Political districting to optimize the Polsby-Popper compactness score

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 with optimum Polsby-Popper score. Specifically, we propose new mixed-integer second-order cone programs (MISOCPs), which can be solved with existing optimization software. We investigate their tractability by applying them to real-life districting instances in the USA. Experiments show that they are able to: (i) identify the most compact majority-minority districts at the tract level; (ii) identify the most compact districting plans at the county level; and (iii) refine existing tract-level plans to make them substantially more compact. Our techniques could be used by plaintiffs when seeking to overturn maps that dilute the voting strength of minority groups. Our python code and results are publicly 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 already been proposed for measuring the roundness of sand grains in a paleontology journal 64 years prior (Cox 1927). Mathematically, 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. 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 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 (or “demonstration districts”) showing that such a majority-minority district is possible. 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.

At the time of writing, the Supreme Court is deciding the case *Merrill v. Milligan* about whether Alabama must draw a second Black-opportunity district to satisfy the VRA. Much of the expert testimony concerns 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, 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).

Our Contributions. In light of the foregoing discussion, it is clear that political districting involves numerous tradeoffs between competing criteria and that illuminating 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 (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: (i) identify the most compact majority-minority districts at the tract level; (ii) identify the most compact districting plans at the county level; and (iii) refine existing tract-level plans to make them substantially more compact. These abilities may assist in Section 2 litigation when satisfying the first *Gingles* precondition or assist in the drawing of remedial districts. Our python code and computational results are publicly available at: `https://github.com/AustinLBuchanan/Polsby_Popper_optimization`.

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 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, cf. 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 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 puts 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 has 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 is 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 ultimately require the maps to be redrawn (Persily 2004). In the remedial districts, 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 but is nevertheless able to elect preferred candidates with crossover support from other groups (Grofman et al. 2001, Lublin et al. 2020).

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 has been 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 from 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*.

Presently, the Supreme Court is deciding the case *Merrill v. Milligan*. Accordingly, the court's precedent regarding the VRA and the Equal Protection Clause may change at any moment (Chen and Stephanopoulos 2021, Duchin and Spencer 2021, 2022, Gordon and Spencer 2022).

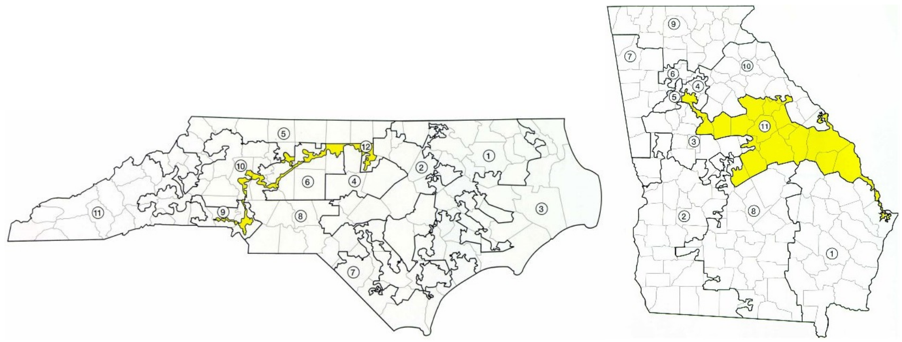

Figure 1 The US Supreme Court overturned North Carolina's 12th district in *Shaw v. Reno* (1993) and Georgia's 11th district in *Miller v. Johnson* (1995).

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.

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

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

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

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. Interestingly, the Schwartzberg score was proposed back in 1965, decades before the Polsby-Popper score, and Schwartzberg expressed concern in accurately measuring district perimeters and proposed a discretized "gross perimeter" instead.

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). Duchin (2021), in her expert report in the case *Milligan v. Merrill* (before it was appealed to the US Supreme Court), 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). Schutzman (2020) also uses the average Polsby-Popper score. Barnes and Solomon

(2021), when calculating the Polsby-Popper score of a non-contiguous district, take the average across the district's connected components.

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.

Chikina et al. (2017) and Vagnozzi (2020) use the equivalent *sum* of inverse Polsby-Popper scores.

Later, we observe that the average Polsby-Popper score may sacrifice the compactness of a few districts to perfect the others. We also show that Barnes-Solomon averaging leads to pathological behavior, preferring bizarre "Swiss cheese" districts.

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 framework (MCMC) that moves from one districting plan to a neighboring or similar districting plan in a random walk (Fifield et al. 2015, Liu et al. 2016, 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 minority opportunity districts (Becker et al. 2021, Cannon et al. 2023).

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), 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. (2022) 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, either 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 we solve them without linearization. With the additional expressiveness afforded by the second-order cone constraints, we can propose the first optimization models for handling the Polsby-Popper score and its 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 $(D_1, D_2, \dots, D_k)$. 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

Throughout 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) and final processing conducted by Daryl DeFord. In our computational experiments, we typically use an ordinary desktop PC (Windows 10 enterprise, Intel Core i7-4790 CPU at 3.6 GHz, 16 GB RAM). To solve MISOCPs, we use Gurobi v10.0.1. (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, reads input data using the GerryChain package, and is available at `https://github.com/AustinLBuchanan/Polsby_Popper_optimization`. We typically consider congressional redistricting instances for which the number of districts k is publicly known and determined after the 2020 Census. We typically impose a 1% total population deviation (±0.5%), setting $L = \lceil 0.995 \cdot p(V)/k \rceil$ and $U = \lfloor 1.005 \cdot p(V)/k \rfloor$, under the relatively standard hypothesis that 1-person deviation could be achieved with minor tweaks.

3. The Perils of Straight Averaging

In this section, we make some qualitative observations about straight averaging. First, we observe that averaging the Polsby-Popper scores of the districts, as done by many (DRA 2023, McCartan et al. 2022a, Princeton Gerrymandering Project 2023, Schutzman 2020), encourages one to sacrifice the compactness of a few districts to perfect the shape of the others. This may lead to districting plans that fail the eyeball test but nevertheless score well. Second, we observe that the practice of averaging the Polsby-Popper scores of a disconnected district's components, as done by Barnes and Solomon (2021), leads the Polsby-Popper score to prefer bizarre "Swiss cheese" districts.

3.1. Sacrificial Districts

Consider the problem of drawing equal-area districts for the Euclidean plane to maximize the average Polsby-Popper score. One approach is to tile the plane with equal-sized squares, which would lead to an average Polsby-Popper score of $\pi/4 \approx 0.7854$. Better yet, we could tile the plane with equal-sized hexagons, as in Figure 2, giving an average Polsby-Popper score of $\pi\sqrt{3}/6 \approx 0.9069$.

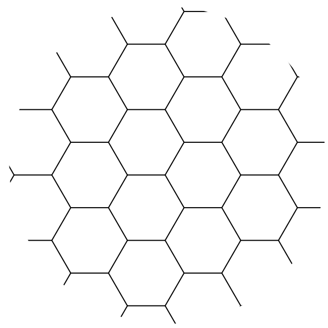

Figure 2 Tiling the plane by hexagon-shaped districts gives an average Polsby-Popper score of $\pi\sqrt{3}/6 \approx 0.9069$.

Can we do better? To do so, at least some of the district shapes would need to be "nearer" to a circle than that of a hexagon. While it is not possible to tile the plane with disks themselves, one can get close to a tiling, with more than 90% of the plane covered. Specifically, this is achieved by the hexagonal packing arrangement in which each circle is surrounded by six other circles, as in Figure 3. This arrangement achieves the highest density among lattice packings with a density of $\pi\sqrt{3}/6 \approx 0.9069$, which turns out to equal the Polsby-Popper score of a hexagon. With the remaining "scraps" of the plane, one can piece together highly non-compact "districts" whose Polsby-Popper scores are small, but nevertheless positive. Thus, a circles-and-scraps districting plan can achieve a slightly higher *average* Polsby-Popper score than a tiling by hexagons, as 90.69% of the districts achieve a perfect score of one, and the remaining 9.31% of the districts achieve a positive score. One could complain that the scraps are disconnected and that their districts will be as well. In response to this complaint, one could adjust the districts by keeping the circle centers where they are, but shrinking their radii by $\varepsilon > 0$ to connect the scraps to each other.

In a more concrete example, consider the rectangle given in Figure 4 which has height $H = 2 + 2\sqrt{3}$ and width $W = 8$. It admits a packing with 11 circles of radius $r = 1$. The rectangle itself has an area of $H \cdot W \approx 43.7128$ which is slightly less than 14π. So, by slightly elongating its height and width, we can generate a rectangular districting instance in which the task is to partition the rectangle into 14 districts, each with an area of π. In a circles-and-scraps districting plan, we can have 11 circles and 3 scrap districts, so the average Polsby-Popper score would be larger than $11/14 \approx 0.7857$. Meanwhile, any partition of this state into rectangular districts would have

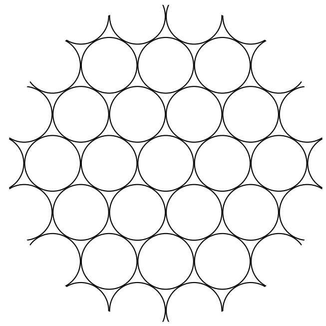

Figure 3 Packing circular districts in a hexagonal lattice arrangement (and constructing the remaining districts from the scraps) gives an average Polsby-Popper score of more than $\frac{\pi\sqrt{3}}{6} \approx 0.9069$.

a worse average Polsby-Popper score, as the highest-scoring rectangle is in fact a square which has Polsby-Popper score $\pi/4 \approx 0.7854$. This shows the extent to which the average Polsby-Popper score may sacrifice the compactness of a few districts to perfect the others.

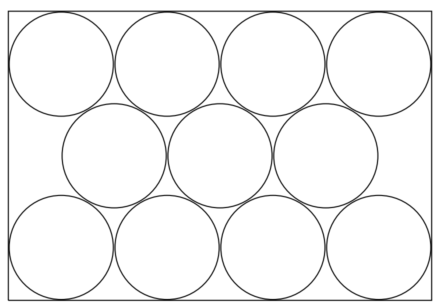

Figure 4 Circle packing for rectangle with side lengths $H = 2 + 2\sqrt{3}$ and $W = 8$.

3.2. Swiss Cheese Districts

As previously mentioned, Barnes and Solomon (2021) compute the Polsby-Popper score of a non-contiguous district by taking the average over its connected components. Here we show an undesirable consequence of this choice. Consider the rectangular state in Figure 5 that is split into $k = 2$ square districts. The Polsby-Popper score of each district is $\pi/4 \approx 0.7854$.

Are higher scores possible? Yes, in fact we can get arbitrarily close to one. The idea is to cut circles from the left and right sides and assign these holes to the opposite sides. As the number of holes grows, the Polsby-Popper scores of these "Swiss cheese" plans approach one. For example, the districts in Figure 5 have Polsby-Popper scores greater than 0.90 under Barnes-Solomon averaging.

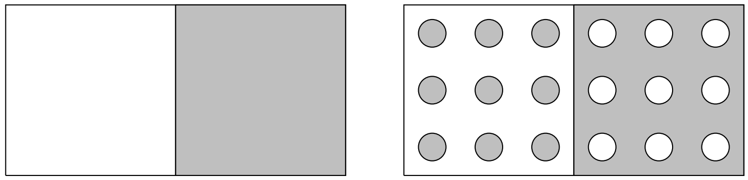

Figure 5 Straight averaging prefers "Swiss cheese" plans

Meanwhile, if we directly compute the Polsby-Popper scores, ignoring the fact that the Swiss cheese districts are disconnected, then we get scores that better reflect their bizarre shapes; if the squares have side lengths of 6 and the holes have radius 1/2, then each district has area $A = 36$ and perimeter $P = 4(6) + 9\pi + 9\pi \approx 80.5487$, giving a Polsby-Popper score of roughly 0.0697.

4. MISOCP for Drawing a Single District

As a warm-up, we propose a mixed-integer second-order cone program (MISOCP) for drawing a single district with optimum Polsby-Popper score. To understand its computational performance, we apply it to various instances in the USA. First, we find an optimally compact district for Oklahoma. Second, we draw an Asian-majority district in California, similar to the current district CA-17. Next, we show how to draw a second Hispanic-majority district in Chicago, Illinois and a third Native-American-majority state senate district in South Dakota. Finally, motivated by the pending Supreme Court case *Merrill v. Milligan*, we draw Black-majority districts for Alabama. We strongly suspect that each of these districts can be extended into feasible districting *plans*, see, e.g., McDonald (2019). Recall that the VRA does not require such *majority-minority* districts to be enacted, but demonstrating that they are possible to draw is nevertheless a step that plaintiffs must fulfill to have their case heard, per the Supreme Court's rulings in *Gingles* and *Bartlett*.

4.1. The MISOCP Model

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., $L \leq p(D) \leq U$). 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.

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

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 forces it to hold at equality). Constraints (1c) and (1d) capture the perimeter and area of the district, respectively. Constraints (1e), which are imposed for each orientation of edge $e \in E$, indicate that if one endpoint of $e = \{u, v\}$ belongs to the district but the other endpoint does not, then the edge e between them is cut. Strictly speaking, 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 achieve contiguity that we experiment with later. Also, we will see that the compactness objective often achieves contiguous solutions on its own.

4.2. A compact district for Oklahoma

In our first application, we draw a single congressional district for Oklahoma with optimum Polsby-Popper score. We use census tracts as our geographic units, of which there are $1,205$. 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 44 seconds of computation on our desktop PC, the MISOCP solver Gurobi v10.0.1 returns an optimal solution with an inverse Polsby-Popper score of roughly 1.3466, i.e., a Polsby-Popper score of roughly 0.7426, as depicted in Figure 6. By chance, the district mostly follows the boundaries of Oklahoma County, which has a rectangular shape. It was not necessary to impose explicit contiguity constraints; the compactness objective achieved contiguity "for free".

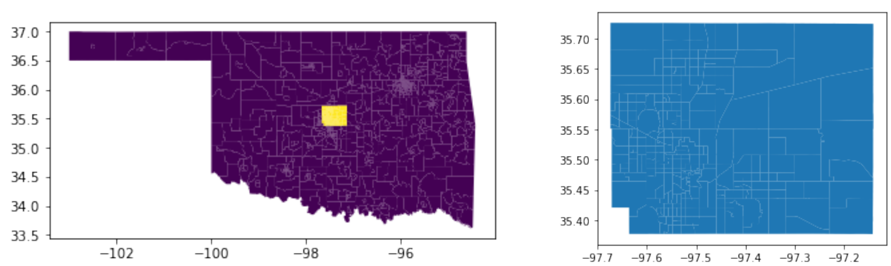

Figure 6 A district for Oklahoma (with optimum Polsby-Popper score).

4.3. Asian-majority district in California

In our second application, we draw an Asian-majority congressional district for California. Following the Supreme Court's opinion in *Bartlett*, we impose that the minority group's voting age population (VAP) is at least 50% of the district's voting age population. Mathematically, denote by m_i^{VAP} as *this* minority group's VAP and p_i^{VAP} as the total VAP in geographic unit $i \in V$.

When calculating m_i^{VAP}, we use the "any part" racial classification, which was also used in Moon Duchin's expert report for *Milligan v. Merrill* (Duchin 2021). For example, someone who checked both the "Asian" and "White" boxes on their Census form would be included in the Asian m_i^{VAP}. Guidance from the US Department of Justice for Section 2 enforcement 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). The Citizen Voting Age Population (CVAP) also plays a role in VRA case law. Since the decennial census does not have a citizenship question, this information is typically inferred using American Community Survey (ACS) data. However, we use only VAP data in this paper. This is one reason why our analysis should be considered preliminary and not legally conclusive.

Then, to obtain a majority-minority district, we impose the constraint:

$$\sum_{i \in V} m_i^{\text{VAP}} x_i \geq \frac{1}{2} \sum_{i \in V} p_i^{\text{VAP}} x_i.$$

We also require the district to be contained within Santa Clara County, as does the current Asian-majority district CA-17. Santa Clara County has a total of 408 census tracts. After two seconds, we obtain the district depicted in Figure 7 that has a Polsby-Popper score of 0.5933. Again, explicit contiguity constraints were not required.

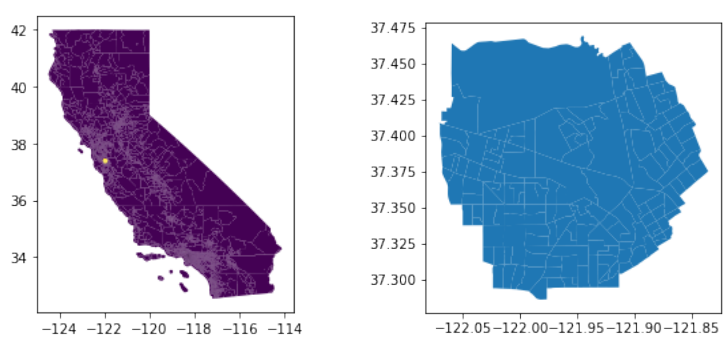

Figure 7 An Asian-majority district in California, similar to CA-17.

4.4. Hispanic-majority districts in Chicago, Illinois

In our third application, we consider Hispanic-majority districts in Chicago, Illinois. Presently, there is one Hispanic-majority district IL-4 which touches Cook County and DuPage County. In the previous redistricting cycle, this district was the famous "earmuff" district, drawn to connect two Latino communities—largely Puerto Rican in the north and Mexican-American in the south—without diluting the vote of a largely African-American community in-between (Levine 2018). The 2010 and 2020 vintages of this district are depicted in Figure 8.

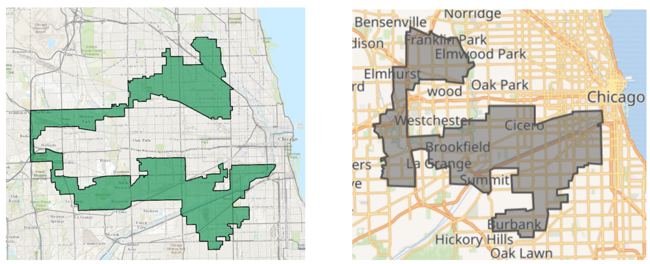

Figure 8 The Hispanic-majority district IL-4 in Illinois, past and present.

Using the MISOCP model, we draw a Hispanic-majority district within Cook and DuPage Counties that has a Polsby-Popper score of 0.7363, as depicted in Figure 9. Curious if we could draw two Hispanic-majority districts, we attempted to find a Hispanic-majority "double district", i.e., with twice the requisite population, under the hypothesis that it could subsequently be split into two Hispanic-majority districts. Indeed, we successfully obtained a double district, also depicted in Figure 9. Both the single and double districts are fully contained within Cook County, and neither needed explicit contiguity constraints.

4.5. Native-American-majority districts in South Dakota

In our fourth application, we consider drawing Native-American-majority districts in South Dakota. Due to its small population, South Dakota receives just one ("at-large") congressional seat. However, South Dakota has 35 seats in their state senate that *do* rely on districts. In the early 2000s, there was one Native-majority ($> 90\%$) senate district, which was overturned under the VRA for

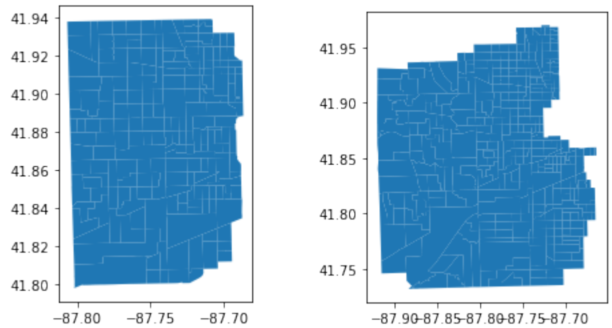

Figure 9 An alternative Hispanic-majority district and double district for Illinois.

packing too many Native Americans into a single district and diluting their vote. In the remedial plan, two districts were drawn to each have a Native-American majority. This remains true today, with the 26th district roughly corresponding to the Rosebud Indian Reservation and the 27th district roughly corresponding to the Pine Ridge Indian Reservation.

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. Pertinently, the current number of Native-majority state senate districts (2) is less than a proportionality baseline of $3.15 = 0.09 \times 35$.

So, we were curious whether a third Native-majority district could be drawn and attempted to draw a Native-majority "triple district" using the MISOCP. Initially, our code returns the triple district on the left of Figure 10, which has a hole in the middle around Pierre, which is the state's capital and seat of Hughes County. In a second attempt, we forbid any part of Hughes County to be in the triple district and add explicit contiguity constraints as a, b-separator inequalities in a cut callback (Validi et al. 2022, Validi and Buchanan 2022). We obtain the triple district on the right side of Figure 10, which separates the state into left and right components. Luckily, the left and right components have populations that are suitable to build 7 districts and 25 districts, respectively. Interestingly, this triple district largely covers not only the Rosebud and Pine Ridge Indian Reservations that the 26th and 27th districts currently touch, but also brings in much of the Standing Rock and Cheyenne River Indian Reservations to their north.

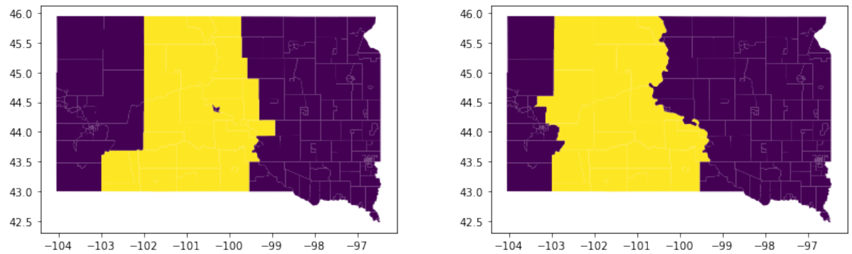

Figure 10 First and second attempts at a Native-American-majority triple district for the South Dakota Senate.

4.6. Black-majority districts in Alabama

In our fifth application, we draw Black-majority congressional districts for Alabama, motivated by the pending Supreme Court case *Merrill v. Milligan*. In 2022, the Alabama legislature enacted a congressional districting plan in which one of their seven districts had a Black majority (district 7). Plaintiffs remarked that the state is more than 27% Black and that the enacted plan places counties from the state's "Black Belt" into majority-White districts. To satisfy the first *Gingles* precondition, mathematician and expert witness Moon Duchin drew alternative maps in which two districts had a Black majority (districts 2 and 7). Figure 11 shows the enacted map and Duchin's Plan D which have average Polsby-Popper scores of 0.222 and 0.249, respectively. The enacted plan splits six counties once each, for a total of six splits. Duchin's Plan D splits four counties once each, and one county twice, again for a total of six splits.

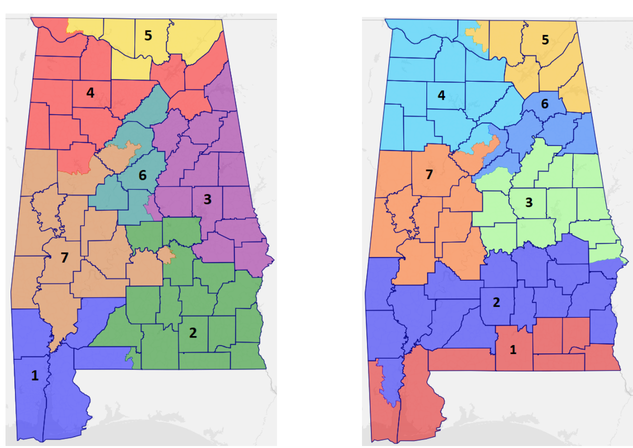

Figure 11 Alabama's enacted plan and Duchin's Plan D

We use the MISOCP to build a Black-majority "double district" using census tracts. To impose contiguity, we use the flow-based constraints of Shirabe (2005, 2009). Given the emphasis on political subdivisions in *Merrill v. Milligan*, we also require that the district consist of whole counties plus two partial counties. To model this restriction, we introduce a binary variable for each county indicating whether it is *split* (i.e., if *some* of the county, but not *all*, is picked in the district). Likewise, we introduce a binary variable indicating whether *some* of the county is picked in the district, and another indicating whether *all* of it is picked. The split variable equals the difference between them (*some* minus *all*). Straightforward logical inequalities relate the *some* and *all* variables to the original x_i variables.

Initially, we obtain the district on the left side of Figure 12, which separates the Gulf Coast from the rest of the state. In our second attempt, we forbid any portion of Baldwin County or Escambia County to be in the double district. We obtain the Black-majority double district on the right of Figure 12. This double district is similar to the two Black-majority districts from Duchin's Plan D. One difference is that Duchin's districts also include the Black Belt counties of Barbour and Russell, better preserving this community of interest.

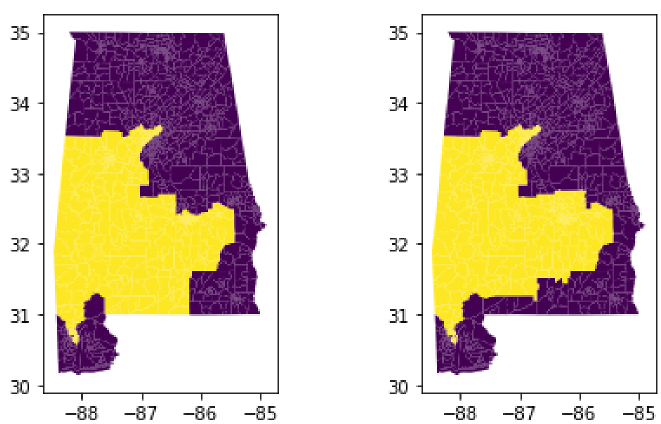

Figure 12 First and second attempts at a Black-majority double district for Alabama.

5. MISOCPs for Drawing 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.

5.1. 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 $j \in [k] = \{1,2,\dots,k\}$. 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:

$$
\begin{aligned}
\min \quad & \frac{1}{k}\sum_{j=1}^{k} z_j && && \text{(2a)}\\
\text{s.t.} \quad & P_j^2 \le 4\pi A_j z_j && \forall j \in [k] && \text{(2b)}\\
& P_j = \sum_{e\in E} b_e y_e^j + \sum_{i\in V} b_i x_{ij} && \forall j \in [k] && \text{(2c)}\\
& A_j = \sum_{i\in V} a_i x_{ij} && \forall j \in [k] && \text{(2d)}\\
& x_{uj} - x_{vj} \le y_e^j && \forall e = \{u,v\} \in E,\ \forall j \in [k] && \text{(2e)}\\
& L \le \sum_{i\in V} p_i x_{ij} \le U && \forall j \in [k] && \text{(2f)}\\
& D_j = \{i \in V \mid x_{ij} = 1\} \text{ is connected} && \forall j \in [k] && \text{(2g)}\\
& \sum_{j=1}^{k} x_{ij} = 1 && \forall i \in V && \text{(2h)}\\
& x_{ij} \in \{0,1\} && \forall i \in V,\ \forall j \in [k] && \text{(2i)}\\
& y_e^j \in \{0,1\} && \forall e \in E,\ \forall j \in [k]. && \text{(2j)}
\end{aligned}
$$

The objective (2a) minimizes the average inverse Polsby-Popper score. The constraints of model (2) are analogous to those of model (1) but are written for k districts instead of one. Another change is that the assignment constraints (2h) 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).

5.2. Polsby-Popper Score

To maximize the average Polsby-Popper score, we would like to

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

subject to the same constraints from model (2). 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_i. 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_{ij} 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.

$$\sum_{i=1}^{t} 2^{-i} b_{ij} = \frac{1}{z_j}.$$

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

$$\sum_{i=1}^{t} 2^{-i} b_{ij} z_j \leq 1.$$

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

$$z_j + M(b_{ij} - 1) \leq w_{ij} \leq M b_{ij}$$

$$0 \leq w_{ij} \leq z_{ij}.$$

The left inequalities ensure that $b_{ij}z_j \le w_{ij}$, while the right inequalities ensure that $w_{ij} \le b_{ij}z_j$. 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:

$$
\begin{aligned}
\max \quad & \frac{1}{k}\sum_{j=1}^{k}\sum_{i=1}^{t} 2^{-i} b_{ij} && && \text{(3a)}\\
\text{s.t.} \quad & \sum_{i=1}^{t} 2^{-i} w_{ij} \le 1 && \forall j \in [k] && \text{(3b)}\\
& z_j + M(b_{ij} - 1) \le w_{ij} \le M b_{ij} && \forall i \in [t],\ \forall j \in [k] && \text{(3c)}\\
& 0 \le w_{ij} \le z_j && \forall i \in [t],\ \forall j \in [k] && \text{(3d)}\\
& b_{ij} \in \{0,1\} && \forall i \in [t],\ \forall j \in [k] && \text{(3e)}\\
& (x,y,z) \text{ satisfies (2b)} - \text{(2j)}. && && \text{(3f)}
\end{aligned}
$$

5.3. 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 (2). This objective is the minimization of a (separable) concave function, which we reformulate to be compatible with 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 $z_j \le s_j^2$. 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 to denote the user-chosen precision, say $M=17$ and $t=20$. Again, introduce binary variables b_{ij} for each $i=1,2,\ldots,t$ and each $j\in[k]$. See that the expression $\sum_{i=1}^{t} 2^{-i} b_{ij}$ takes values between zero and (nearly) one, so we set the inverse Polsby-Popper score z_j to be

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

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

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

$$\min \quad \frac{1}{k}\sum_{j=1}^{k} s_j \tag{4a}$$

$$\text{s.t.} \quad 1+(M-1)\sum_{i=1}^{t} 2^{-i} b_{ij}^2 \leq s_j^2 \qquad \forall j\in[k] \tag{4b}$$

$$z_j \leq 1+(M-1)\sum_{i=1}^{t} 2^{-i} b_{ij} \qquad \forall j\in[k] \tag{4c}$$

$$b_{ij}\in\{0,1\} \qquad \forall i\in[t],\ \forall j\in[k] \tag{4d}$$

$$(x,y,z) \text{ satisfies } (2b)-(2j). \tag{4e}$$

The second-order cone constraints (4b) and linear constraints (4c) together impose that $z_j \leq s_j^2$. Note that $b_{ij}=b_{ij}^2$ as b_{ij} is binary, so the left side of inequality (4b) will equal the right of (4c).

5.4. Application to County-Level Districting

Now, we apply the MISOCP models to county-level districting instances in the USA. 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 apply them to ten states that could conceivably draw county-level plans while satisfying 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 the flow-based constraints of Shirabe (2005, 2009), cf. Oehrlein and Haunert (2017), Validi et al. (2022), which sufficed for county-level instances in prior work (Validi et al. 2022, Validi and Buchanan 2022). 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 that is developed later in Section 6.1. 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 results for the average: inverse Polsby-Popper score (PP^{-1}), Polsby-Popper score (PP), and Schwartzberg score ($\mathrm{PP}^{-1/2}$). 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		$\mathrm{PP}^{-1/2}$	
state	$\lvert C\rvert$	k	obj	time	obj	time	obj	time
AR	75	4	3.3176	108.55	[0.32,0.44]	TL	[1.74,1.85]	TL
ID	44	2	3.4966	0.68	0.3127	1.43	1.8584	1.29
IA	99	4	2.1725	162.03	[0.47,0.68]	TL	[1.37,1.46]	TL
KS	105	4	2.3903	438.23	[0.46,0.72]	TL	[1.32,1.57]	TL
ME	16	2	2.8979	0.06	0.3500	0.31	1.6993	0.30
MS	82	4	3.3847	197.47	[0.34,0.46]	TL	[1.64,1.84]	TL
MT	56	2	2.9482	0.69	0.3400	2.58	1.7165	1.22
NE	93	3	2.4400	21.58	0.4128	2615.25	1.5606	196.70
NM	33	3	2.4268	0.45	0.4135	10.40	1.5571	1.27
WV	55	2	5.0111	0.72	0.2360	1.71	2.2264	1.53

First, we observe that MISOCPs solve quickest with the inverse Polsby-Popper objective, with each instance solving in under eight minutes. The average Polsby-Popper and Schwartzberg objectives take longer, reaching the time limit on four instances (where $k = 4$ and $|C| \geq 75$). The average Polsby-Popper objective seems to be the most challenging, e.g., taking 44 minutes for Nebraska.

Several of these instances can be solved with more time. For example, after 11.6 hours, we obtain a maximum average Polsby-Popper score of 0.3318 for Arkansas. After 1.6 hours, we obtain a minimum average Schwartzberg score of 1.8016. The associated maps are depicted in Figure 13. We observe that the middle map, obtained from the average Polsby-Popper score, is different from the others. In particular, its eastern district includes the "Arkansas Delta" counties that border the Mississippi River. This jagged river border gives the eastern district a poor Polsby-Popper score, but permits the other districts to have shorter perimeters, resulting in a higher average score. This matches the qualitative observation from Section 3 that the average Polsby-Popper score may sacrifice the compactness of some districts to improve that of others.

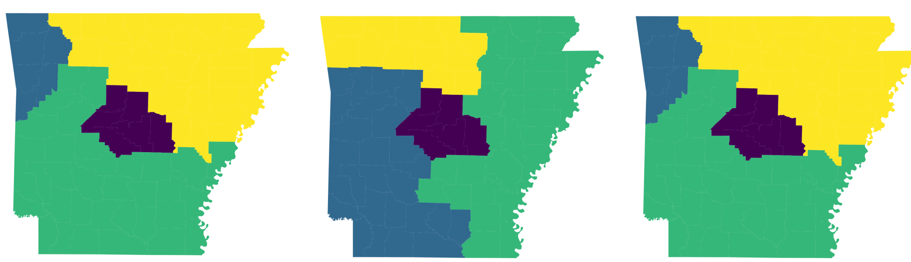

Figure 13 County-level plans for Arkansas that are optimal for the average: inverse Polsby-Popper score, Polsby-Popper score, and Schwartzberg score.

6. Handling Large Instances

To handle large districting instances, we propose a local search heuristic. 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 the 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 simply solve the associated MISOCP. We apply the approach to the notable VRA case *Merrill v. Milligan* in Alabama.

6.1. MISOCP Local Search

In graph theory, the h-hop neighborhood of a vertex $i \in V$ is the set of vertices $N^h(i) = \{v \in V \mid \text{dist}(i,v) \leq h\}$ within h hops of i. Relatedly, we can define the h-hop neighborhood of a districting plan. Below, we consider a districting plan $d : V \to [k]$ as a map from each vertex to a district.

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 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.

Below is a high-level implementation of the MISOCP local search heuristic.

1. Begin with a feasible districting plan d.
2. Using the MISOCP, find a more compact districting plan d' in the h-hop neighborhood $N^h(d)$; if none exist, then terminate.
3. Update $d \leftarrow d'$ and goto Step 2.

There are multiple ways to implement the MISOCP local search neighborhood. We simply write out the full MISOCP and update the upper bounds on the x_{ij} variables in each iteration. That is, impose an upper bound of one on x_{ij} if there is a vertex $v \in N^h(i)$ in its h-hop neighborhood that currently belongs to district j; otherwise, impose an upper bound of zero. By running breadth-first search once from each district, we can update these nk upper bounds in time $O(km + nk)$ which is $O(nk)$, as planar graphs are sparse and satisfy $m = O(n)$. This will ultimately fix most x_{ij}

variables to zero, which the solver removes in preprocessing. This simple implementation suffices for tract-level instances.

If working with larger block-level instances, then one may want to merge vertices together where possible, similar to Henzinger et al. (2020). Namely, let $I_j^h(d)$ be the interior of district j in districting plan d, i.e., the subset of vertices whose h-hop neighborhood is fully contained within district j. Then, collapse each connected component of $G[I_j^h(d)]$ into a single node, with the new weights (e.g., populations, areas, and border lengths) set appropriately. This is slightly different than the approach of Henzinger et al. (2020), as they did not impose contiguity constraints and could collapse the entirety of $I_j^h(d)$ into a single node. However, in our initial experiments, we encountered significant numerical stability issues on block-level instances. Intuitively, this is because some borders between census blocks are quite small (e.g., in cities), whereas borders in rural areas can be significantly longer, particularly along bodies of water. The situation is arguably worse for the areas a_i, which may be proportional to the *square* of the border lengths. The dramatic differences in coefficients in the perimeter constraints (2c) and area constraints (2d) make the block-level MISOCPs prone to numerical instability, leading MISOCP solvers to actively replace tiny values with zeros. Accordingly, we do not apply the MISOCPs to block-level instances and do not bother merging district interiors in our implementation.

6.2. Complexity of the Local Search Neighborhood

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 proposed 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 sound 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, but we show that the problem remains hard *even when these constraints are relaxed*, indicating that the inverse Polsby-Popper objective it 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 $a_1, a_2, \ldots, a_q$, 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 $a_1, a_2, \ldots, a_q$ 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 14.

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

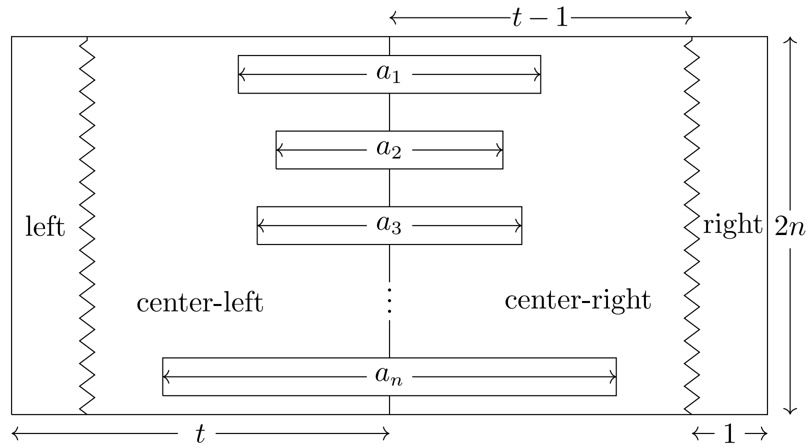

Figure 14 NP-hardness reduction from the "gift" subset sum problem.

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 $a_1, a_2, \ldots, a_{n-2}$). 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

$$\frac{1}{2}\left(\frac{P_1^2}{4\pi A_1} + \frac{P_2^2}{4\pi A_2}\right) = \frac{P^2}{8\pi}\left(\frac{1}{A_1} + \frac{1}{A_2}\right) \tag{5a}$$

$$= \frac{P^2}{8\pi}\left(\frac{1}{2nt-1} + \frac{1}{2nt+1}\right) \tag{5b}$$

$$< \frac{P^2}{8\pi}\left(\frac{2}{2nt-1}\right) = \frac{4(n+t)^2}{\pi(2nt-1)}. \tag{5c}$$

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 (5a), 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". □

6.3. Application to Alabama

Motivated by the pending Supreme Court case *Merrill v. Milligan*, we apply the local search procedure to Alabama. We begin with a set of demonstration districts drawn by *Milligan's* expert witness Moon Duchin, see Plan D in Figure 11. Recall that, for *Gingles* reasons, this plan was drawn to have two Black-majority districts. Duchin also sought to follow traditional redistricting principles, using just six county splits and attaining an average Polsby-Popper score of 0.249, already improving upon the enacted plan's score of 0.222. Due to different implementation choices, Dave's Redistricting App (DRA) reports these average Polsby-Popper scores as 0.2515 and 0.2203, respectively. As proof of concept, we apply the proposed local search procedure to make Duchin's demonstration districts even more compact.

As we have seen, the proposed MISOCPs are challenged by block-level districting plans due to their large size and numerical issues. So, we first approximate Duchin's plan with census tracts using an integer programming model. To do so, define a binary variable x_{ij} indicating whether tract $i \in V$ is assigned to district $j \in [k]$. We require the tract-based districts to satisfy contiguity (using a,b-separator inequalities in a cut callback) and a $\pm 0.5\%$ population deviation. Following Duchin, districts 2 and 7 are required to have a Black-majority among the voting age population. Similar to Abrishami et al. (2020), we use a transportation-based objective that depends on the hop-based distance $\text{dist}(D_j, b)$ from each district D_j to each block b in the block-level graph. Specifically, define the cost of assigning tract $i \in V$ to district $j \in [k]$ as the sum

$$c_{ij} = \sum_{b \in B_i} (1 + p_b)\text{dist}(D_j, b),$$

where B_i is the set of blocks in tract i and p_b is the population of block b. This cost will be zero if the tract i's blocks are all already assigned to district j. The objective is to minimize the transportation cost $\sum_{i \in V} \sum_{j=1}^{k} c_{ij} x_{ij}$. In one second of compute time, we obtain our tract-level approximation which has an average Polsby-Popper score of 0.2347 according to DRA.

In a first attempt, we apply 2-hop local search, using the average *inverse* Polsby-Popper score for best computational performance. We impose hard constraints on population balance ($\pm 0.5\%$), contiguity, and Black majorities in districts 2 and 7. In five iterations of local search (with each iteration limited to one minute of compute time), we obtain the plan on the left of Figure 15. According to DRA, this plan has an average Polsby-Popper score of 0.3636, which is a substantial improvement from the initial average score of 0.249 (or 0.2515 according to DRA). Unfortunately, the number of county splits has increased from six (in the tract-level approximation) to nineteen.

In our second attempt, we add county splitting constraints. Specifically, we disallow tract i from being assigned to district j if no other tract in its county was previously assigned to district j. This is a somewhat restrictive constraint, but preserves the number of splits within each county (and

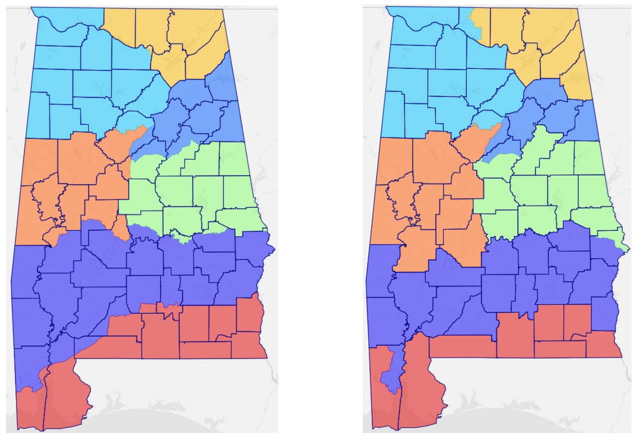

Figure 15 First and second attempts at making Plan D more compact

thus the total number of splits). In three local search iterations, we obtain the plan on the right of Figure 15. According to DRA, this plan has an average Polsby-Popper score of 0.2897, which is a 15% improvement from 0.2515.

Inspecting the two maps in Figure 15, we see that the first attempt, which was not bound by county splitting requirements, ultimately found a more compact pair of districts on the state's northern border. Further, the number of splits in them is still small. Meanwhile, the second attempt better preserved counties in the southern part of the state. To get the best of both plans, we grafted them together and applied local search with county splitting constraints.

Finally, we arrive at the plan in Figure 16 which preserves two Black-majority districts and six county splits. Moreover, this plan has an average Polsby-Popper score of 0.3068, which is nearly a 40% improvement over Alabama's enacted plan and a 22% improvement over Duchin's Plan D. If not already clear from Duchin's demonstration districts, it is compatible with traditional redistricting principles to draw two Black-majority (or Black-opportunity) districts.

Before concluding, we mention a few caveats. First, Alabama's enacted plan and Duchin's plan were both drawn to satisfy a 1-person deviation (not a $\pm 0.5\%$ deviation). They also consider municipality preservation and build districts largely from precincts (instead of tracts). Recognizing these limitations, we acknowledge that the map shown in Figure 16 would not hold up in court, but nevertheless feel that this has been a successful proof-of-concept exercise. With suitable tweaks

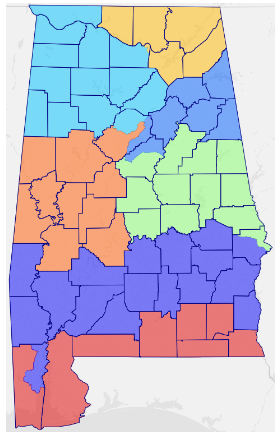

Figure 16 Demonstration districts that are 22% more compact than Plan D

(e.g., using precinct-level data and post-processing to achieve 1-person deviation), this approach could assist Section 2 litigants in satisfying the first *Gingles* precondition when seeking to overturn districting plans that dilute the voting power of minority groups or when drawing remedial districts.

7. 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 1993, 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 districts using counties or census tracts. By adding a single constraint, they can also identify optimally compact majority-minority districts. This newfound ability may find applications in Section 2 litigation when using the VRA to overturn political districting plans that dilute the voting power of minority groups. Specifically, the MISOCPs can be used by plaintiffs to identify districts that meet the first *Gingles* precondition. As proof of concept, we apply the MISOCP to draw majority-minority districts in California, Illinois, South Dakota, and Alabama.

To handle very large instances, we propose a local search neighborhood called the h-hop neighborhood, in the spirit of Henzinger et al. (2020). We show that it is NP-complete to find a more compact districting plan (with respect to the inverse Polsby-Popper score) within the h-hop neighborhood of a given plan. This problem remains hard even when the contiguity and population balance constraints are relaxed, there are only $k = 2$ districts, and the hop parameter is $h = 1$, showing that the (inverse) Polsby-Popper objective is itself hard. However, we find that the proposed MISOCPs are up for the task. Again, this ability may aid Section 2 litigants to refine an initial set of (demonstration) districts to make them substantially more compact. As proof of concept, we apply the MISOCP-based local search procedure to Alabama. Building on Duchin's plan D, we find that it is compatible with traditional redistricting principles for Alabama to draw two Black-majority (or Black-opportunity) districts.

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.

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 and Moon Duchin for sharing block assignment files for her demonstration districts in *Milligan v. Merrill*. This analysis was conducted using data from the Redistricting Data Hub.

References

Abrishami T, Guillen N, Rule P, Schutzman Z, Solomon J, Weighill T, Wu S (2020) Geometry of graph partitions via optimal transport. *SIAM Journal on Scientific Computing* 42(5):A3340–A3366.

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.

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, 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. *Journal of 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 preprint 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* (National Conference of State Legislatures).

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 (2021) Recombination: A family of Markov chains for redistricting. *Harvard Data Science Review* 3(1).

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 the 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, Spencer D (2022) Blind justice: Algorithms and neutrality in the case of redistricting. *Proceedings of the 2022 Symposium on Computer Science and Law*, 101–108.

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. *Mathematical Programming Computation* 9(2):203–229.

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 (ACDA21)*, 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`.

Levine S (2018) It's not as easy as you think to spot a Gerrymandered map. *HuffPost* URL `https://www.huffpost.com/entry/gerrymandering-strange-maps_n_5a848498e4b0ab6daf454f1e`.

Liu YY, Cho WKT, Wang S (2016) PEAR: a massively parallel evolutionary computation approach for political redistricting optimization and analysis. *Swarm and Evolutionary Computation* 30:78–92.

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. *Annals of Applied Statistics* To appear.

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 United States. *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.

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, University of Illinois at Urbana-Champaign.

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.

Polsby DD, Popper RD (1993) Ugly: An inquiry into the problem of racial Gerrymandering under the Voting Rights Act. *Mich. L. Rev.* 92:652.

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).

Redistricting Data Hub (2021) Download Data. `https://redistrictingdatahub.org/data/download-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. Available on 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 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.

Vagnozzi AM (2020) *Detecting Partisan Gerrymandering through Mathematical Analysis: A Case Study of South Carolina*. Master's thesis, Clemson University.

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 techiques. *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 (2022) Linear-size formulations for connected planar graph partitioning and political districting. Available on Optimization-Online.

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