Mathematics of Nested Districts: The Case of Alaska

Sophia Caldera*
Harvard University

Daryl DeFord
CSAIL, Massachusetts Institute of Technology

Moon Duchin
Department of Mathematics, Tufts University

Samuel C. Gutekunst
Cornell University

Cara Nix
University of Minnesota

October 15, 2019

Abstract

In eight states, each state Senate district is required to be exactly composed of two state House districts—this is sometimes known as a nesting rule. In this paper we investigate the potential impacts of these nesting rules with a focus on Alaska, where Republicans have a 2/3 majority in the Senate while a Democratic-led coalition controls the House. If we treat the current House districts as fixed and consider all possible pairings, we find that the Senate plan has a Republican advantage compared to a typical valid pairing, amounting to about one seat out of 20. The analysis enables other insights into Alaska districting, including the partisan latitude available to districters with and without strong rules about nesting and contiguity.

Keywords: Redistricting, Gerrymandering, Markov Chains, Nesting

*Author order is alphabetical. The authors acknowledge the Prof. Amar G. Bose Research Grant at MIT and the Jonathan M. Tisch College of Civic Life at Tufts University for ongoing support. MD is partially supported by NSF DMS-1255442. SG is partially supported by NSF DGE-1650441. SG thanks David P. Williamson and Kenrick Bjelland for helpful discussions. We thank Coly Elhai, Mallory Harris, Claire Kelling, Samir Khan, and Jack Snoeyink for substantial joint work and conversations on other approaches to modeling Alaska redistricting, and particularly acknowledge Samir Khan for his initial implementation of prune-and-choose. Hakeem Angulu, Ruth Buck, and Max Hully provided excellent data and technical support. Finally, we thank Anchorage Assemblyman Forrest Dunbar for bringing this problem to our attention.
# Contents

1 Introduction: Nesting 3  
   1.1 Perfect matching interpretation 3  
   1.2 Mathematical literature on perfect matchings 6  
   1.3 Paper outline 6  

2 Alaska electoral politics 7  
   2.1 Partisanship in Alaska 7  
   2.2 Racial demographics and the Voting Rights Act 10  
   2.3 Redistricting rules and practices 11  
   2.4 Our choice of election data 12  

3 Data and methods 15  
   3.1 Election results, shapefiles, and dual graphs 15  
   3.2 FKT algorithm for enumerating perfect matchings 16  
   3.3 Prune-and-choose algorithm for constructing perfect matchings 17  
   3.4 Markov chains for generating alternative House plans 20  

4 Enumerating matchings 21  

5 Alternative matchings in current House plan 23  

6 Alternative House and Senate districts 25  
   6.1 Partisan outcomes 25  
   6.2 Native population 27  
   6.3 Re-matching the new plans 27  

7 Conclusions 29  

A Prune-and-choose algorithm validity 31  

B Sampling and extremization over matchings 33
1 Introduction: Nesting

A great deal of recent attention has been given to the problem of detecting gerrymandering using mathematical and statistical tools. Most attention has been restricted to gerrymandering in its classical form: the manipulation of district boundaries to favor one party or another. However, some states’ rules of redistricting create other opportunities to extract partisan advantage from control of the process. For example, many states favor plans that keep counties and cities intact rather than splitting them between districts; Iowa even requires that congressional plans keep all of its counties intact within districts. Some observers worry about whether such seemingly neutral rules would turn out to have partisan or racial consequences for representation. (See, for instance, [12]). In this paper, we will focus on a class of redistricting principles called nesting rules, which require or encourage that members of the state-level House or Assembly be elected from districts that nest inside of the state Senate districts.

Our present case study is the state of Alaska, where 40 House districts are paired into 20 Senate districts. We start by focusing on the scenario in which House districts are fixed first, then subsequently paired into Senate districts. We select two recent elections to get a baseline of partisan preference at the precinct level, then compare the current Senate plan to all others that can be formed from the current House districts by pairing. Across all these scenarios, we will discuss when and why the choice of pairing, or perfect matching, can have a sizeable impact on electoral outcomes.

1.1 Perfect matching interpretation

There are eight states that currently have two single-member House/Assembly districts nested in each Senate district. In six of those (AL, IL, MN, MT, OR, WY), nesting is required by State Constitution or statute, and in the remaining two (IA, NV), there are provisions explaining possible exemptions. There are an additional two states (OH, WI) that require nesting of three single-member House districts within each Senate district.\footnote{The article of the Ohio constitution with this requirement was in effect in 2011 but has now been repealed, effective 2021.} Additionally, California, Hawaii, and New York call for nesting “if possible.”
Alaska 40 House → 20 Sen    Illinois 118 House → 59 Sen
Iowa 100 House → 50 Sen    Minnesota 134 House → 67 Sen
Montana 100 House → 50 Sen  Nevada 42 House → 21 Sen
Oregon 60 House → 30 Sen    Wyoming 60 House → 30 Sen
Ohio 99 House → 33 Sen     Wisconsin 99 House → 33 Sen

From the perspective of election administration, nesting is convenient because it reduces the number of different ballot styles needed. From the perspective of redistricting, nesting means that the composition of one house of the legislature massively constrains the space of possible districting plans for the other, arguably cutting down the latitude for gerrymandering.

When nesting is mandated, procedures can still vary. According to the Brennan Center’s Citizen’s Guide to Redistricting [22]: “Sometimes, a nested redistricting plan is created by drawing Senate districts first, and dividing them in half to form Assembly districts; sometimes the Assembly districts are drawn first, and clumped together to form Senate districts.” This paper will focus on the second case: matching, rather than splitting.

Proof of concept

We begin by constructing a toy example to illustrate that matchings matter. Consider the map shown in Figure 1, where each square cell represents a voter. The 56 voters are grouped into eight equally-sized, contiguous House districts, each of which is indicated by a different color. Geographically adjacent House districts (those overlapping on an edge) are to be paired to form four Senate districts. It is convenient to represent the geographic relationship of the districts with a dual graph: each node (or vertex) corresponds to a single district, and two nodes are connected by an edge if the corresponding districts are geographically adjacent. Pairing these House districts into four Senate districts corresponds to choosing a perfect matching in the graph: a set of four edges that, together, cover each of the eight nodes exactly once. (See Figure 2.)

There are exactly eight perfect matchings of this graph; in other words, given these House districts, there are only 8 ways to form four Senate districts while respecting nesting.
Figure 1: At left, an illustrative map of 56 voters in eight equally-sized House districts to be paired into four Senate districts. At right, the dual graph that encodes districts adjacency.

By contrast, there are 2,332,394,150 ways to create four (contiguous, equal-size) Senate districts from these 56 units without that restriction.\footnote{This is the number of partitions of a $7 \times 8$ grid graph into four contiguous “districts” of 14 nodes each. See mggg.org/table.html for a discussion of enumeration patterns for districting problems on grids, and a link to enumeration code.} If we name the colors White, Blue, Red, Orange, Yellow, Magenta, Cyan, and Violet, we can represent the matchings in the table below.

<table>
<thead>
<tr>
<th>Matching</th>
<th>Results</th>
<th># D</th>
</tr>
</thead>
<tbody>
<tr>
<td>WB/RC/OY/MV</td>
<td>D/R/R/R</td>
<td>1</td>
</tr>
<tr>
<td>WB/RO/YM/CV</td>
<td>D/R/R/D</td>
<td>2</td>
</tr>
<tr>
<td>WB/RV/CM/OY</td>
<td>D/R/D/R</td>
<td>2</td>
</tr>
<tr>
<td>WB/RC/OV/MY</td>
<td>D/R/D/R</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Matching</th>
<th>Results</th>
<th># D</th>
</tr>
</thead>
<tbody>
<tr>
<td>WO/BY/RC/VM</td>
<td>D/D/R/R</td>
<td>2</td>
</tr>
<tr>
<td>WO/BY/RV/CM</td>
<td>D/D/R/D</td>
<td>3</td>
</tr>
<tr>
<td>WR/BO/YM/CV</td>
<td>D/D/R/D</td>
<td>3</td>
</tr>
<tr>
<td>WR/BY/OV/CM</td>
<td>D/D/D/D</td>
<td>4</td>
</tr>
</tbody>
</table>

Figure 2: The districts can be matched the eight different ways listed here, leading to the Democratic party getting anywhere from 25\% to 100\% of the Senate seats. The two perfect matchings corresponding to the extreme outcomes are shown here.

In this toy example, we discover that the choice of matching can swing the outcome for Democrats from 1 seat to 4 seats out of four. Below, we carry out a similar analysis on real-world data.
1.2 Mathematical literature on perfect matchings

In our motivating example, we considered the Senate outcomes for every possible perfect matching in a small graph. Enumerating all perfect matchings in a given graph is a classical problem in the mathematical field of combinatorics; it has captured significant attention because it is at once quite elementary and extremely difficult to compute for arbitrary graphs [33]. The matching problem is also of great interest to physicists studying dimer coverings (domino tilings) of lattices, which are used to estimate thermodynamic behavior of liquids [20]. In 1961, three statistical physicists, Temperley, Fisher, and Kasteleyn, independently and nearly simultaneously derived the formula for the number of perfect matchings of an $m \times n$ grid [18, 31] and subsequently proposed the FKT algorithm for efficiently computing the number of perfect matchings of any planar graph (that is, in any graph that can be drawn in the plane without edges crossing). The algorithm is discussed below in §3.2. For surveys on the mathematics of matching, see [24] and Volume A of [28].

1.3 Paper outline

The central research question here is to quantify the partisan advantage available to an agent who is empowered only to select a House-to-Senate pairing. In Alaska, where there are only 40 House districts and they are patterned in a not very dense manner, it might seem that there is only limited advantage to be gained. However, we will demonstrate that the choice of pairings alone can create a swing of 4-5 seats out of 20 against recent voting patterns. In fact, we will see that even though pairings give a far simpler model of how to create Senate districts, they give just as much partisan latitude as making Senate districts from scratch.

We begin by reviewing pertinent background on Alaska politics, demographics, and redistricting rules in §2, culminating in the selection of two recent elections—the Governor and U.S. House races of 2018—to serve as our electoral baselines for the remainder of the

---

3This problem is also intimately related to a second enumeration problem, that of counting the spanning trees of a graph. The number of spanning trees is sometimes called the complexity of a graph. Temperley defined a transformation that starts with a graph and creates a new associated graph called its $T$-graph. A series of remarkable theorems tell us that if $G$ is the $T$-graph associated to $\overline{G}$, then the number of spanning trees of $\overline{G}$ is exactly equal to the number of perfect matchings of $G$ [30, 9, 19, 21].
analysis. In §3, we begin by describing the construction of dual graphs that model the adjacencies of geographical units—in this case, House districts. Next, we overview the algorithmic approaches we apply to those graphs in the rest of the paper. These methods include enumerating matchings with a classic algorithm called FKT, constructing sets of matchings with a depth-first algorithm we call prune-and-choose, and finally varying the underlying districts with a Markov chain. The proof of validity for prune-and-choose is found in Appendix A.

In the remainder of the paper, we report on the results of these algorithmic investigations. In §4 we enumerate perfect matchings in each of the eight states that mandate two-to-one nesting (see Figure 3. For several of these states, it is computationally infeasible to construct the set of matchings because it is prohibitively large; nonetheless, Appendix B describes how they can be efficiently sampled. In Alaska, we find that the interpretation of redistricting rules (in particular, geographic adjacency when regions are connected by water) has a substantial impact on the number of matchings. With the current House districts fixed, §5 measures the partisan tilt of the pairing itself among the full set of matchings. Finally, in §6, we vary the House and Senate districts themselves by randomly assembling them from precinct building blocks with a method that provides heuristic assurances of representative sampling. By exploring the space of valid plans, and evaluate expected partisan properties and matchability of alternative plans.

2 Alaska electoral politics

2.1 Partisanship in Alaska

Alaska is an outlier in U.S. political geography for several reasons. Alaska’s uniquely wide array of viable minor parties features in both local and statewide races. The state officially recognizes the secessionist Alaskan Independence Party, which succeeded in electing Wally Hickel as Governor in 1990. In addition, there are nine organized “political groups” that are seeking official recognition, and meanwhile are entitled to run candidates for statewide office: the Libertarian, Constitution, Progressive, Moderate, Green, and Veterans Parties,
Figure 3: Dual graphs for the states that require two-to-one nesting. The left-hand column shows the House districts, with the dual graph overlaid. The center column abstracts the graph, and the right-hand column shows a nearly-planar embedding with accurate district labels. Given these House districts, valid Senate plans in these states are perfect matchings of these graphs.
together with the more fringe OWL Party, Patriot’s Party, and UCES Clowns Party.\textsuperscript{4}

The current Governor of Alaska is Republican Mike Dunleavy, whose predecessor Bill Walker ran as an Independent to become the only recent U.S. Governor not from one of the two major parties. (Walker had previously left the Republican Party and then successfully ran as an Independent candidate, with a Democratic candidate for Lieutenant Governor.) In 2014, Alaska had the first U.S. Senator in more than 50 years to win election as a write-in candidate, Senator Lisa Murkowski.

![Figure 4: Trump share of major-party Presidential vote per precinct, sized by number of votes cast.](image)

Alaska is also unique in its geographic distribution of the major parties’ strengths. Unlike the contiguous United States, where urban areas tend to be most reliably Democratic, Alaska has Democratic strength in the rural areas to the north and west of the state. These areas are the homes of significant levels of Native Alaskan population, which constitutes the largest minority in the state. Conversely, the Republican vote is often stronger in suburban areas. In fact, even the city of Anchorage—by far the most populous in the state with 291,826 out of Alaska’s 710,231 residents in Census 2010—votes Republican overall in recent presidential races, making it a rare city of its size to do so.\textsuperscript{5}

Although it is one of the “reddest” states in national terms, the Republican-Democratic split is not the fundamental divide in Alaskan politics. Extremely conservative Republicans are sometimes balanced by a tenuous coalition of moderate Republicans, Democrats, and Independents, which currently aligns to give net Democratic control in the state House. In

\textsuperscript{4}See http://www.elections.alaska.gov/Core/politicalgroups.php.

\textsuperscript{5}In the 2016 Presidential race, Trump’s share of the major party vote was 58.4% statewide and 53.1% in Anchorage. The trend holds up across elections in the last cycle, with Republican performance in Anchorage trailing statewide levels by only about four points.
2018, an Independent, Libertarian, or Nonpartisan candidate ran in nine of the 40 House districts; an Independent won in one district and one Democratic candidate changed his affiliation to undeclared after winning [4]. In areas where the Democratic party label is an obstacle to election, running as an Independent can be a successful political strategy. The majority caucus in the House originally consisted of 25 members: all 15 Democrats, the two unaffiliated members, and eight Republicans [4]. On the other hand, one state Senator elected as a Democrat caucuses with the Republican majority in that body [2].

2.2 Racial demographics and the Voting Rights Act

The 2010 Census reports Alaska’s racial demographics as roughly 6% Hispanic, with non-Hispanic population comprising 63% White, 3.5% Black, 5.5% Asian, and 15% Alaska Native or other Native American as shares of the total. An additional 9% of residents are recorded as belonging to other races, or to two or more races. Figure 5 shows the proportion of Alaska Native or other Native American residents across the state.

![Figure 5: Proportion of Alaska Native or other Native American residents across Alaska. The color scale is in equal intervals of 20%; the darkest shade marks precincts that are 80-100% Native.](image)

The large Alaska Native population has long been singled out for federal protection under the Voting Rights Act of 1965, specifically Section 5 of the VRA, which required covered jurisdictions to seek prior federal approval (or “preclearance”) for any changes to districts or other voting laws. Alaska’s inclusion owed to a long history of discriminatory

---

6Twenty-one members (15 Democratic, 4 Republican, and two unaffiliated) voted together to elect the current Speaker (who ran for his House seat as a Democrat but became unaffiliated just days before being elected Speaker). Four more Republicans joined to establish the majority caucus. In May 2019, however, one Republican left the House majority coalition [4, 3].
“literacy tests”—in this case, English-language tests used to deny voting eligibility to Native residents—making Alaska one of only nine states covered in full by the special protections [32].\(^7\) Though the Supreme Court ended the practice of preclearance with Shelby v. Holder (2013), all states are still bound by the VRA requirement to afford minority groups the ability to elect a candidate of their choice where possible.\(^8\) Issues of fair representation and ballot access for the rural Native population are still highly active in Alaska \([10]\).

2.3 Redistricting rules and practices

Following a 1998 state constitutional amendment, a five-member Alaska Redistricting Board was formed to draw new district lines after each decennial census \([16]\). The House speaker, Senate president, and Chief Justice of the state Supreme Court each choose one member of the board, and the Governor chooses two. At least three members of the board must approve a redistricting plan for it to be adopted. The board must draw maps in accordance with the state Constitution, which requires that House districts be “contiguous and compact territory containing as nearly as practicable a relatively integrated socio-economic area.... [and] contain[ing] a population as near as practicable to the quotient obtained by dividing the population of the state by forty” while Senate districts are simply “...composed as near as practicable of two contiguous house districts” without further constraints \([29]\).\(^9\)

Balancing the requirements of the VRA and the guidelines of the state Constitution—compactness in particular—means that Alaska’s House and Senate districts have to be drawn in a coordinated fashion in most of the state. But Alaska’s relatively urban centers of Anchorage and Fairbanks are both predominantly white and made up of small, regular pieces, leaving room to form the House districts first.

Allegations of partisan intent have frequently been leveled at the redistricting process in Alaska. The maps drawn after the 2000 Census were accused of being a Democratic

\(^7\)Since 1971, the indigenous people of Alaska are organized into thirteen regional Tribal Corporations to administer land and finances. The current legal landscape gives the corporations substantial financial clout, which does not translate to commensurate political representation for the broader Native Alaskan population.

\(^8\)State courts have established that Alaska redistricters must consider the state’s constitutional requirements for districts before considering the requirements of the Voting Rights Act. Alaska’s courts have enforced this hierarchy several times, including most recently in a 2012 ruling that invalidated the maps used in that year’s elections \([25]\).

\(^9\)As far as we are aware, the socio-economic clause has never been operationalized or enforced.
gerrymander, while Democrats have called the post-2010 maps (drawn by a board with a 4-1 Republican majority) a Republican gerrymander [25]. The fact that a Democratic-led caucus controls the House while Republicans have 2/3 control of the Senate lends credence to the possibility that not the House districts themselves, but their pairing to form Senate districts, is chosen for Republican advantage. That possibility is investigated below.

2.4 Our choice of election data

In Alaska, three types of races occur statewide. The entire state votes for a Governor and Lieutenant Governor, elected on a single ticket, every four years; they elect one member to the U.S. House of Representatives every two years; and they elect a U.S. Senator for a term of 6 years in the Class 2 and Class 3 cycles.

We consider only those elections which occurred after the implementation of new maps in July 2013. (A map approved for temporary use in 2012 was replaced after litigation.) Seven statewide races occurred in this time period:

<table>
<thead>
<tr>
<th>Year</th>
<th>Race</th>
<th>Candidate A</th>
<th>Vote %</th>
<th>Candidate B</th>
<th>Vote %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gov14</td>
<td>Walker</td>
<td>(I/D) 48.10%</td>
<td>Parnell(R)</td>
<td>45.88%</td>
<td></td>
</tr>
<tr>
<td>Cong14</td>
<td>Young</td>
<td>(R) 50.97%</td>
<td>Dunbar (D)</td>
<td>40.97%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>McDermott (L)</td>
<td>7.61%</td>
<td></td>
</tr>
<tr>
<td>Sen14</td>
<td>Sullivan</td>
<td>(R) 47.96%</td>
<td>Begich (D)</td>
<td>45.83%</td>
<td></td>
</tr>
<tr>
<td>Cong16</td>
<td>Young</td>
<td>(R) 50.32%</td>
<td>Lindbeck (D)</td>
<td>36.02%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>McDermott (L)</td>
<td>10.31%</td>
<td></td>
</tr>
<tr>
<td>Sen16</td>
<td>Murkowski</td>
<td>(R) 44.36%</td>
<td>Miller (L)</td>
<td>29.16%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stock (I)</td>
<td>13.23%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Metcalfe (D)</td>
<td>11.62%</td>
<td></td>
</tr>
<tr>
<td>Cong18</td>
<td>Young</td>
<td>(R) 53.08%</td>
<td>Galvin (D)</td>
<td>46.50%</td>
<td></td>
</tr>
<tr>
<td>Gov18</td>
<td>Dunleavy</td>
<td>(R) 51.44%</td>
<td>Begich (D)</td>
<td>44.41%</td>
<td></td>
</tr>
</tbody>
</table>

The list includes all candidates with at least 5% of the vote in any race.10

We will use the Cong18 and Gov18 races as the fundamental electoral data for the analysis below. These two contests feature a Democratic and Republican candidate without major third-party presence and are interesting because they have very different spatial patterns of party support but similar bottom-line partisan shares. The two-party vote share for the Democrat in those races was 46.7% for Alyse Galvin against Don Young for

---

10Walker ran for Gov14 as an Independent, but with a Democratic running mate. In Gov18, Walker dropped out and ultimately received just 2% of the vote.
Figure 6: These choropleth images show that party preferences in the Governor race are spatialized very differently from the U.S. House race, even though the statewide party share is nearly identical. On the other hand, there is little visible change with and without including absentee ballots (marked with A and N, respectively), though this does have a significant bottom-line partisan impact.

Congress and 46.3% for Mark Begich against Mike Dunleavy for Governor.\textsuperscript{11}

Alaskan elections receive a high proportion of ballots not reported through individual precincts. These unprecincted ballots include absentee, provisional, and early votes, which are all reported by legislative House district. For federal races such as the U.S. House, a small number of overseas military ballots are also reported on a statewide basis. In the 2018 U.S. House race, 33.36% of reported ballots were unprecincted. In the 2018 Governor race, 34.18% of ballots were unprecincted. (For ease of reference, we will call all unprecincted ballots “absentee” below.) We report results for each election both including and excluding the absentee ballots. Thus our four election treatments can be labeled Gov18-N, Gov18-A, Cong18-N, and Cong18-A, where the N versions drop absentee ballots from the tally and

\textsuperscript{11}We note that the question of preferring endogenous or exogenous election data for redistricting analysis is a live one in political science, as reflected for instance in the article, rejoinder, and response between Best et al and McGhee in the March 2018 issue of the Election Law Journal \cite{7, 26, 6}. Our research group inclines to the use of well-chosen exogenous (statewide) election data in general, but we further note that using endogenous data would be forbidding in the Alaska legislature. Besides a significant number of uncontested races, these legislative races also feature a proliferation of minor parties (described above), making a regression analysis particularly inadequate to cleanly model voters’ preferences between the two major parties.
the A versions include them. In §6, we need to know the precinct location of the votes; for this, we assign absentee ballots to precincts in numbers proportional to precinct population.

Table 1 shows clearly that absentee ballots collectively favor Democrats by roughly two percentage points per House district. Only districts 37 and 38 have a Republican lean in the absentee ballots in either election.

Table 1: This table shows how the currently enacted House and Senate districts fall with respect to the voting patterns from our two selected elections. The individual rows correspond to the 40 numbered state House districts, paired into the 20 state Senate districts labeled A-T. The highlighting in the votes columns shows which party received the majority of the two-way vote share in the corresponding Senate district.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>R</td>
<td>D</td>
<td>R</td>
<td>D</td>
<td>R</td>
<td>D</td>
<td>R</td>
<td>D</td>
</tr>
<tr>
<td>A 1</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>B 2</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>C 3</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>D 4</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>E 5</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>F 6</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>G 7</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>H 8</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>I 9</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>J 10</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>K 11</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>L 12</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>M 13</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>N 14</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>O 15</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>P 16</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>Q 17</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>R 18</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>S 19</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>T 20</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>U 21</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>V 22</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>W 23</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>X 24</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>Y 25</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>Z 26</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
<tr>
<td>A-T</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
<td>0.5095</td>
<td>0.5095</td>
<td>0.5085</td>
<td>0.5085</td>
</tr>
</tbody>
</table>
3 Data and methods

All experiments in this paper were performed on an Ubuntu 16.04 machine with 64 GB memory and an Intel Xeon Gold 6136 CPU (3.00GHz).

3.1 Election results, shapefiles, and dual graphs

Election data were gathered from the Alaska elections website [1] and demographic data were obtained from the 2010 Census. Absentee and early voting information was only available by House district, so precinct-level data was assigned by prorating by population. We prepared the geospatial data with the MGGG Preprocessing Suite, which uses areal interpolation for blocks not fully contained in precincts [35]. The cleaned and processed version of the data is available on GitHub [27].

In the 2010 Census, Alaska had 45,292 census blocks, of which over a third (18,263) are water-only. Alaska has 441 precincts, ranging in population from a minimum of 44 people (Pedro Bay) to 7994 people (JBER1, in Anchorage) [1]. Six precincts have over 5000 people, and sixteen have under 100.

Beginning with a shapefile of the geography, in this instance precincts, we use geospatial libraries in Python to create a dual graph in GerryChain whose nodes are the geographic units, and where two units are connected by an edge if their units share a positive-length boundary in the shapefile [34]. We then adjust edges to better correspond to plausible notions of adjacency, especially when water is involved, as described below.

Water adjacency

For areas connected only by water, a decision must be made about whether to count them as adjacent. To illustrate the impact of this seemingly minor issue, we construct three different dual graphs of the precinct map, which we call the tight, restricted, and the permissive graphs.

Permissive adjacency is the closest match to the AK Division of Elections precinct shapefile. The dual graph of those precincts is nearly connected using this approach, except for one gap in the Kodiak archipelago and five additional island precincts of the
West coast. We manually added all visually reasonable connections in these cases. Among the 441 precincts, this process produces 1151 edges. Aggregating the precincts into the 40 current House districts produces a House district dual graph with 100 edges.

Figure 7: The Cook Inlet is a body of water stretching up from the Gulf of Alaska; its Knik Arm and Turnagain Arm surround the densest part of Anchorage, separating it from rural precincts to the north and south. Following precinct adjacencies provided by the state would allow districts to jump across the water, while a more restrictive notion of adjacency would not. On the right, we see that the precinct shapefile gives no guidance on how the islands are allowed to be connected to the mainland by districts.

To construct our more restricted notion of adjacency, we consulted the Census Bureau Cartographic Boundary shapefile, which is clipped to land, i.e., excludes water from its geographic units. With this as a guide, we removed certain connections across water (see Figure 7). This reduces the number of edges modestly, from 1151 to 1109 for the precinct dual graph and from 100 to 92 for the House dual graph.

Finally, we create the tightest version of the graph by using the current House map as a guide, keeping the fewest water adjacencies that would allow the current districts to be considered valid. This gives us a tight dual graph with 1105 precinct edges and 89 House edges. As we will see below, these small changes to the underlying dual graph can have large consequences for the number of possible matchings. Figure 8 shows the resulting graphs, which we use in the remainder of the analysis.

3.2 FKT algorithm for enumerating perfect matchings

As described above, Fisher, Kasteleyn, and Temperley designed a method now known as the FKT algorithm for enumerating the perfect matchings of a planar graph. This allows us to quickly determine the number of perfect matchings in a dual graph, allowing us to
evaluate whether it is computationally tractable to explicitly list all matchings. FKT takes
a planar embedding of a graph $G$ as input, then assigns signs to the edges in what is called
a Pfaffian orientation (every face should have an odd number of counterclockwise edges)
to create a signed, skew-symmetric adjacency matrix $A$. Then $\sqrt{\det A}$ counts the perfect
matchings.\footnote{The Pfaffian is a general matrix operation that agrees with $\sqrt{\det A}$ for skew-symmetric matrices. See \cite{23} for more mathematical details.}

This algorithm runs in fractions of a second on each graph, which is fast enough to
incorporate at each step of a Markov chain. Our implementation is freely available at \cite{36},
and timing details are provided in §4.

3.3 Prune-and-choose algorithm for constructing perfect matchings

In order to evaluate the partisan properties of the pairings of House districts, it is not suf-
ficient to count matchings; we also need to generate and examine the full list of matchings.
In this section, we describe a simple method to create a list of all possible perfect matchings
of a graph. This is a recursive method that simplifies the search by looking for forced pairs.

The first step is to prune the graph. This means finding all leaves of the graph (nodes of
degree one, i.e., House districts that are only connected to one other district) and matching
each with its only neighbor. We call these matches forced pairs. One round of pruning may
create new nodes of degree one in the resulting graph, and so we iteratively prune forced
pairs until there are no nodes of degree one left.

The second step is a simple check to rule out a parity obstruction to the existence of
a matching. If any connected component has an odd number of nodes, then it cannot be perfectly matched, so the whole graph also fails to have a perfect matching. If all connected components have even numbers of nodes, then we proceed.

Next is a choice step. From the remaining graph, we choose a node of minimum degree, then consider pairing it with each of its neighbors. For each of those pairings, we remove both nodes from the graph and apply our algorithm to what remains. We prune, check, choose, and iterate, in sequence, until the process terminates at a connected graph of two nodes, producing a perfect matching of the original graph.
Example

A run of our algorithm on a sample graph with ten nodes proceeds as follows.

```
A
B

D
E
C
F

G
H
I
J
```

Initiate.

A has degree one, so we record 
AB as a forced pair. Remove 
A, B from the graph. Test for 
parity. There is one remaining 
component with eight nodes, 
so we pass the parity check.

```
A
B

D
E
C
F

G
H
I
J
```

C is lowest-indexed node of de-
gree 2, so try to pair with E. 
However, one complementary 
component (FIJ) has an odd 
number of nodes, so we aban-
don this branch of the decision 
tree.

```
A
B

D
E
C
F

G
H
I
J
```

Pairing C instead with F and 
removing both from the graph 
forces the IJ pairing.

```
A
B

D
E
C
F

G
H
I
J
```

Now E is the lowest-indexed 
ode of minimal degree. By 
pairing ED, the next forced 
pairing completes the matching. Similarly for EH.

```
A
B

D
E
C
F

G
H
I
J
```

In the end, we find the two perfect matchings AB/CF/IJ/ED/GH and AB/CF/IJ/EH/GD.
3.4 Markov chains for generating alternative House plans

Markov chain Monte Carlo, or MCMC, is the leading method in scientific computing for searching large spaces and studying properties of complex systems. Numerous research groups now use MCMC implementations to study the universe of possible districting plans, once the basic units have been set [5, 11, 13]. We use the open-source software package GerryChain, created by the Voting Rights Data Institute [34].

This algorithm generates new plans iteratively, making modifications to the districting assignments of some of the units at each step. To choose how to modify districts at random, we use a proposal called Recombination (ReCom), introduced in [13].\(^\text{13}\) At each step, this proposal merges the units of two adjacent districts and then uses a balanced cut of a spanning tree to generate a new division of the merged districts. The user can elect to impose hard constraints in the form of requirements for plans, or can choose a weighted random walk that preferentially selects plans with properties deemed to better comport with the districting principles. The tree method itself promotes the selection of compact districts, so the plans generated in this way tend to have comparable compactness statistics to human-approved plans and to comfortably pass the “eyeball test” for district shape.

We ran our Markov chains on Alaska’s 441 precincts as basic units, seeking new legally valid plans for forming them into 40 House districts. As outlined above in §2.3, the law requires that the districters aim to produce equipopulous, compact, and contiguous districts.

The ideal population of a House district is 710,231/40, or between 17,755 and 17,756 people. By federal law and common practice, legislative districts can deviate by up to 5% from ideal size without a special reason, so we have imposed that limit on population balance, allowing 16,868–18,643 people per district. The same level of population balance was imposed on the Senate ensembles.

We generated ensembles of 100,000 distinct House and Senate districting plans, varying the definitions of contiguity (tight/restricted/permissive). A district-level dual graph of the sampled plan was pulled every step. Using FKT, we counted the number of matchings for

\(^{13}\)The scientific advantage of using Markov chains to sample districting plans is that they have a theoretical guarantee of producing representative samples (with respect to their stationary distributions) if run for long enough. In our case, we run the chains until we obtain strong heuristic evidence of mixing, which is a common and effective standard in scientific computing. See [13].
each districting plan and edges between House districts, and we stored plans with extreme matching statistics. The goal was to learn whether the new plans could have markedly different partisan outcomes, either on their own or when matched to form a Senate plan, from the current districts.

Replication code for producing and analyzing these ensembles is available on GitHub [36].

4 Enumerating matchings

In this section, we report the count of matchings of the nesting states’ House plans and describe computational experiments evaluating our implementations. We begin with Alaska.

<table>
<thead>
<tr>
<th></th>
<th>ALASKA</th>
<th>Tight</th>
<th>Restricted</th>
<th>Permissive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual edges</td>
<td>89</td>
<td>92</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Matchings</td>
<td>14,446</td>
<td>29,289</td>
<td>108,765</td>
<td></td>
</tr>
<tr>
<td>FKT runtime</td>
<td>0.022 sec</td>
<td>0.022 sec</td>
<td>0.027 sec</td>
<td></td>
</tr>
<tr>
<td>Prune–and–choose runtime</td>
<td>14.2 sec</td>
<td>28.5 sec</td>
<td>105.1 sec</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: On Alaska, FKT (which only counts matchings) runs in a fraction of a second; the prune-and-choose algorithm (which stores matchings) runs in under two minutes.

Table 2 already highlights the fact that toggling a small number of edges in the Alaska House dual graph (due only to reasonable interpretations of water adjacency) can change the number of perfect matchings very substantially; in this case, the number of matchings jumps by nearly a factor of eight.\textsuperscript{14}

We next apply FKT to compute the number of potential matchings for each of the eight states that require Senate districts to be formed from adjacent pairs of House districts. We use the standard Census shapefiles to generate dual graphs for each state, making them parallel to the permissive graph for Alaska. These were shown above in Figure 3.

Alaska’s 108,765 (permissive) matchings are the fewest among the eight states. This is partially due to the fact that Alaska has fewer House districts than the other states, and

\textsuperscript{14}It is worth emphasizing that the number of matchings is sensitively dependent on the precise edge structure as well as simply the number of edges. This fact is explored below in §6.3.
Table 3: Number of matchings possible with respect to the current House plan for each state (with dual graphs generated from census shapefiles) and timings for computing them with FKT.

The number of matchings varies greatly across the other states, with Minnesota having the most at $6.1 \times 10^{18}$, or over six quintillion. The sheer number of ways to pair MN House districts into Senate districts offers a reminder of the formidable size and complexity of the full redistricting problem.

Extrapolating the prune-and-choose timing from Nevada (299 seconds) and Wyoming (851 seconds) suggests that generating all of the matchings for some of the other states would take prohibitively long—even with linear scaling, the Wyoming timing suggests that the Minnesota computation would take some 180 million years. However, it is possible to sample matchings from planar graphs uniformly, allowing for good estimates of relevant statistics. We have implemented the technique suggested in [17] for this purpose [36] and in Appendix B we validate this approach on Alaska.

---

15 The four districts in the Southeast corner of the state must be paired (33–34 and 35–36), further restricting the possible matchings.

16 By contrast, the number of matchings in a $10 \times 10$ grid with 100 nodes is 258,584,046,368 (fewer than Iowa, which has 100 House districts) and a $12 \times 12$ grid with 144 nodes has 53,060,477,521,960,000 (fewer than Minnesota, which has 134 House districts). To understand why the states’ matching numbers exceed those of comparably sized grids, consider the impact of just a few extra edges. Adding just four edges to the $10 \times 10$ grid—a single diagonal edge from each of the four corner vertices to its diagonal neighbor—increases the number of matchings by 745,241,088.
5 Alternative matchings in current House plan

We consider each of the four election treatments with respect to the sets of matchings described above, comparing the number of Senate districts with a Democratic majority in the actual plan to the average number over the Senate plans formed by all possible matchings. It bears emphasizing that the Democratic seats reported across the table refers to the number of Senate districts in which Galvin votes outnumbered Young votes for Congress, or Begich votes outnumbered Dunleavy votes for Governor.

<table>
<thead>
<tr>
<th></th>
<th>D Senate districts</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Enacted</td>
<td>Tight</td>
</tr>
<tr>
<td>Cong18-N</td>
<td>6</td>
<td>6.90</td>
</tr>
<tr>
<td>Cong18-A</td>
<td>7</td>
<td>8.30</td>
</tr>
<tr>
<td>Gov18-N</td>
<td>7</td>
<td>7.69</td>
</tr>
<tr>
<td>Gov18-A</td>
<td>7</td>
<td>8.24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Competitive districts</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Enacted</td>
<td>Tight</td>
</tr>
<tr>
<td>Cong18-N</td>
<td>13</td>
<td>12.12</td>
</tr>
<tr>
<td>Cong18-A</td>
<td>12</td>
<td>12.33</td>
</tr>
<tr>
<td>Gov18-N</td>
<td>11</td>
<td>10.54</td>
</tr>
<tr>
<td>Gov18-A</td>
<td>10</td>
<td>9.78</td>
</tr>
</tbody>
</table>

Table 4: Partisan outcome and competitiveness in the current (Enacted) Senate plan compared to the average over alternative matchings. A competitive district is defined here as one with a D share between 40 and 60 percent. Here, every partisan outcome is more favorable to Republicans than the neutral expectation. Compare Table 5, which varies the underlying House plan and shows the opposite partisan tendency.

We find that the pairings in the current enacted plan do indeed exhibit a moderate Republican tilt, falling about one seat to the Republican side of the typical matching.\(^{17}\) (At the same time, we can observe the substantial effect of discarding absentee ballots: it shifts outcomes towards Republicans by about 1 seat.) The histograms in Figure 9 add detail by showing the full distribution of Democratic seats with respect to each race.

For even more granular detail, at the level of individual districts, we can study box-and-whisker plots (Figure 10). In these images, the districts are ordered from lowest to highest

---

\(^{17}\) The actual Senate composition also has six or seven Democrats, depending on how you count: seven state Senators were elected as Democrats, but Sen. Lyman Hoffman caucuses with the Republican majority.
Figure 9: The number of Senate districts with a D majority, as the matching varies across the permissive set. The green line marks the average number of Democratic seats over all matchings and the red line shows the outcome in the current plan, showing a one-seat advantage for Republicans in the current matching and a four- to five-seat swing overall. These two histograms look nearly identical for the different elections, despite the substantial differences in how the vote was spatialized.

Democratic vote share in order to make them comparable over the ensembles. The boxes show the 25th-75th percentile range and the whiskers show every result achieved over the full set of (permissive) matchings. Similar histograms and boxplots for the remainder of the elections and dual graphs are available in our supplemental material [36].
Figure 10: Democratic vote share in current Senate districts (red dots), compared to range in comparable districts over the full set of matchings (box and whiskers). With district-by-district detail, the differences between the two elections’ voting patterns are more visible. For instance, the 13th-indexed districts in the state have a Galvin (Congressional) share and a Begich (Governor) share just under the median of the respective ensembles, while nearly 75% of the ensemble in each case had a Democratic majority in the corresponding district. Where boxes have degenerated to a single value, it is because some matchings are forced, thinning the number of possibilities.

6 Alternative House and Senate districts

6.1 Partisan outcomes

Using the Markov chain ensembles of 100,000 plans each as a neutral counterfactual for drawing districts, we first report the number of House districts out of 40 with more Democratic than Republican votes.
Table 5: The number of House districts with a D majority and the number of competitive districts, as the House plan varies. A competitive district is defined as one in which the D share is between 40 and 60 percent. Compare Table 4.

Beyond the averages, we can view the full histograms to gauge the extent to which the current plan is an outlier.

Recall that the current majority House caucus includes 24 members: 16 members who were elected as Democrats, seven members who were elected as Republicans, and one who was elected as an Independent.

The Senate ensemble gives another interesting vantage point. With respect to both vote patterns, the bulk of plans assembled by the Markov chain process have 7-10 Democratic Senate seats, and the full range observed in the ensemble is 6-11. Compare this to simply matching the current House plan, where we can fully exhaust the possibilities instead of
Figure 12: The number of districts with a D majority in the indicated election, over the ensemble of (permissive) Senate plans.

sampling. Matchings give us Democratic seat outcomes of 6-10 in either vote pattern, with a small number that achieve a 5-seat outcome against the Governor returns. This means that mere control over the matchings gives essentially just as much partisan latitude as the right to draw plans from scratch with the most permissive notion of precinct adjacency.

6.2 Native population

We also find that the number of districts with an Alaska Native population majority is typically 3-4 in our randomly produced House plans, compared to three in the current House plan. Furthermore, the ability to form districts more permissively across water makes a very noticeable difference, boosting the likelihood of forming a fourth majority-Native district by random selection.

<table>
<thead>
<tr>
<th>Number of majority-Native House districts</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permitless ensemble</td>
<td>444</td>
<td>53,596</td>
<td>45,960</td>
</tr>
<tr>
<td>Restricted ensemble</td>
<td>1,053</td>
<td>97,069</td>
<td>1,878</td>
</tr>
<tr>
<td>Tight ensemble</td>
<td>1,135</td>
<td>97,507</td>
<td>1,358</td>
</tr>
</tbody>
</table>

6.3 Re-matching the new plans

Over each ensemble of 100,000 House plans, we computed the number of dual edges and the number of perfect matchings. The (nonzero) numbers of matchings varied from 74–165,344 (tight), 42–194,588 (restricted), and 852–961,176 (permissive). To a first approximation, more edges means more matchings, but the scatterplots in Figure 13 show that there is
also substantial dependence on the specific placement of the edges.\textsuperscript{18}

Figure 13: The relationship between the number of edges (x axis) and the number of matchings (y axis). Each point is a House plan, varying over the tight (top left), restricted (lower left), and permissive (right) ensembles.

To illustrate the sensitive dependence of the number of matchings on the precise topology of the graph, we focus on three examples found in the permissive House ensembles in Figure 14. The leftmost dual graph has 103 edges and 961,176 matchings while the next one has more edges but less than half the number of matchings. The disparity in these matching counts is almost entirely due to the fact that there are two ways to pair the districts in the “panhandle” of the 103-edge graph—[(29,30), (39,40)] or [(29,39), (30,40)]—compared to the unique pairing [(17,40), (35,36)] in the corresponding region of the 105-edge graph. This accounts for a doubling in the number of overall matchings, assuming a comparable number of ways to match the remaining 36 vertices. Forced pairings play an important role

\textsuperscript{18}One notable feature of Fig 13 is the prevalence of plans with low numbers of matchings. It is easy to construct graphs with any number of edges and zero matchings, simply by having any two leaves (vertices of degree one) connected to a single common neighbor—and this can easily occur by chance. There were no matchings at all in 2319, 4274, and 3504 of the dual graphs found by the ensembles, respectively. It is similarly easy to randomly construct graphs with very few matchings simply by having many leaves and thus many forced edges. On the other hand, there are graphs with 2n vertices, roughly $n^2/2$ edges, and only a single matching: start with a complete graph on n vertices (i.e., with all edges present) and add a single leaf connected to each of those. Each leaf vertex is forced to match to its unique neighbor, leaving no more vertices to pair.
in plans with few matchings. The rightmost graph shows an example with 101 edges but only 852 matchings, the lowest nonzero number ever observed. This is due to the many forced pairs—[(11,13), (12,34), (2,3), (1,36), (4,33), (8,24), (7,26), (17,25)]—which limits the flexibility in pairing the remaining vertices.

Figure 14: A selection of three House plans from our ensembles whose dual graphs have various extremal properties. The violin plot shows the number of Democratic districts with respect to Gov18-A vote data, and the colored regions show the relative sizes of the matching sets.

The analysis demonstrates that the matchability of the underlying House plan can have a significant downstream partisan impact on the Senate plans that can be formed. Figure 14 shows examples of this behavior by comparing the distributions over the possible perfect matchings for three plans from the permissive ensemble. For each of the three House plans, a typical perfect matching has 5-8 Democratic districts out of 20. However, by choosing a House plan with more district adjacencies or more matchings, it is possible to get as few as 3 or as many as 12.

7 Conclusions

Numerous studies have sought to quantify the partisan advantage secured by the selection of a particular districting plan. In that vein, we find that the current Alaska House plan
favors Democrats by an estimated 1-2 seats out of 40 when compared to other (contiguous, compact, population-balanced) ways of forming districts from precincts. The core of the paper, however, is a novel application of rigorous mathematics to redistricting in the case of a nesting rule for state legislatures. For that, we can apply the theory of perfect matchings of graphs, learning that the Alaska Senate plan secures a Republican advantage of 1 seat out of 20 when compared to other ways to match the House districts. Other findings:

- The choice of matching of a fixed House plan gives as much latitude to control partisan outcomes as drawing a Senate plan from scratch: approximately a five-seat swing out of 20.

- The significant number of absentee/early/provisional ballots in Alaska skew markedly Democratic. Different choices of how to assign them to precincts will impact findings about the consequences of moving district boundaries, and should be further studied. However, this has no effect on our analysis of matchings.

- Well-chosen statewide races, in this case the 2018 Governor and Congressional elections, gave partisan measurements that are closely compatible with each other and qualitatively concordant with the Legislative outcomes.

- Contiguity rules are not completely straightforward, and can have a major role in shaping the space of districting possibilities. For instance, permissive water adjacency makes nearly half of neutrally generated House plans have a fourth majority-Native district, while less than 2% of plans do with more restricted adjacency (§6.2).
A Prune-and-choose algorithm validity

In this section we formally describe the prune-and-choose method and provide a proof of correctness. Pseudo-code for the algorithm is given here and our implementation in Python is available at [36]. We introduce some additional notation to describe the method. The subgraph of $G$ induced by deleting nodes $u$ and $v$ will be denoted $G \setminus \{u, v\}$. We will represent a matching as a set of edges $M = \{(u_1, v_1), (u_2, v_2), \ldots, (u_\ell, v_\ell)\}$. We assume that the vertices of $G$ are ordered in order to provide a deterministic algorithm. To generate the full set of matchings for a graph $G$, we would call \texttt{FindMatchings}(G, \emptyset).

\begin{algorithm}
\caption{Pseudo-code for Prune-and-Choose Algorithm to Find All Perfect Matchings in a Graph $G$}
\begin{algorithmic}[1]
\Procedure{FindMatchings}{$G, M$} \Comment Input a graph $G$ and the current set of matched edges $M$
\If{$G$ is connected and has exactly two vertices $u, v$}
\State \Return $G \setminus \{u, v\}, M \cup (u, v)$
\ElsIf{$G$ has any vertex with exactly one neighbor}
\State prune: let $u$ be the first degree-one vertex; let $v$ be its neighbor
\State \Return \texttt{FindMatchings}(\(G \setminus \{u, v\}, M \cup (u, v)) $\triangleright$ Pair forced vertices and recurse.
\ElsIf{$G$ contains a component with an odd number of vertices}
\State \Break $\triangleright$ There are no perfect matchings in $G$
\Else
\For{$1 \leq i \leq k$, let $G_i = G \setminus \{u, v_i\}$ and $M_i = M \cup (u, v_i)$}
\State \Return $\bigcup_{i=1}^{k} \texttt{FindMatchings}(G_i, M_i)$ $\triangleright$ Recurse to find all perfect matchings with each pair
\EndFor
\EndIf
\EndProcedure
\end{algorithmic}
\end{algorithm}

We next show that the algorithm returns the correct set of perfect matchings on any graph. They key idea of the algorithm and the proof is that for any edge of the graph, the set of perfect matchings that contain edge $(u, v)$ can be computed by finding all perfect matchings in the subgraph $G \setminus \{u, v\}$. This is an example of the self-reducible nature of the perfect matching problem which is discussed in more detail below.

Theorem 1. The prune-and-choose algorithm correctly finds all perfect matchings in the input graph.

Proof. We consider any graph $G$ with $n = 2k$ vertices and proceed by induction on $k$.  

31
When \( k = 1 \), \( G \) is either connected (in which case the algorithm correctly finds the unique perfect matching at lines 2-3) or has two isolated vertices and no perfect matchings (which the algorithm correctly reports in lines 7-8).

For \( k > 1 \) the algorithm proceeds according to exactly one of the following three cases:

1. If \( G \) contains a leaf \( u \) with neighbor \( v \), then \( u \) must be matched to \( v \) in any perfect matching. Line 6 then calls FindMatchings on \( G \setminus \{u, v\} \) which returns the correct set of matchings of \( G \setminus \{u, v\} \), by our inductive hypothesis. Adding \((u, v)\) to each matching returned by this function gives the full set of matchings for \( G \).

2. If \( G \) contains no leaves and some connected component of \( G \) has an odd number of vertices, then there are no perfect matchings in \( G \) and the algorithm correctly terminates at lines 7-8.

3. If \( G \) contains no leaves and each connected component of \( G \) has an even number of vertices, then there exists a vertex of minimal index \( u \) which has a minimum number of neighbors. Since \( G \) has no leaves and no odd components, \( u \) has degree at least 2. In any perfect matching, it must be matched to one of its neighbors \( v_1, \ldots, v_k \). The algorithm considers each possibility calling FindMatchings on \( G \setminus \{u, v_i\} \) which returns the correct set of matchings by our inductive hypothesis. As in step 1, adding \((u, v_i)\) to the matchings returned on \( G \setminus \{u, v_i\} \) provides a complete set of matchings for \( G \).

Thus for any graph \( G \), the first pass through the algorithm either returns \( \emptyset \), which only occurs if \( G \) has no perfect matchings, or it calls the algorithm recursively on a graph of size \( 2(k - 1) \). These recursive calls satisfy our inductive hypothesis and hence we obtain the complete set of matchings for \( G \).

We note that well-known classes of planar graphs have exponentially many perfect matchings. For example, this is true of the \( n \times n \) grids \([18, 31, 8]\). This trivially implies that there is no polynomial-time algorithm to list them all as output. As we discuss in §4, the dual graphs of real-world districting plans often have more perfect matchings than a grid graph of comparable size. In that section we also provide timing results for our
algorithm that demonstrate that it is adequately fast for several problems at realistic scale, but not all.

Next, we discuss how our techniques can be adapted using sampling to those settings in which listing all perfect matchings is computationally infeasible.

B Sampling and extremization over matchings

For Minnesota’s 6.1 quintillion matchings, it would be prohibitively inefficient to list them all, no matter the algorithmic design. On the other hand, we can construct uniform samples of the full set of matchings by making use of the self-reducible structure in the perfect matching problem \cite{17} as follows. We can compute the likelihood that a given edge appears in a perfect matching by deleting the edge from the graph and enumerating the matchings on the remaining nodes with FKT. The ratio of matchings on the leftover to total matchings is the probability that the edge is used. With this, we can iterate, starting with the original graph and adding a single edge to the matching at each step with appropriate probability. Since FKT runs in polynomial time, so does our sampling procedure, since a perfect matching requires \(n^2\) edges and finding the probabilities used to select each edge requires at most \(\binom{n}{2}\) FKT evaluations.

We next demonstrate that the uniform sampling method can attain good accuracy with a surprisingly small number of samples, using the case of Alaska where we can compare to the ground truth from the full matching set. For each of our three dual graphs, we sample 100 matchings uniformly and compare the resulting statistics to those of the full set of matchings. Figure 15 shows these comparisons. Although the distributions are not identical, they are quite similar and the sample means vary only by small fractions of a seat from the actual values.

This example shows that even a sample of modest size produces a good estimate of the full distribution. This provides support for our assertion that this procedure can be carried out successfully on states like Minnesota, where it would be computationally infeasible to generate all matchings. We note that all materials are available in our code repositories for others to perform this sampling for the other matching states, but there will be a non-trivial data setup cost in choosing appropriate election data and cleaning it for the analysis.
Figure 15: Comparison of number of Democratic Senate districts in a uniform sample of 100 (permissive) matchings to the full collection of matchings. The table shows the absolute error in the average seats total for this and the other two Alaska dual graphs. The histograms show more detail, and illustrate how close the averages are. The fidelity is striking considering that only 1/1000 of the space has been sampled.

Though the histograms above are remarkably accurate, the sample fails to capture the full range of seat outcomes in the Governor’s race: a small number of possible matchings result in five D seats, but that is never observed in the sample. A second algorithm may be employed to provably find the correct range of seats outcomes possible, again without fully listing the matchings. Finding perfect matchings of extremal weight, given an edge-weighted graph, is a classic problem in combinatorics, solved for instance with the Blossom algorithm developed by Edmonds in the 1960s [14, 15]. To apply that in this setting, we use any given pattern of votes to assign a weight to each edge of our dual graph: an edge \{u, v\} linking two House districts u and v is given weight 1 if there are more D than R votes in the hypothetical Senate district that combines u and v. Otherwise, assign weight 0. The weight of the perfect matching is defined as the sum of the weights of its edges. By construction, this is the number of D seats in that matching. For more background on extremal perfect matchings, see for instance Chapters 25-26 of [28].

As a final note, knowing these extremes also informs the size of a uniform sample necessary to estimate the true distribution to a desired precision. A detailed discussion of the precise number of samples needed for various estimates is presented in [37]. In particular, Theorem 5.3 shows that with failure rate \(\delta\), taking max \(\left(\frac{4}{\varepsilon^2}, \frac{4\ln(\frac{1}{\delta})}{\varepsilon^2}\right)\) samples suffices to es-
timate the probability of each individual outcome to within \( \varepsilon \) (i.e., an \( L^\infty \) bound) whereas \( \max \left( \frac{4n}{\varepsilon^2}, \frac{8 \ln(\frac{1}{\delta})}{\varepsilon^2} \right) \) samples suffice to bound the sum of the absolute differences between the individual estimates and the true probabilities (an \( L^1 \) bound).

References


