geographic units, so they are usually applied in a branch-and-cut fashion (Oehrlein and Haunert 2017, Validi et al. 2022, Validi and Buchanan 2022), see also Carvajal et al. (2013), Buchanan et al. (2015), Wang et al. (2017), Fischetti et al. (2017). Another popular approach uses the tree-based constraints of Zoltners and Sinha (1983) or the subsequent distance-based (Mehrotra et al. 1998, Cova and Church 2000, Caro et al. 2004, Önal and Patrick 2016, Gurnee and Shmoys 2021) or DAG-based generalizations (Shahmizad and Buchanan 2023). These constraints are fast in practice and always return contiguous solutions, but are invalid in the sense that they cut off some solutions. Zhang et al. (2024) propose a *perfect* or *integral* linear-size model for partitioning the n nodes of a planar graph into k contiguous districts, but unfortunately struggle to impose population balance efficiently.

All of these optimization models are mixed-integer linear programs, originally or after reformulation. In contrast, this paper proposes mixed-integer second-order cone programs (MISOCPs) which have nonlinear constraints (Drewes 2009, Belotti et al. 2013, Benson and Sağlam 2013) and are solved without linearization. With the additional expressiveness afforded by the second-order cone constraints, we propose the first models for handling the Polsby-Popper score and variants.

2.4. Terminology and Notation

Consider a simple graph $G=(V,E)$ whose vertices represent a state's geographic units (e.g., counties, tracts, blocks) and whose edges indicate which geographic units are adjacent on the map. Each geographic unit $i \in V$ has an associated population p_i, and the population of a district $D \subset V$ is indicated by the shorthand $p(D) := \sum_{i\in D} p_i$. The number of districts is k. The ideal district population equals the total population $p(V)$ divided by k. The smallest and largest populations permitted in a district are given by L and U. A districting plan is a partition of the vertices into k contiguous and population-balanced districts $(D_1, D_2, \ldots, D_k)$. That is, each district D_j should induce a subgraph $G[D_j]$ that is connected, and its population $p(D_j)$ should lie between L and U.

To calculate district areas, we need the area a_i of each geographic unit $i \in V$. The area of district D is then $\sum_{i \in D} a_i$. For perimeters, we need the border length b_e between adjacent geographic units $e = \{u, v\} \in E$, as well as the border length b_i between geographic unit $i \in V$ and the state's exterior (which is zero if i belongs to the state's interior). The subset of edges with one endpoint in district D is denoted by $\delta(D)$. Then, the total perimeter of district D is that which lies in the state's interior, $b(\delta(D)) = \sum_{e \in \delta(D)} b_e$, plus that which coincides with the state's border, $b(D) = \sum_{i \in D} b_i$.

2.5. Computational Setup

In this paper, we apply our techniques to districting instances from the USA. The raw input data comes from the 2020 US Census (US Census Bureau 2021a,b), with initial processing conducted by the (Redistricting Data Hub 2021), subsequent processing conducted by Daryl DeFord, and final state-specific map projections by us using MGGG (2024a). In our computational experiments, we use a desktop PC with Windows 11 enterprise, Intel Core i9-13900K CPU at 3.00 GHz (5.80 GHz turbo), 64 GB RAM. To solve MISOCPs, we use Gurobi v11.0.1. (In initial testing, we also used the FICO Xpress Solver, with substantially similar running times.) Our code is written in Python, handles graphs using NetworkX, and is available at: `https://github.com/AustinLBuchanan/Polsby_Popper_optimization`.

3. The MISOCP Models

In this section, we propose a mixed-integer second-order cone program (MISOCP) for drawing a single district with optimum Polsby-Popper score, which we apply to precinct-level instances from Oklahoma and Alabama. We then extend the MISOCP to draw optimally compact districting *plans*, which we apply to several county-level instances.

3.1. MISOCP for a Single District

We seek a single district $D \subseteq V$ from graph $G = (V, E)$ with optimum Polsby-Popper score that is contiguous (i.e., the subgraph $G[D]$ is connected) and population-balanced (i.e., $L \le p(D) \le U$). To model this, we introduce a binary variable x_i for each geographic unit $i \in V$ indicating whether it is selected in the district. We also use a binary variable y_e indicating whether the edge $e \in E$ belongs to the cut $\delta(D)$ between D and $V \setminus D$. Last, we use continuous variables A and P for the area and perimeter of district D, respectively.

The Polsby-Popper objective $4\pi A/P^2$ is not directly permitted in an MISOCP; however, a satisfactory reformulation exists. Specifically, introduce a variable z representing the inverse Polsby-Popper score $P^2/(4\pi A)$ and minimize z subject to the rotated second-order cone (or Lorentz cone) constraint $P^2 \leq 4\pi Az$. In an optimal solution, this constraint will hold at equality, meaning that z will equal the inverse Polsby-Popper score, and $1/z$ will equal the Polsby-Popper score.

So, our MISOCP for drawing a single district is as follows. Its continuous relaxation is an SOCP which admits specialized interior-point algorithms and the use of outer approximation cuts.

$$
\begin{aligned}
\min \quad & z && && \text{(1a)}\\
\text{s.t.} \quad & P^2 \leq 4\pi Az && && \text{(1b)}\\
& P = \sum_{e\in E} b_e y_e + \sum_{i \in V} b_i x_i && && \text{(1c)}\\
& A = \sum_{i\in V} a_i x_i && && \text{(1d)}\\
& x_u - x_v \leq y_e \quad \text{and} \quad x_v - x_u \leq y_e && \forall e = \{u,v\} \in E && \text{(1e)}\\
& L \leq \sum_{i\in V} p_i x_i \leq U && && \text{(1f)}\\
& D = \{i \in V \mid x_i = 1\} \text{ is connected} && && \text{(1g)}\\
& x_i \in \{0,1\} && \forall i \in V && \text{(1h)}\\
& y_e \in \{0,1\} && \forall e \in E. && \text{(1i)}
\end{aligned}
$$

The objective (1a) minimizes the inverse Polsby-Popper score. Constraint (1b) ensures that z is *at least* the inverse Polsby-Popper score (and the objective will cause it to hold at equality). Constraints (1c) and (1d) capture the perimeter and area of the district, respectively. Constraints (1e) indicate that if one endpoint of edge $e = \{u, v\}$ belongs to the district but the other endpoint does not, then the edge e between them is cut. This constraint permits $y_e = 1$ even when $x_u = x_v$, although this will not occur in an optimal solution. Constraints (1f) impose population balance.

Constraint (1g) states that the district should be connected. As written, this is not a linear constraint, but there are many ways to do this. For example, the flow-based constraints of Shirabe (2005, 2009) are arguably the most popular contiguity constraints in the literature, especially if a root vertex has been pre-selected. However, in our experiments, we use i,j-separator constraints:

$$x_i + x_j \leq 1 + \sum_{s \in S} x_s, \tag{2}$$

where i and j are nonadjacent vertices and S is an i,j-separator (i.e., i and j become disconnected if the vertices S are removed from the graph). Since there are exponentially many of these constraints, our implementation adds them on-the-fly, only as needed. In particular, we separate infeasible integer points x^* using the linear-time algorithm of Fischetti et al. (2017), which identifies a violated inequality of the form (2) in which the separator S is inclusion-wise minimal. This approach has worked well in previous research (Validi et al. 2022, Validi and Buchanan 2022) and suits our needs.

3.1.1. Application to Oklahoma In a quick example, we draw a single congressional district for Oklahoma with optimum Polsby-Popper score. We use census VTDs (i.e., precincts) as our geographic units, of which there are 1,947. According to the 2020 Census, Oklahoma had a population of $3,959,353$ to be divided over $k = 5$ congressional districts. We permit a total population deviation of 1% (±0.5%), setting bounds of L =787,912 and U =795,829. After two minutes of computation on our desktop PC, the MISOCP solver Gurobi v11.0.1 returns an optimal solution with an inverse Polsby-Popper score of roughly 1.3184, i.e., a Polsby-Popper score of roughly 0.7585, as depicted in Figure 2. By chance, the district largely follows the boundaries of Oklahoma County, which has a rectangular shape.

3.1.2. Application to Alabama In a second example, we draw congressional districts for Alabama. Motivated by the Supreme Court case *Allen v. Milligan* (2023), we seek districts that are majority-Black. Specifically, following the Supreme Court's opinion in *Bartlett v. Strickland*

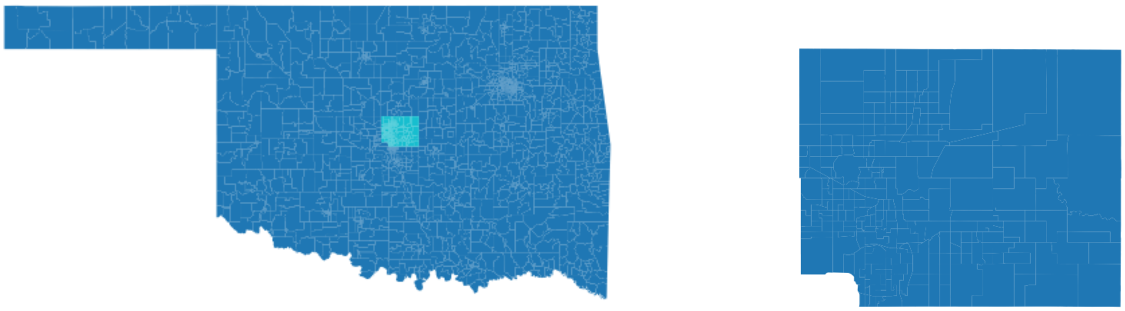


Figure 2 A congressional district for Oklahoma with optimum Polsby-Popper score.

(2009), we impose that the Black voting age population (BVAP) is at least 50% of the district's VAP[1]. We denote the BVAP and VAP of geographic unit $i \in V$ by BVAP_i and VAP_i, respectively. Our constraint is then:

$$\sum_{i \in V} \text{BVAP}_i x_i \geq \frac{1}{2} \sum_{i \in V} \text{VAP}_i x_i. \tag{3}$$

Given the emphasis on political subdivisions in *Allen v. Milligan* (2023), we also require that any district consist of whole counties plus at most one partial (or split) county. To model this, let C be the set of counties and introduce binary variables α_c and β_c for each county $c \in C$ indicating whether all or some of county c is selected in a district, respectively. The difference between these variables, $\gamma_c = \beta_c - \alpha_c$, indicates whether the district contains only part of county c. For each county $c \in C$, denote by V_c as the subset of V that lies within c. This leads to the logical constraints:

$$\alpha_c \leq x_i \leq \beta_c \qquad \forall i \in V_c,\ \forall c \in C \tag{4a}$$

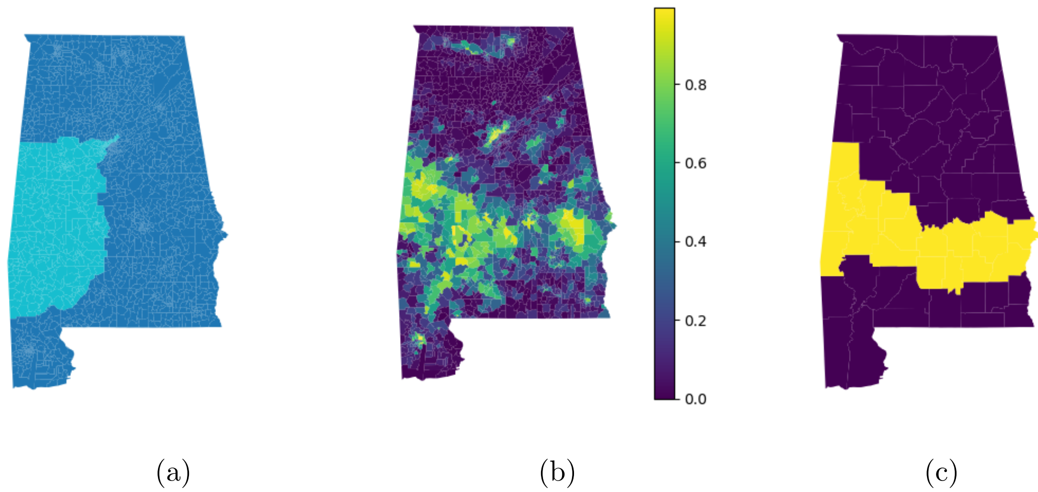
$$\gamma_c = \beta_c - \alpha_c \qquad \forall c \in C \tag{4b}$$

$$\sum_{c \in C} \gamma_c \leq 1 \tag{4c}$$

$$\alpha_c,\ \beta_c,\ \gamma_c \in \{0, 1\} \qquad \forall c \in C. \tag{4d}$$

If we solve the single district MISOCP model (1) with the majority-Black constraint (3) and county preservation constraints (4), we obtain the district in Figure 3a with Polsby-Popper score 0.3908.

Of course, the issue at hand in *Milligan* was whether Alabama was required by the VRA to draw not one, but *two* majority-Black districts. With the MISOCP machinery described thus far, how

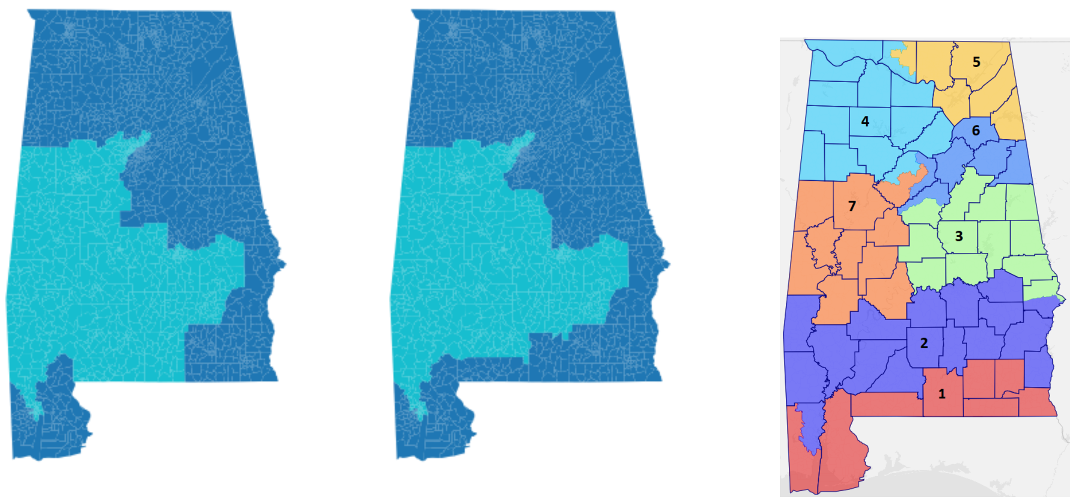


(a) (b) (c)

Figure 3 (a) majority-Black district with optimum Polsby-Popper score; (b) BVAP proportion by precinct; (c) 18 Black Belt counties

might we do this? First, recognize that Alabama's Black population lies primarily in the central and southern portions of the state, specifically in the state's largely rural Black Belt (known for its rich, black soil that was farmed by Black slaves for many years) and in the cities of Birmingham (near the center) and Mobile (near the southwest corner). To illustrate, Figure 3b shows the Black proportion of each precinct, $\text{BVAP}_i/\text{VAP}_i$ for $i \in V$, and Figure 3c shows the 18 counties that are considered part of the Black Belt (several others sometimes join). So, we suppose that if two majority-Black districts could be drawn then they might be adjacent to each other.

With this in mind, we may seek a majority-Black "double district", i.e., with twice the requisite population, under the hypothesis that it could subsequently be divided into two majority-Black districts. Intuitively, this—or something close—should be possible by discrete versions of the pancake theorem, see DeFord et al. (2023). Given that we seek *two* districts, we may allow for more county splitting, permitting, say, two partial counties in constraint (4c) instead of one. When we solve this MISOCP, we indeed obtain a double district, see Figure 4a. However, its complement is disconnected, and we cannot draw from it five other contiguous, population-balanced districts. In a second attempt, we require that both the double district *and its complement* be contiguous. To ensure the complement's contiguity, we apply the same i,j-separator constraints as before (2), but over the "complement" variables $\overline{x}_i = 1 - x_i$. We then obtain the double district in Figure 4b. This double district is somewhat similar to the two majority-Black districts that were prepared by the *Milligan* expert witness Moon Duchin in her Plan D, see Figure 4c. However, Duchin's districts also include the Black Belt counties of Barbour and Russell, better preserving this community of interest. Later, we will propose MISOCP techniques that can be used to extend *Gingles* districts and double districts into full districting plans.



(a) (b) (c)

Figure 4 (a) majority-Black double district (first attempt); (b) majority-Black double district (second attempt); (c) Duchin's Plan D

3.2. MISOCP for a Districting Plan

Now, we seek districting *plans* that are optimally compact with respect to the Polsby-Popper score. That is, instead of drawing a single compact district within a state, we seek to partition the state into k compact districts. In the following, we seek to optimize the average: (1) inverse Polsby-Popper score; (2) Polsby-Popper score; and (3) Schwartzberg score.

3.2.1. Inverse Polsby-Popper Score. We find that the inverse Polsby-Popper score is mathematically more convenient than the Polsby-Popper and Schwartzberg scores, so we first propose an MISOCP for it. The variable definitions are similar to before, but with an additional index j for the district number. The binary variable x_{ij} equals one when geographic unit $i \in V$ is assigned to district $j \in [k] = \{1, 2, \ldots, k\}$. The binary variable y_e^j equals one when edge $\{u, v\} \in E$, with $u < v$, is cut because geographic unit $u \in V$ is assigned to district j but geographic unit $v \in V$ is not. The continuous variables P_j, A_j, and z_j capture the perimeter, area, and inverse Polsby-Popper score of district j, respectively.

Our MISOCP to minimize the average inverse Polsby-Popper score is:

$$\min \quad \frac{1}{k} \sum_{j=1}^{k} z_j \tag{5a}$$

$$\text{s.t.} \quad P_j^2 \leq 4\pi A_j z_j \qquad \forall j \in [k] \tag{5b}$$

$$P_j = \sum_{e \in E} b_e y_e^j + \sum_{i \in V} b_i x_{ij} \qquad \forall j \in [k] \tag{5c}$$

$$A_j = \sum_{i \in V} a_i x_{ij} \qquad \forall j \in [k] \tag{5d}$$

$$x_{uj} - x_{vj} \leq y_e^j \qquad \forall e = \{u, v\} \in E, \ \forall j \in [k] \tag{5e}$$

$$L \leq \sum_{i \in V} p_i x_{ij} \leq U \qquad \forall j \in [k] \tag{5f}$$

$$D_j = \{i \in V \mid x_{ij} = 1\} \text{ is connected} \qquad \forall j \in [k] \tag{5g}$$

$$\sum_{j=1}^{k} x_{ij} = 1 \qquad \forall i \in V \tag{5h}$$

$$x_{ij} \in \{0, 1\} \qquad \forall i \in V, \ \forall j \in [k] \tag{5i}$$

$$y_e^j \in \{0, 1\} \qquad \forall e \in E, \ \forall j \in [k]. \tag{5j}$$

The objective (5a) minimizes the average inverse Polsby-Popper score. The constraints of model (5) are analogous to those of model (1) but are written for k districts instead of one. Another change is that the assignment constraints (5h) require each geographic unit $i \in V$ to be assigned to one district. As written, this model suffers from symmetry, but this can be ameliorated with the extended formulation for partitioning orbitopes of Faenza and Kaibel (2009), cf. Validi and Buchanan (2022).

3.2.2. Polsby-Popper Score. To maximize the average Polsby-Popper score, we want to

$$\max \quad \frac{1}{k}\sum_{j=1}^{k}\frac{1}{z_j}$$

subject to the same constraints from model (5). This is a mixed-integer nonlinear programming (MINLP) problem that does not seem to admit a direct reformulation as a MISOCP. Its continuous relaxation is nonconvex as it requires the maximization of a (separable) convex function. It could be solved using a spatial branch-and-bound algorithm (Smith and Pantelides 1999) by branching on the continuous variables z_j. However, we propose a different approximate approach that avoids spatial branching through a binary expansion of the $1/z_j$ terms, similar to Temiz et al. (2010). Specifically, express the Polsby-Popper score $1/z_j$ using binary variables b_{hj} as follows, where the user can set the desired precision through the parameter t. We choose $t = 20$, meaning that the binary expansion of the Polsby-Popper score will be off by less than one part in a million.

$$\sum_{h=1}^{t} 2^{-h} b_{hj} = \frac{1}{z_j}.$$

By the maximization objective, it suffices to impose this equation as a less-than-or-equal inequality. Then, after multiplying both sides by z_j, we obtain

$$\sum_{h=1}^{t} 2^{-h} b_{hj} z_j \leq 1.$$

Next, we introduce new variables w_{hj} to replace the terms $b_{hj}z_j$. We impose the equation $w_{hj} = b_{hj}z_j$ with a suitable big-M using the following constraints:

$$z_j + M(b_{hj} - 1) \leq w_{hj} \leq M b_{hj}$$
$$0 \leq w_{hj} \leq z_j.$$

The left inequalities ensure that $b_{hj}z_j \leq w_{hj}$, while the right inequalities ensure that $w_{hj} \leq b_{hj}z_j$. For them to be valid, it is required that each district have a Polsby-Popper score of at least $1/M$. In our experiments, we use $M = 16$.

Finally, our MISOCP to maximize the average Polsby-Popper score is:

$$\max \quad \frac{1}{k}\sum_{j=1}^{k}\sum_{h=1}^{t} 2^{-h} b_{hj} \tag{6a}$$

$$\text{s.t.} \quad \sum_{h=1}^{t} 2^{-h} w_{hj} \leq 1 \qquad \forall j \in [k] \tag{6b}$$

$$z_j + M(b_{hj} - 1) \leq w_{hj} \leq M b_{hj} \qquad \forall h \in [t],\ \forall j \in [k] \tag{6c}$$

$$0 \leq w_{hj} \leq z_j \qquad \forall h \in [t],\ \forall j \in [k] \tag{6d}$$

$$b_{hj} \in \{0,1\} \qquad \forall h \in [t],\ \forall j \in [k] \tag{6e}$$

$$(x,y,z) \text{ satisfies (5b)} - \text{(5j)}. \tag{6f}$$

3.2.3. Schwartzberg Score. To minimize the average Schwartzberg score, we would like to

$$\min \quad \frac{1}{k}\sum_{j=1}^{k}\sqrt{z_j}$$

subject to the same constraints from model (5). This objective is the minimization of a (separable) concave function, which we reformulate for MISOCP solvers. We begin by introducing a continuous variable s_j for the Schwartzberg score of each district $j \in [k]$ and change the objective to

$$\min \quad \frac{1}{k}\sum_{j=1}^{k} s_j.$$

By the minimization objective, it suffices to relate the Schwartzberg s_j variables to the inverse Polsby-Popper variables z_j with inequalities of the form $z_j \leq s_j^2$. While these constraints are non-convex, we can apply binary expansion on z_j and convert these inequalities into second-order cone constraints. Similar to before, we assume that the inverse Polsby-Popper score z_j of each district is less than M and use t for the user-chosen precision, say $M = 17$ and $t = 20$. Again, we introduce binary variables b_{hj} for each $h = 1, 2, \ldots, t$ and each $j \in [k]$. See that the expression $\sum_{h=1}^{t} 2^{-h} b_{hj}$ takes values between zero and (nearly) one, so we set the inverse Polsby-Popper score z_j to be

$$1 + (M-1)\sum_{h=1}^{t} 2^{-h} b_{hj},$$

which can take values between 1 and (nearly) M.

Finally, our MISOCP to minimize the average Schwartzberg score is:

$$\min \quad \frac{1}{k}\sum_{j=1}^{k} s_j \tag{7a}$$

$$\text{s.t.} \quad 1+(M-1)\sum_{h=1}^{t} 2^{-h} b_{hj}^2 \le s_j^2 \qquad \forall j \in [k] \tag{7b}$$

$$z_j \le 1+(M-1)\sum_{h=1}^{t} 2^{-h} b_{hj} \qquad \forall j \in [k] \tag{7c}$$

$$b_{hj} \in \{0,1\} \qquad \forall h \in [t],\ \forall j \in [k] \tag{7d}$$

$$(x,y,z) \text{ satisfies } (5b)-(5j). \tag{7e}$$

The second-order cone constraints (7b) and linear constraints (7c) together impose that $z_j \le s_j^2$. Note that $b_{hj} = b_{hj}^2$ as b_{hj} is binary, so the left side of inequality (7b) will equal the right of (7c).

3.3. County-Level Experiments

We now apply the MISOCP models to county-level instances, which have a vertex for each county and edges indicate county adjacencies. In these experiments, we seek to answer the following questions. How quick are the models? Does the running time heavily depend on which objective is used? We run tests on ten states that could draw county-level plans under a $\pm 0.5\%$ deviation: Arkansas, Idaho, Iowa, Kansas, Maine, Mississippi, Montana, Nebraska, New Mexico, and West Virginia. After the 2010 and 2020 censuses, only Iowa and West Virginia enacted county-level plans; the others are considered out of curiosity and to supplement our testbed.

To impose contiguity, we use a strengthened form of the i,j-separator constraints called length-U i,j-separator constraints that gives speedups on county-level instances (Validi et al. 2022). To separate them, we follow the procedure outlined by Validi et al. (2022), which first applies the linear-time algorithm of Fischetti et al. (2017) to get a minimal i,j-separator and then pares it down to a minimal length-U i,j-separator. We inject a warm start solution obtained by solving the model of Hess et al. (1965), which is very quick for county-level instances (Validi et al. 2022), and then apply the local search procedure (*Cleanup*) that is developed later. We also safely fix some variables to zero or one by exploiting the population balance constraints, see the L-fixing and U-fixing procedures for labeling models due to Validi and Buchanan (2022). For symmetry handling, we use the extended formulation for partitioning orbitopes due to Faenza and Kaibel (2009), cf. Validi and Buchanan (2022). Results are given in Table 1.

Table 1 MISOCP results on county-level instances for the average: inverse Polsby-Popper score (PP^{-1}), Polsby-Popper score (PP), and Schwartzberg score ($PP^{-1/2}$). We report the optimal objective value, or best lower and upper bounds $[LB, UB]$, and solve time in a 3600s time limit (TL).

			PP^{-1}		PP		$PP^{-1/2}$	
state	$\lvert C\rvert$	k	obj	time	obj	time	obj	time
AR	75	4	3.2251	16.46	[0.33,0.39]	TL	1.7796	294.83
ID	44	2	3.5377	0.24	0.3211	0.65	1.8595	0.60
IA	99	4	2.0570	34.64	[0.50,0.62]	TL	1.4280	1165.13
KS	105	4	2.3471	354.40	[0.42,0.67]	TL	[1.42,1.54]	TL
ME	16	2	3.0941	0.09	0.3320	0.24	1.7531	0.12
MS	82	4	3.3692	34.32	0.3861	2033.26	1.8216	1913.51
MT	56	2	2.7285	0.28	0.3716	1.03	1.6510	0.42
NE	93	3	2.2460	2.05	0.4487	132.02	1.4972	21.07
NM	33	3	2.3593	0.17	0.4258	3.17	1.5351	0.47
WV	55	2	4.9502	0.24	0.2395	1.32	2.2165	1.99

First, we observe that MISOCPs solve most quickly with the inverse Polsby-Popper objective, with most instances solving in under 35 seconds (and KS near six minutes). The average Polsby-Popper and Schwartzberg objectives take longer, reaching the time limit on three instances and one instance, respectively (all of which have $k \geq 4$ and $|C| \geq 75$). The average Polsby-Popper objective seems to be the most challenging, for example, taking more than one hour for AR, while the other objectives solve in 16 seconds and 295 seconds. Given these computational advantages—and the practical advantages noted by Chikina et al. (2017) and Duchin (2018)—the procedures that we propose next for drawing *Gingles* demonstration plans will use the inverse Polsby-Popper objective.

4. Extensions for Generating Gingles Demonstration Plans

We now turn to the problem of drawing compact plans with a large number of majority-minority districts. Recall that this is the first legal hurdle faced by plaintiffs to bring a vote dilution claim under Section 2 of the VRA.

We remark that each VRA case is unique, and there is no one-size-fits-all approach to drawing *Gingles* demonstration plans. Differing state laws and other criteria (e.g., communities of interest, like Alabama's Black Belt) may influence their development. Likewise, recognizing that courts are prone to adopt "least change" plans as remedies, VRA litigators may prefer demonstration districts that keep most of the enacted districts intact or that only scramble districts within certain portions of a state to permit "modular" remedies (see Moon Duchin's export report in *Georgia State Conference of the NAACP v. Raffensperger* for an example of this).

With these caveats in mind, we do not expect any computer-generated plan to be immediately court-ready, but hope that our approach can still be of assistance, analogous to that provided by short bursts (Cannon et al. 2023). Helpful in this regard is that our approach often outperforms enacted maps by significant margins on traditional criteria, leaving plenty of "cushion" for practitioners to work with.

Our approach has three phases. The first phase, *Carve*, repeatedly carves compact majority-minority districts from the state. Each iteration of *Carve* solves a single-district MISOCP. The second phase, *Complete*, districts the rest of the state. Each iteration of *Complete* solves a MIP to break part of the state into nearly equal halves, and this idea is applied recursively until each part is a single district. The third phase, *Cleanup*, makes the completed plan more compact while retaining its other key properties. Each iteration of *Cleanup* solves a multi-district MISOCP.

4.1. Phase 1: Carve

In *Carve*, we use the single-district MISOCP model (1) to identify a compact majority-minority district. The district is removed from the state, and the process is repeated.

Although the idea behind *Carve* is simple, intuitive, and not new (DeFord et al. 2018, McDonald 2019, McCartan and Imai 2023), we have found that naïve implementations of it perform poorly, both in terms of computational speed and in terms of solution quality. In an analogy, imagine taking a cookie cutter to dough. After cutting out several cookies, we may be left with disconnected, non-compact dough scraps. Normally, we would ball up the dough, roll it out, and start again; however, we do not get second chances like this in district carving. So, to improve the performance, we make a few tweaks to the MISOCP model, as described below.

First, like DeFord et al. (2018) and McCartan and Imai (2023), we require that the district's complement be connected, in our case using i,j-separator inequalities written over the complement variables $\overline{x}_i = 1 - x_i$. Without this, we may be left with unworkable scraps after a few carves.

Second, we penalize minority VAPs that are in excess of the 50% threshold, since otherwise the MISOCP sometimes unintentionally "packs" Black voters into districts with 80% to 95% BVAP, thus diluting Black voting strength. Specifically, we introduce a nonnegative surplus variable ζ in the original BVAP constraint (3), writing:

$$\sum_{i \in V} \text{BVAP}_i x_i - \zeta = \frac{1}{2} \sum_{i \in V} \text{VAP}_i x_i. \tag{8}$$

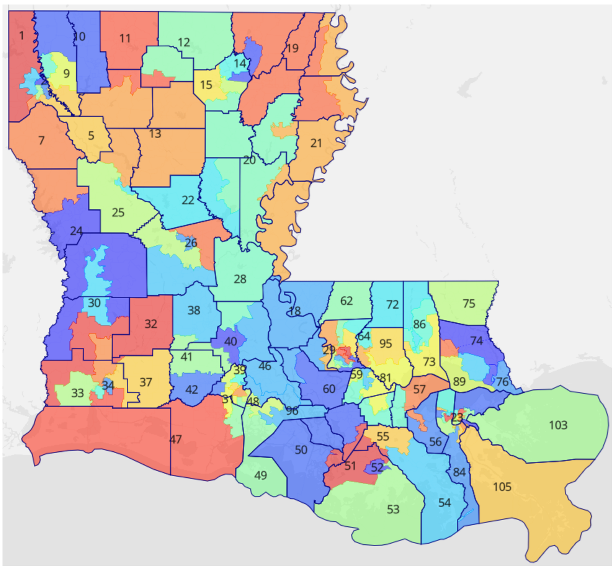
By default, we penalize the excess ζ in our minimization objective with the term $+0.001\zeta$. The interpretation here is that one penalty point is charged for every 1,000 people beyond the 50% threshold. The inverse Polsby-Popper score z usually lies in the single-digit range for our instances, and we have found the objective $\min z + 0.001\zeta$ to be a reasonable scalarization to balance the competing objectives of compactness and avoiding excessive packing.

Third, we promote county preservation. Specifically, we first carve districts that are built entirely from whole counties, then we seek districts with one partial county, and last we seek districts with two partial *adjacent* counties. For maximum speed, we coarsen the original nodes into county nodes where possible. Specifically, when carving districts built entirely from whole counties, we work with the county-level graph, i.e., the graph in which each county's nodes have been merged into a single node. Next, when carving districts with one partial county, we coarsen all but one county. In initial experiments, it seemed best to solve these problems sequentially, starting with the least-populous counties and working our way towards the most-populous counties. Intuitively, this is because the least-populous counties have fewer nodes in them, thus allowing us to first tackle smaller subproblems, possibly carving a large number of nodes from the graph, before tackling the larger subproblems. When seeking districts with two partial counties, we restrict ourselves to adjacent counties. Because the county-level graphs are planar (or nearly so), this means that the number of subproblems remains linear in the number of counties. Again, we solve them sequentially, starting with adjacent counties with the least combined population, in an effort to tackle smaller subproblems and reduce the size of the graph before attempting the largest subproblems.

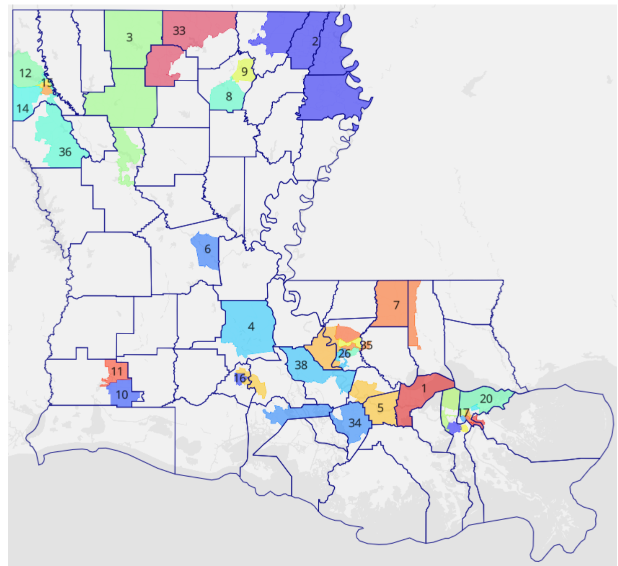
Fourth, we place a hard constraint on district compactness. Our default implementation requires that any carved district have a Polsby-Popper score of at least $1/8 = 0.125$. This is a reasonable restriction for most states and avoids costly MISOCP solves, especially where a majority-Black district could be drawn but would have too bizarre a shape.

For illustration purposes, we apply *Carve* to Louisiana's House of Representatives, which has 105 single-member districts. We choose this instance because it was the primary test case in the short bursts paper (Cannon et al. 2023). In the enacted plan, depicted in Figure 5a, there are 29 districts with a Black majority. Meanwhile, 31.25% of the voting age population is Black, and 33.13% of the total population is Black. So, somewhere between 33 and 35 districts would be proportional. Why does this matter? Under the Supreme Court's ruling in *Johnson v. De Grandy* (1994), a vote dilution claim is likely to be unsuccessful if the number of minority-opportunity districts is already proportional to the minority group's statewide population. In this case, however, the enacted plan does not achieve proportionality, a fact that was cited by the plaintiffs in *Nairne v. Landry*, which challenged both the state senate and state house plans. This case was decided in favor of the plaintiffs in February 2024, but the state has appealed the district court's ruling.

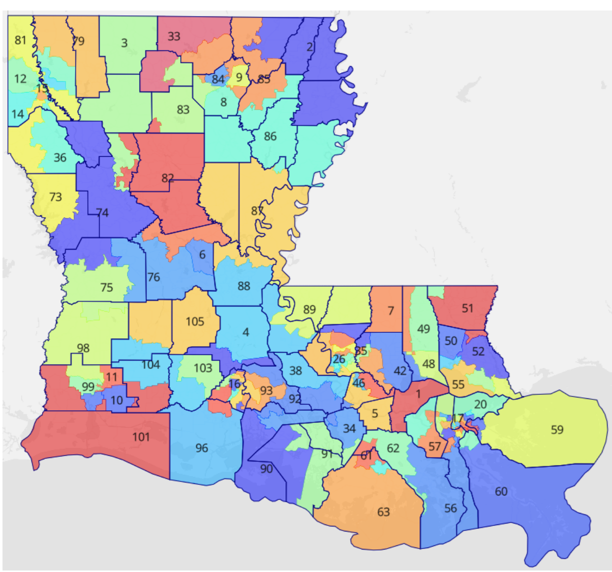
Applying *Carve* to the precinct-level graph, we obtain 38 majority-Black districts, as depicted in Figure 5b, with an average Polsby-Popper score of 0.3918. Compare this to the enacted plan's 29 majority-Black districts, which have an average score of 0.2670. Later, we provide a more extensive set of comparisons against enacted plans and the short bursts approach across several states on precinct-level and block-group-level instances.



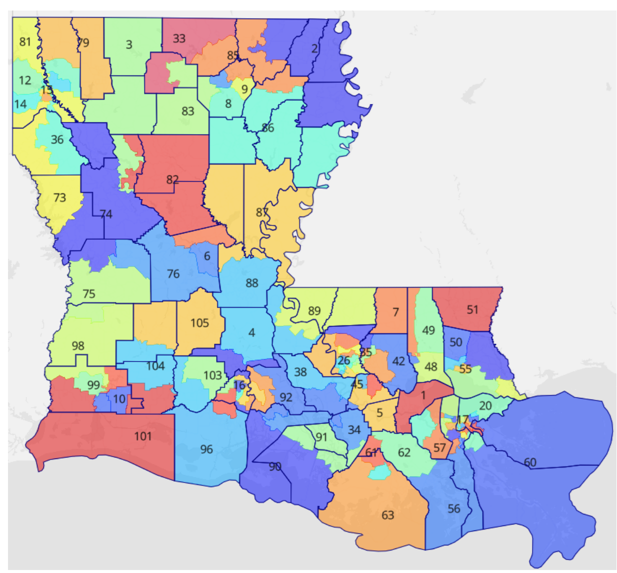
(a) Enacted



(b) *Carve*



(c) *Complete*



(d) *Cleanup*

Figure 5 Louisiana House of Representatives: Enacted plan versus *Carve*, *Complete*, and *Cleanup*

4.2. Phase 2: Complete

In *Complete*, we take the partial districting plan from *Carve* and complete it. Each iteration breaks part of the state into contiguous, (nearly) equal halves, repeating until each part is a single district. For this task, we use the single-district model (1) to identify one side of the bipartition, and its complement is taken as the other half. For example, after carving 38 *Gingles* districts, we seek $105 - 38 = 67$ additional districts for Louisiana's House of Representatives. The first iteration breaks the remainder into contiguous, nearly equal halves with populations suitable for 33 and 34 districts. Next, each half is broken in half, with populations suitable for 16 or 17 districts, etc.

While the idea behind *Complete* is simple and intuitive, we have found that naïve implementations of it perform poorly. Specifically, the biggest issue is that the bipartition problem can become infeasible, often because of the contiguity or population balance constraints. One possible approach to the population balance issue is to impose stricter population balance constraints for "larger" subproblems and progressively relax them to the originally intended deviations for the "smallest" subproblems. In this vein, Gurnee and Shmoys (2021) propose to use the tighter deviation tolerance of $\pm\varepsilon' := \varepsilon/\lceil \log_2(k') \rceil$, where k' is the size (e.g., $k' = 33$) and ε is the original deviation tolerance (e.g., $\varepsilon = 0.05$). We take a different approach to the population balance issue. Rather than impose stricter "hard" constraints for larger subproblems, we penalize deviations from the ideal population in the objective. This is similar to the approach we outlined for discouraging excessive BVAP concentrations, see (8), except that we penalize *both* too-high and too-low populations. As before, we use a penalty coefficient of 0.001, with the interpretation that one penalty point is charged for every 1,000-person deviation from ideal. In our experience, this nudge does the trick.

We also observe that if the Polsby-Popper objective is used for the bipartition step, then the smallest subproblems sometimes have graphs that are barely connected, also leading to infeasibility. So, instead of the inverse Polsby-Popper objective (1a), we minimize the number of cut edges, i.e., that cross between the two sides. This is equivalent to maximizing the number of edges that are preserved, thus saving them for subsequent subproblems.

Last, we promote county preservation. Using the same logical constraints as before (4), we require that only one county can be split across the two sides. In the event that this constraint makes the bipartition problem infeasible, we re-optimize without it. In some circumstances, the bipartition problem continues to be infeasible. In this case, we forgo the bipartition step and take its vertices as a multi-district. This allows the code to terminate gracefully, and the user can divide the multi-district into individual districts manually using more granular geographic units like blocks.

Applying *Complete*, we obtain 67 additional districts for the Louisiana House of Representatives, as depicted in Figure 5c. Across the entire plan, the average Polsby-Popper score is 0.3254, which is already an improvement over the enacted plan's average score of 0.2911, and its compactness will get even better in the next phase.

4.3. Phase 3: Cleanup

The last phase, *Cleanup*, takes a districting plan and improves its compactness while retaining its other key properties. At its core, *Cleanup* is a local search heuristic, with each iteration solving a restricted form of the MISOCP model (5).

As with any local search heuristic, one must define a local search neighborhood. In the spirit of Henzinger et al. (2020), our local search neighborhood permits nodes to relocate to nearby districts, say, to any district that lies within h hops. Meanwhile, nodes that are deep in a district's interior are forced to stay in their current district. To find an improved districting plan within this local search neighborhood (with respect to compactness), we solve the associated MISOCP.

Before proceeding, we recall two concepts from graph theory. Denote by $\text{dist}(i,j)$ as the hop-based distance between vertices $i,j \in V$ in graph G. Then, the (closed) h-hop neighborhood of vertex $i \in V$ is the set of vertices $N^h[i] = \{j \in V \mid \text{dist}(i,j) \leq h\}$ within h hops of i. For example, the 0-hop neighborhood of i is the singleton $\{i\}$; the 1-hop neighborhood is the usual (closed) neighborhood; and the 2-hop neighborhood includes i, its neighbors, and its neighbors' neighbors.

Relatedly, we can define the h-hop neighborhood *of a districting plan*. Below, we suppose that a districting plan $d : V \to [k]$ maps each vertex $i \in V$ to a district number from $[k] := \{1, 2, \dots, k\}$.

Definition 1 (h-hop neighborhood of a districting plan). Let d be a districting plan. The districting plan d' belongs to its h-hop neighborhood $N^h[d]$ if, for every vertex $i \in V$, there is a vertex $v \in V$ satisfying

1. $v \in N^h[i]$, i.e., v belongs to $i's$ h-hop neighborhood, and
2. $d'(i) = d(v)$, i.e., i's new district $d'(i)$ is v's old district $d(v)$.

Observe that the h-hop neighborhood of a districting plan always contains itself, i.e., $d \in N^h[d]$, provided that $h \geq 0$. The only plan in a districting plan's 0-hop neighborhood is itself, $N^0[d] = \{d\}$. The 1-hop neighborhood permits each vertex to remain in its current district or to move to any of its neighbors' districts.

The parameter h allows the user to control the size of the local search neighborhood and the difficulty of the associated MISOCPs. For small values of h, like $h = 1$ or $h = 2$, the MISOCPs are still relatively easy to solve and give a much larger search neighborhood than the usual *flip* and *swap* neighborhoods that only relocate one or two nodes per iteration.

For maximum speed, we should exploit the fact that most vertices stay in the same district from one iteration to the next. So, we do not create all $|V|k$ assignment variables of the form x_{ij}. Rather, we create x_{ij} only if i has an h-hop neighbor that presently belongs to district j. Likewise, not all cut edge variables y_e^j need to be created. Even better, given that large portions of a map may be frozen in the h-hop neighborhood, we may merge their nodes into a single super-node, similar to Henzinger et al. (2020). In this way, we need not worry about a memory crash when applying *Cleanup* to our running example for the Louisiana House of Representatives, which would normally require $3,540 \times 105 = 371,700$ assignment variables and nearly two million cut edge variables. As we will see later, this permits us to handle block-level instances with nearly 200,000 nodes.

By default, we limit local search to ten iterations, and limit each iteration to ten minutes. For county preservation reasons, we also disallow a vertex i from being assigned to district j if no other vertex its county was assigned to j in the previous iteration. This ensures that no new county splits will be introduced during local search (and, in fact, sometimes reduces splits).

Applying *Cleanup* to our running example, we improve the average Polsby-Popper score from 0.3254 to 0.3622. Before *Cleanup*, 42 counties (or Parishes) were split a total of 118 times. After *Cleanup*, 39 counties are split a total of 113 times (zero precinct splits). This slightly outperforms the enacted plan, which splits 41 counties a total of 116 times (three precinct splits).

4.3.1. Complexity of Cleanup What is the computational complexity of *Cleanup*? That is, given a districting plan, how difficult is it to find a more compact plan (say, with respect to the inverse Polsby-Popper score) in its local search neighborhood? For typical neighborhoods, like flip and swap, this is polynomial-time solvable by brute force. However, for the h-hop neighborhood, this is less clear as there may be exponentially many plans to sift through.

We show that this problem is NP-complete, even when there are only $k=2$ districts and the hop parameter is set to $h=1$. This may seem unsurprising given that districting plans usually must be contiguous and population-balanced, and these constraints are often thought to be the source of the difficulty. However, we show that the problem remains hard *even when these constraints are relaxed*, indicating that *Cleanup* under the inverse Polsby-Popper objective is itself hard.

Theorem 1. *Given a districting plan, it is NP-complete to determine whether there is a more compact plan (with respect to the average inverse Polsby-Popper score) in its h-hop neighborhood. This problem remains NP-complete even when $h=1$, $k=2$, and the contiguity and population balance constraints are relaxed.*

Proof. First, we provide a reduction from the subset sum problem to a "gift" version of subset sum. In the subset sum problem, we are given a list of positive integers $a_1, a_2, \dots, a_q$, and the question is whether there is a subset of them whose sum equals half of the total $\sum_{i=1}^q a_i$. Let $t_q = \sum_{i=1}^q a_i/2$ be this target value. The gift subset sum problem is nearly the same, except that the input *also* includes a subset whose sum is just one unit below the target value. To reduce the subset sum problem to the gift subset sum problem, simply add two new integers to the list $a_{q+1} = a_{q+2} = t_q + 1$, meaning that there are now $n = q+2$ integers and the new target value is $t = 2t_q + 1$. Observe that the first q integers $a_1, a_2, \dots, a_q$ sum to $2t_q = t-1$, which is one less than the new target t. It can also be seen that the original subset sum instance is a "yes" if and only if the gift subset sum instance is a "yes".

Next, observe that our local search problem belongs to NP because a districting plan in the h-hop neighborhood that is more compact (with respect to the average inverse Polsby-Popper score) is a suitable witness. To show NP-hardness, we provide a reduction from gift subset sum in which positive integers $a_1, a_2, \ldots, a_n$ are given as input. Again, denote by $t = \sum_{i=1}^{n} a_i/2$ the target value. We construct the districting instance in Figure 6.

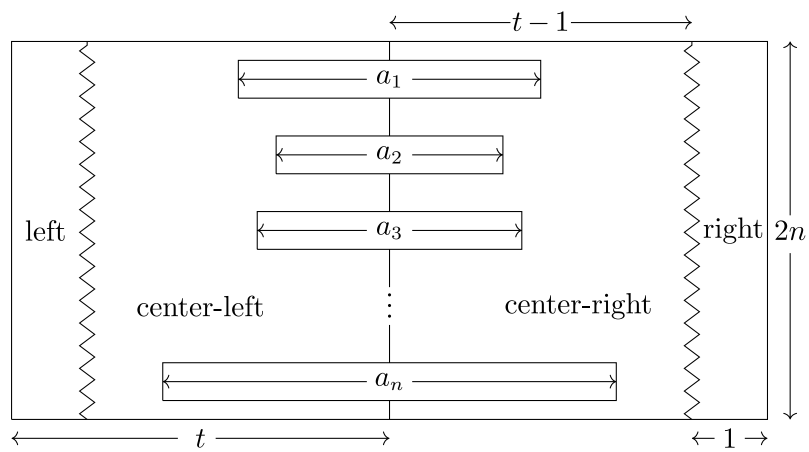


Figure 6 NP-hardness reduction from the "gift" subset sum problem.

The rectangular geographic units down the center have widths $a_1, a_2, \ldots, a_n$ and height one. Thus, their areas are $a_1, a_2, \ldots, a_n$, totaling $2t$. There is one unit of space between rectangles i and $i+1$, and one-half unit of space above rectangle 1 and below rectangle n. Thus, the total height of the bounding rectangle is $2n$, giving the entire state an area of $A = 4nt$. There are four other geographic units named left, center-left, center-right, and right. There are jagged (river) boundaries between left and center-left, and between center-right and right, so that these jagged borders have a large length, say, $M = \sqrt{128nt(n+t)^2}$. These jagged boundaries lie sufficiently to the left or right, say $t-1$ units from the center, so as not to touch the center rectangles. The leftmost and rightmost borders are each t units from the center. The left and right units have equal areas ($2n$), as do the center-left and center-right units ($2nt - 2n - t$).

For local search purposes, define the initial districts as follows. Create the first district from the left unit, center-left unit, and center rectangles from the "gift" solution to the subset sum problem (i.e., with widths $a_1, a_2, \dots, a_{n-2}$). Create the second district from what remains. The perimeter of each district is

$$P = P_1 = P_2 = 4(n+t),$$

and the first district's area is two units less than the right district's area ($A_1 = 2nt - 1$ versus $A_2 = 2nt + 1$). So, the average inverse Polsby-Popper score is

$$\frac{1}{2}\left(\frac{P_1^2}{4\pi A_1} + \frac{P_2^2}{4\pi A_2}\right) = \frac{P^2}{8\pi}\left(\frac{1}{A_1} + \frac{1}{A_2}\right) \tag{9a}$$

$$= \frac{P^2}{8\pi}\left(\frac{1}{2nt-1} + \frac{1}{2nt+1}\right) \tag{9b}$$

$$< \frac{P^2}{8\pi}\left(\frac{2}{2nt-1}\right) = \frac{4(n+t)^2}{\pi(2nt-1)}. \tag{9c}$$

As the left and right units lie within their districts' interiors, they must remain in the first and second district, respectively. Meanwhile, all other geographic units lie on a district boundary and are permitted to switch districts in the 1-hop neighborhood. However, because of the long border between the left and center-left units, $M = \sqrt{128nt(n+t)^2}$, it would be undesirable for them to be in different districts, as the average inverse Polsby-Popper score would be

$$\frac{1}{2}\left(\frac{P_1^2}{4\pi A_1} + \frac{P_2^2}{4\pi A_2}\right) > \frac{P_1^2}{8\pi A_1} \geq \frac{M^2}{8\pi(4nt)} = \frac{128nt(n+t)^2}{32\pi nt}$$
$$= \frac{4(n+t)^2}{\pi} > \frac{4(n+t)^2}{\pi(2nt-1)}.$$

Likewise, the right and center-right units must belong to the same district. Then, for any assignment of center rectangles to districts, the perimeter of each district will again be $P = 4(n+t)$. So, the average inverse Polsby-Popper score is again given by (9a), where the areas A_1 and A_2 satisfy $A_1 + A_2 = A = 4nt$. Removing the leading constant $P^2/(8\pi)$, we are left with the internal expression

$$\frac{1}{A_1} + \frac{1}{A_2} = \frac{1}{A_1} + \frac{1}{A - A_1},$$

which is strictly convex in the (integer) variable A_1 (for $0 < A_1 < A$) and symmetric about the unique minimizer $A_1 = A/2$. So, the only improved solution would come with $A_1 = A/2$, whose solution would evenly split the center rectangles' areas between the two districts. So, there is a more compact plan within the 1-hop neighborhood of the initial plan (with respect to the average inverse Polsby-Popper score) if and only if the gift subset sum instance is a "yes". $\square$

5. Experiments for Notable VRA Instances

In this section, we apply the proposed approach to notable VRA instances. Specifically, we consider state legislative districts, i.e., state senate ("SS") and state house ("SH"), from Louisiana ("LA"), Mississippi ("MS"), Alabama ("AL"), and Georgia ("GA"). Each of these eight enacted maps was challenged under Section 2 of the VRA in the 2020 redistricting cycle for minority vote dilution (Redensky and Leaverton 2023). Lawsuits brought by Black voters include: *Nairne v. Landry* in Louisiana; *Mississippi State Conference of the NAACP v. State Board of Election Commissioners* in Mississippi; *Milligan v. Allen* and *Stone v. Allen* in Alabama; and *Alpha Phi Alpha Fraternity Inc v. Raffensperger* and *Grant v. Raffensperger* in Georgia.

As previously mentioned, each VRA case is unique, and we do not believe that any computer-generated plan will be immediately court-ready. However, given that our approach often outperforms enacted maps by significant margins on traditional criteria, our hope is that it can assist in the drawing *Gingles* demonstration plans by giving a starting point with plenty of cushion, analogous to how short bursts and other computational methods have provided assistance.

To demonstrate the potential of our approach, we compare it against short bursts (Cannon et al. 2023), particularly its recent "Gingleator" implementation in the GerryChain software package (MGGG 2024c,a). We apply it both with default settings and with region-aware settings to promote county preservation (MGGG 2024b). In initial testing, GerryChain had trouble finding an initial districting plan, especially for state house instances, so we apply *Complete* to find one. Another stumbling block is that precincts can be quite large, sometimes having more than half a state house district's population. This includes a 14,258-person precinct in Mississippi, a 28,753-person precinct in Alabama, and a 45,590-person precinct in Georgia. For this reason, we build state house districts from block groups, which also follows Cannon et al. (2023). Meanwhile, we build state senate districts from precincts. We run short bursts for a total of 100,000 steps and a burst length of 10, following Cannon et al. (2023). We also compare against the enacted 2022 state legislative districts, as collected from the US Census Bureau (2022).

Table 2 reports results across four states and two district types. The reported criteria are the number of majority-Black districts, the average Polsby-Popper score, and the total number of county splits. We compare the performance of the 2022 enacted plan ("Enacted"), short bursts ("SB"), region-aware short bursts ("SB-RA") with a county surcharge of 0.5 (as 1.0 often caused GerryChain to get stuck), and our proposed approach *Carve-Complete-Cleanup* ("CCC"). Recall that the short bursts variants run for 100,000 steps, generating a total of 100,000 plans beyond the initial plan. Which of them should we report scores for? We choose the one with the largest number of majority-Black districts, breaking ties by largest average Polsby-Popper score. We also report the number of majority-Black districts that would be proportional based on the state's total Black population, rounded to the nearest integer. On three instances (MS/SS, MS/SH, GA/SH), our default implementation generates one multi-district that is subdivided into districts by hand.

First, we consider the number of majority-Black districts. We observe that all eight enacted plans are sub-proportional. This comes as no surprise given that all were challenged in court. We see that short bursts and region-aware short bursts usually find more majority-Black districts than the enacted plans (excepting SB-RA for LA/SS and GA/SS), although not always to the level of proportionality. Short bursts and region-aware short bursts find proportional plans for 5/8 and 4/8 of the instances, respectively, while our default implementation finds proportional plans for 6/8.

Next, we turn to district compactness, measured by the average Polsby-Popper score. We remark that, although *Gingles* demonstration plans must be compact and reasonably configured, it is not a *requirement* that they match or outperform the enacted plan on the usual metrics, but this is preferred to be most compelling. We see that short bursts and region-aware short bursts outperform 3/8 and 7/8 of the enacted plans on compactness. Meanwhile, our default implementation outperforms all eight enacted plans, often substantially (e.g., by double for LA/SS, 0.3863 vs. 0.1837).

Instance	Number of Majority-Black Districts					Average Polsby-Popper Score				Total Number of County Splits			
State/Type	Proportional	Enacted	SB	SB-RA	CCC	Enacted	SB	SB-RA	CCC	Enacted	SB	SB-RA	CCC
LA/SS	13	11	12	11	13	0.1837	0.2306	0.2955	0.3863	77	95	46	36
LA/SH	35	29	35	33	39	0.2911	0.2664	0.2960	0.4007	116	173	123	115
MS/SS	20	15	25	21	22	0.2630	0.2215	0.3250	0.3377	64	131	53	46
MS/SH	46	42	51	50	53	0.2644	0.2739	0.3225	0.4137	181	206	134	118
AL/SS	10	8	9	9	9	0.2568	0.2165	0.3156	0.3620	35	88	36	28
AL/SH	29	28	30	30	31	0.2445	0.2507	0.2865	0.3857	115	191	125	105
GA/SS	18	16	20	19	19	0.2870	0.2009	0.2753	0.3530	60	157	70	47
GA/SH	59	54	56	53	57	0.2784	0.2458	0.2902	0.4204	209	388	218	180

Table 2 Comparison of 2022 enacted plans, short bursts, region-aware short bursts, and *Carve-Complete-Cleanup*

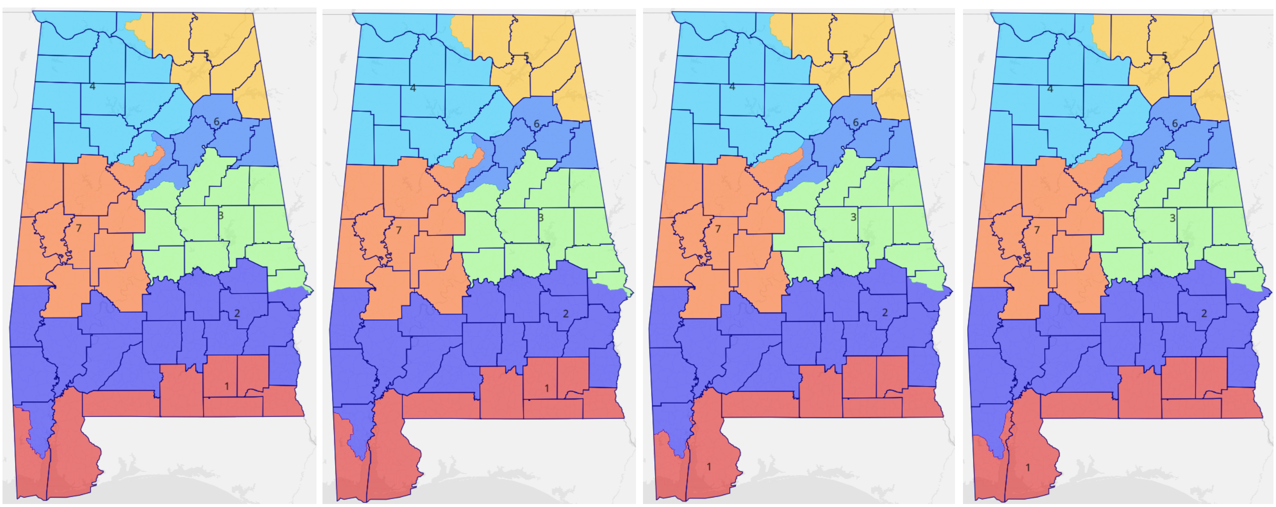
Last, we turn to county preservation. We see that short bursts splits counties more times than the enacted plans, sometimes by double or triple; see, for example, MS/SS, AL/SS, and GA/SS. As expected, the region-aware implementation performs better on county preservation, although still not matching 5/8 of the enacted plans. We suspect that the region surcharge can be tuned somewhere between 0.5 and 1.0 to get better county preservation performance without getting stuck or sacrificing majority-Black districts. Meanwhile, our default implementation outperforms all enacted plans, often by substantial margins (e.g., by double for LA/SS, 36 vs. 77 total splits).

We reiterate that, for the state senate (SS) instances, all precincts are kept whole. Meanwhile, on the state house (SH) instances, these proof-of-concept experiments use block groups, so many precincts end up being split. In our experience, our default implementation can be applied to the precinct level for the SH instances as well (as we already saw for LA/SH in the running example), although sometimes generating a few more multi-districts. These multi-districts could be subdivided by hand. Alternatively, our approach (e.g., *Complete*) could subdivide these multi-districts by handling them at a finer level of granularity (e.g., at the block level).

5.1. Cleanup for Alabama Congressional (Plan D)

We remark that one can pick and choose which phases of our approach to use. For example, one could draft a map by hand (or with alternative computer methods such as GerryChain's short bursts implementation) and subsequently polish it using our MISOCP-based *Cleanup* procedure. To illustrate, we apply *Cleanup* to Moon Duchin's Plan D, which she drew for *Gingles* demonstration purposes when challenging Alabama's enacted congressional districts in the Supreme Court case *Allen v. Milligan* (2023). We work with the block-level graph, which is quite large, having 185,976 vertices. Despite this, our implementation manages because not all variables of model (5) are created, as most of them are fixed to one or zero in the h-hop neighborhood (or are effectively projected out of the model when the cores of the districts are merged into super-nodes). As before, we maintain the total number of county splits by allowing a block $i \in V$ to be assigned to district

$j \in [k]$ only if another block from its county was assigned to district j. We require that the majority-Black districts remain majority-Black and enforce a 1-person deviation. With 1-hop local search, we improve the average Polsby-Popper score from 0.2520 to 0.2830, already outperforming the "Livingston Plan" (score of 0.2817) that Alabama enacted in 2023 after their previous plan was overturned. By increasing the size of the local search parameter h, we can go further. With $h = 2$, the average Polsby-Popper score increases to 0.3003. With $h = 3$, it increases to 0.3239. With $h = 4$, it increases to 0.3288. Figure 7 shows how the district lines smooth out as h increases.



(a) 1-hop (b) 2-hop (c) 3-hop (d) 4-hop

Figure 7 Application of *Cleanup* to Duchin's Plan D.

6. Conclusion and Future Work

Although the Polsby-Popper score is the most popular compactness score in the academic literature and in expert testimony (Polsby and Popper 1991, Duchin 2018), it has largely been absent from the operations research and mathematical programming literature, presumably because of its nonlinear nature. We show how to capture the Polsby-Popper, inverse Polsby-Popper, and Schwartzberg scores in mixed-integer second-order cone programs (MISOCPs), which optimization solvers began supporting only recently. We demonstrate that the MISOCPs can be used to draw optimally compact plans using counties and optimally compact districts using precincts. With additional constraints, they can also identify optimally compact majority-minority districts.

We then turn to the task of drawing compact districting plans with large numbers of majority-minority districts for *Gingles* demonstration purposes. We show that, with a few intuitive tweaks, the MISOCP can be applied repeatedly to find a large number of majority-Black districts, in a procedure that we call *Carve*. These partial districting plans are extended to full districting plans in the MIP-based procedure called *Complete*. These full plans are made significantly more compact using a MISOCP-based local search procedure that we call *Cleanup*. Although this *Cleanup* task is proven to be NP-hard, we find that commercial optimization solvers can successfully optimize over the proposed local search neighborhood, even for instances with nearly two hundred thousand census blocks such as Alabama. Ultimately, we arrive at a powerful computer tool for drafting *Gingles* demonstration plans. On several notable VRA instances, it performs as well or better than enacted plans and state-of-the-art computer tools (like GerryChain's Gingleator) on criteria such as the number of majority-Black districts, Polsby-Popper compactness, and county preservation.

These newfound abilities may find applications in Section 2 litigation when using the VRA to overturn political districting plans that dilute the voting power of minority groups. We envision several. First, the computer-generated plans could serve as visual guides that mapmakers look to for inspiration. Second, mapmakers could apply the *Cleanup* procedure to their own hand-drawn or computer-generated maps (that were generated, say, using GerryChain) to improve their compactness scores. Third, in the event that the proposed approach generates more majority-Black districts than a proportionality baseline, the user could select which ones they think are most reasonably configured, drop the rest, and apply *Complete* and *Cleanup* to finish and polish the plan. Fourth, the approach need not be applied to the entire state; the user could use it to scramble and redraw districts in select portions of a state to promote least-change or modular remedies.

In future work, it would be interesting (and legally impactful) to use optimization to understand the fundamental tradeoffs between district compactness (e.g., Polsby-Popper score) and other criteria (e.g., population deviation, minority representation, political subdivision preservation, partisan fairness). We suspect this to be a difficult task, requiring integer programming methods beyond those developed in this paper. In recent literature, e.g., Schutzman (2020), Becker and Solomon (2022), Swamy et al. (2023), McCartan (2023), researchers have proposed heuristic and optimization-inspired methods to estimate the associated Pareto frontiers, but without guarantees. Another opportunity for future work is to extend the approach to assist in the drawing of VRA-compliant plans, allowing minority-*opportunity* districts and not necessarily *majority*-minority districts (Grofman et al. 2001, Lublin et al. 2020, Becker et al. 2021).

Endnotes

1. When calculating BVAP_i, we use the "any part" racial classification, which was also used in Moon Duchin's expert report for *Allen v. Milligan* (2023) (Duchin 2021). For example, someone who checked both the "Black or African American" and "White" boxes on their Census form would be included in the BVAP_i tally. Guidance from the US Department of Justice also suggests this practice when a respondent selects a minority race and a white race, but when someone selects multiple minority group boxes the DOJ will "allocate these responses on an iterative basis to each of the component single-race categories for analysis" (U.S. Department of Justice 2021).

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 1942065. We thank Daryl DeFord for sharing the districting instances as `json` files, Peter Rock for help with map projections, and Moon Duchin for sharing block assignment files for her demonstration districts in *Allen v. Milligan* (2023). This analysis was conducted using data from the Redistricting Data Hub.

References

Adler WT, Wang SSH (2019) Response to Cho and Liu, "Sampling from complicated and unknown distributions: Monte Carlo and Markov chain Monte Carlo methods for redistricting". *Physica A: Statistical Mechanics and its Applications* 516:591–593.

Altman M (1997) The computational complexity of automated redistricting: Is automation the answer? *Rutgers Computer & Tech. LJ* 23:81.

Ancheta AN (2007) Language accommodation and The Voting Rights Act. Henderson A, ed., *Voting Rights Act Reauthorization Of 2006: Perspectives on Democracy, Participation, and Power*, 293–325 (Univ of California Regents School).

Arredondo V, Martínez-Panero M, Peña T, Ricca F (2021) Mathematical political districting taking care of minority groups. *Annals of Operations Research* 305(1-2):375–402.

Autry EA, Carter D, Herschlag GJ, Hunter Z, Mattingly JC (2021) Metropolized multiscale forest recombination for redistricting. *Multiscale Modeling & Simulation* 19(4):1885–1914.

Bandle C (2017) Dido's problem and its impact on modern mathematics. *Notices of the AMS* 64(9).

Bar-Natan A, Najt L, Schutzman Z (2020) The Gerrymandering jumble: map projections permute districts' compactness scores. *Cartography and Geographic Information Science* 47(4):321–335.

Barnes R, Solomon J (2021) Gerrymandering and compactness: Implementation flexibility and abuse. *Political Analysis* 29(4):448–466.

Becker A, Duchin M, Gold D, Hirsch S (2021) Computational redistricting and the Voting Rights Act. *Election Law Journal: Rules, Politics, and Policy* 20(4):407–441.

Becker A, Solomon J (2022) Redistricting algorithms. Duchin M, Walch O, eds., *Political Geometry*, 303–340 (Birkhauser).

Belotti P, Kirches C, Leyffer S, Linderoth J, Luedtke J, Mahajan A (2013) Mixed-integer nonlinear optimization. *Acta Numerica* 22:1–131.

Benson HY, Sağlam Ü (2013) Mixed-integer second-order cone programming: A survey. *INFORMS TutORials in Operations Research* 13–36.

Birge JR (1983) Redistricting to maximize the preservation of political boundaries. *Social Science Research* 12(3):205–214.

Buchanan A (2023) Using optimization to support minority representation in Voting Rights Act Cases. *ORMS Today* 50(4):32–35.

Buchanan A, Sung JS, Butenko S, Pasiliao EL (2015) An integer programming approach for fault-tolerant connected dominating sets. *INFORMS Journal on Computing* 27(1):178–188.

Cannon S, Duchin M, Randall D, Rule P (2022) Spanning tree methods for sampling graph partitions. *arXiv preprint arXiv:2210.01401* .

Cannon S, Goldbloom-Helzner A, Gupta V, Matthews J, Suwal B (2023) Voting rights, Markov chains, and optimization by short bursts. *Methodology and Computing in Applied Probability* 25(1):36.

Caro F, Shirabe T, Guignard M, Weintraub A (2004) School redistricting: Embedding GIS tools with integer programming. *Journal of the Operational Research Society* 55(8):836–849.

Carvajal R, Constantino M, Goycoolea M, Vielma JP, Weintraub A (2013) Imposing connectivity constraints in forest planning models. *Operations Research* 61(4):824–836.

Chen J, Stephanopoulos NO (2021) The race-blind future of voting rights. *Yale Law Journal* 130:862–947.

Chikina M, Frieze A, Pegden W (2017) Assessing significance in a Markov chain without mixing. *Proceedings of the National Academy of Sciences* 114(11):2860–2864.

Cho WKT, Liu YY (2018) Sampling from complicated and unknown distributions: Monte Carlo and Markov chain Monte Carlo methods for redistricting. *Physica A: Statistical Mechanics and its Applications* 506:170–178.

Clelland J, Colgate H, DeFord D, Malmskog B, Sancier-Barbosa F (2022) Colorado in context: Congressional redistricting and competing fairness criteria in Colorado. *J Computational Social Science* 5(1):189–226.

Clelland JN, Bossenbroek N, Heckmaster T, Nelson A, Rock P, VanAusdall J (2021) Compactness statistics for spanning tree recombination. *arXiv:2103.02699* .

Cova TJ, Church RL (2000) Contiguity constraints for single-region site search problems. *Geographical Analysis* 32(4):306–329.

Cox E (1927) A method of assigning numerical and percentage values to the degree of roundness of sand grains. *Journal of Paleontology* 1(3):179–183.

Davis M, Strigari F, Underhill W, Wice JM, Zamarripa C (2019) *Redistricting Law 2020* (NCSL).

DeFord D, Becker A, Gold D (2022) Brief of Computational Redistricting Experts as Amici Curiae in support of Appellees and Respondents in *Merrill v. Milligan* and *Merrill v. Caster*.

DeFord D, Duchin M (2022) Random walks and the universe of districting plans. Duchin M, Walch O, eds., *Political Geometry*, 341–381 (Birkhauser).

DeFord D, Duchin M, Solomon J (2018) Comparison of districting plans for the virginia house of delegates. *MGGG Technical Report* 12(II).

DeFord D, Duchin M, Solomon J (2021) Recombination: A family of Markov chains for redistricting. *Harvard Data Science Review* 3(1).

DeFord D, Kimsey E, Zerr R (2023) Multi-balanced redistricting. *Journal of Computational Social Science* 6(2):923–941.

DRA (2023) Dave's redistricting. `https://davesredistricting.org/`, accessed: 2023-03-29.

Drewes S (2009) *Mixed Integer Second Order Cone Programming*. Ph.D. thesis, TU Darmstadt.

Duchin M (2018) Outlier analysis for Pennsylvania congressional redistricting. *LWV vs. Commonwealth of Pennsylvania Docket No. 159 MM 2017* .

Duchin M (2021) Presentation of Alternative Congressional Districting Plans for Alabama. Available in Supplemental Joint Appendix for *Merrill v. Milligan* at `https://www.supremecourt.gov/DocketPDF/21/21-1086/221826/20220425150837756_42140%20pdf%20Bowdre%20IV%20Supplemental%20JA.pdf`.

Duchin M (2022) Explainer: Compactness, by the numbers. Duchin M, Walch O, eds., *Political Geometry*, 29–35 (Birkhauser).

Duchin M, Spencer D (2021) Models, race, and the law. *Yale Law Journal Forum* 130:744–797.

Duchin M, Walch O (2022) Interview: Drawing for the courts. Duchin M, Walch O, eds., *Political Geometry*, 409–414 (Birkhauser).

Dyer ME, Frieze AM (1985) On the complexity of partitioning graphs into connected subgraphs. *Discrete Applied Mathematics* 10(2):139–153.

Faenza Y, Kaibel V (2009) Extended formulations for packing and partitioning orbitopes. *Mathematics of Operations Research* 34(3):686–697.

Fifield B, Higgins M, Imai K, Tarr A (2015) A new automated redistricting simulator using Markov chain Monte Carlo. *Work. Pap., Princeton Univ., Princeton, NJ* .

Fischetti M, Leitner M, Ljubić I, Luipersbeck M, Monaci M, Resch M, Salvagnin D, Sinnl M (2017) Thinning out Steiner trees: a node-based model for uniform edge costs. *Math Prog Comput* 9(2):203–229.

Fravel J, Hildebrand R, Goedert N, Travis L, Pierson M (2024) Optimizing representation in redistricting: Dual bounds for partitioning problems with non-convex objectives. *arXiv:2305.17298v2* .

Freeman N (2014) Nobody lives here: The nearly 5 million census blocks with zero population. `https://tumblr.mapsbynik.com/post/82791188950/nobody-lives-here-the-nearly-5-million-census`.

Garfinkel RS, Nemhauser GL (1970) Optimal political districting by implicit enumeration techniques. *Management Science* 16(8):B–495.

Gordon A, Spencer DM (2022) Explainer: A brief introduction to the Voting Rights Act. *Political Geometry: Rethinking Redistricting in the US with Math, Law, and Everything In Between*, 131–136 (Birkhauser).

Grofman B (1982) For single member districts random is not equal. *Representation and Redistricting Issues*, 55–58 (Lexington Books).

Grofman B, Handley L, Lublin D (2001) Drawing effective minority districts: A conceptual framework and some empirical evidence. *North Carolina Law Review* 79:1383.

Gurnee W, Shmoys DB (2021) Fairmandering: A column generation heuristic for fairness-optimized political districting. *SIAM Conference on Applied and Computational Discrete Algorithms*, 88–99 (SIAM).

Hebert JG, Vandenberg ME, Smith P (2010) *The Realist's Guide to Redistricting: Avoiding the Legal Pitfalls* (American Bar Association).

Henzinger A, Noe A, Schulz C (2020) ILP-based local search for graph partitioning. *Journal of Experimental Algorithmics (JEA)* 25:1–26.

Hess S, Weaver J, Siegfeldt H, Whelan J, Zitlau P (1965) Nonpartisan political redistricting by computer. *Operations Research* 13(6):998–1006.

Hofeller TB (2015) Declaration of Thomas Brooks Hofeller. Available in Joint Appendix Volume IV of IX for *Virginia House of Delegates v. Bethune-Hill (2019)* at `https://www.supremecourt.gov/DocketPDF/18/18-281/77913/20190102174617179_Vol%204%20Part%201.pdf`.

Hojati M (1996) Optimal political districting. *Computers & Operations Research* 23(12):1147–1161.

Laney GP (2008) The Voting Rights Act of 1965, as amended: Its history and current issues. `https://congressionalresearch.com/95-896/document.php`.

Lublin D, Handley L, Brunell TL, Grofman B (2020) Minority success in non-majority minority districts: Finding the "sweet spot". *Journal of Race, Ethnicity, and Politics* 5(2):275–298.

McCartan C (2023) Finding Pareto efficient redistricting plans with short bursts. *arXiv preprint arXiv:2304.00427* .

McCartan C, Imai K (2023) Sequential Monte Carlo for sampling balanced and compact redistricting plans. *The Annals of Applied Statistics* 17(4):3300–3323.

McCartan C, Kenny C, Simko T, Kuriwaki S, Garcia III G, Wang K, Wu M, Imai K (2022a) ALARM project. `https://alarm-redist.github.io/fifty-states/`, accessed: 2023-03-29.

McCartan C, Kenny CT, Simko T, Garcia III G, Wang K, Wu M, Kuriwaki S, Imai K (2022b) Simulated redistricting plans for the analysis and evaluation of redistricting in the US. *Scientific Data* 9(1):689.

McDonald M (2019) The predominance test: A judicially manageable compactness standard for redistricting. *Yale Law Journal Forum* 129:18.

Mehrotra A, Johnson EL, Nemhauser GL (1998) An optimization based heuristic for political districting. *Management Science* 44(8):1100–1114.

MGGG (2024a) Gerrychain v0.3.2. URL `https://gerrychain.readthedocs.io/en/latest/`.

MGGG (2024b) Region-Aware ReCom. URL `https://gerrychain.readthedocs.io/en/latest/user/recom/#region-aware-recom`.

MGGG (2024c) Using Gingleator. URL `https://gerrychain.readthedocs.io/en/latest/user/optimizers/#using-gingleator`.

Niemi RG, Grofman B, Carlucci C, Hofeller T (1990) Measuring compactness and the role of a compactness standard in a test for partisan and racial Gerrymandering. *The Journal of Politics* 52(4):1155–1181.

Oehrlein J, Haunert JH (2017) A cutting-plane method for contiguity-constrained spatial aggregation. *Journal of Spatial Information Science* 2017(15):89–120.

Önal H, Patrick KT (2016) A mathematical programming approach to political redistricting with compactness and community integrity considerations. Technical report, UIUC.

Osserman R (1978) The isoperimetric inequality. *Bulletin of the AMS* 84(6):1182–1238.

Persily N (2004) When judges carve democracies: a primer on court-drawn redistricting plans. *George Washington Law Review* 73:1131.

Polsby DD, Popper RD (1991) The third criterion: Compactness as a procedural safeguard against partisan Gerrymandering. *Yale Law & Policy Review* 9:301.

Princeton Gerrymandering Project (2023) Redistricting report card. https://gerrymander.princeton.edu/redistricting-report-card/, accessed: 2023-02-17.

Procaccia AD, Tucker-Foltz J (2022) Compact redistricting plans have many spanning trees. *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 3754–3771 (SIAM).

Redensky Y, Leaverton C (2023) Ongoing Voting Rights Act Redistricting Litigation After SCOTUS Ruling in Allen v. Milligan. https://www.brennancenter.org/our-work/research-reports/ongoing-voting-rights-act-redistricting-litigation-after-scotus-ruling, accessed: 2024-05-08.

Redistricting Data Hub (2021) Download Data. https://redistrictingdatahub.org/data/.

Ricca F, Scozzari A, Simeone B (2013) Political districting: from classical models to recent approaches. *Annals of Operations Research* 204(1):271–299.

Schutzman Z (2020) Trade-offs in fair redistricting. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 159–165.

Schwartzberg JE (1965) Reapportionment, Gerrymanders, and the notion of compactness. *Minnesota Law Review* 50:443.

Shahmizad M, Buchanan A (2023) Political districting to minimize county splits. Optimization-Online.

Shirabe T (2005) A model of contiguity for spatial unit allocation. *Geographical Analysis* 37(1):2–16.

Shirabe T (2009) Districting modeling with exact contiguity constraints. *Environment and Planning B: Planning and Design* 36(6):1053–1066.

Smith EM, Pantelides CC (1999) A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex MINLPs. *Computers & Chemical Engineering* 23(4-5):457–478.

Swamy R, King DM, Jacobson SH (2023) Multiobjective optimization for politically fair districting: a scalable multilevel approach. *Operations Research* 71(2):536–562.

Temiz N, Trapp A, Prokopyev OA, Camacho CJ (2010) Optimization of minimum set of protein–DNA interactions: a quasi exact solution with minimum over-fitting. *Bioinformatics* 26(3):319–325.

US Census Bureau (2021a) Decennial Census P.L. 94-171 Redistricting Data. `https://www.census.gov/programs-surveys/decennial-census/about/rdo/summary-files.html`.

US Census Bureau (2021b) TIGER/Line Shapefiles. `https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html`.

US Census Bureau (2022) State legislative districts. `https://www.census.gov/programs-surveys/decennial-census/about/rdo/state-legislative-district.html`.

US Department of Justice (2021) Guidance under Section 2 of the Voting Rights Act, 52 U.S.C. 10301, for redistricting and methods of electing government bodies. Available at `https://www.justice.gov/opa/press-release/file/1429486/download`.

Validi H, Buchanan A (2022) Political districting to minimize cut edges. *Mathematical Programming Computation* 14(4):623–672.

Validi H, Buchanan A, Lykhovyd E (2022) Imposing contiguity constraints in political districting models. *Operations Research* 70(2):867–892.

Vickrey W (1961) On the prevention of Gerrymandering. *Political Science Quarterly* 76(1):105–110.

Wang Y, Buchanan A, Butenko S (2017) On imposing connectivity constraints in integer programs. *Mathematical Programming* 166(1-2):241–271.

Weaver JB, Hess SW (1963) A procedure for nonpartisan districting: Development of computer techniques. *Yale Law Journal* 73:288.

Young HP (1988) Measuring the compactness of legislative districts. *Legislative Studies Quarterly* 13(1):105–115.

Zhang J, Validi H, Buchanan A, Hicks IV (2024) Linear-size formulations for connected planar graph partitioning and political districting. *Optimization Letters* 18:19–31.

Zoltners AA, Sinha P (1983) Sales territory alignment: A review and model. *Management Science* 29(11):1237–1256.