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

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.[1] Additionally, California, Hawaii, and New York call for nesting "if possible."

[1]The article of the Ohio constitution with this requirement was in effect in 2011 but has now been repealed, effective 2021.

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.

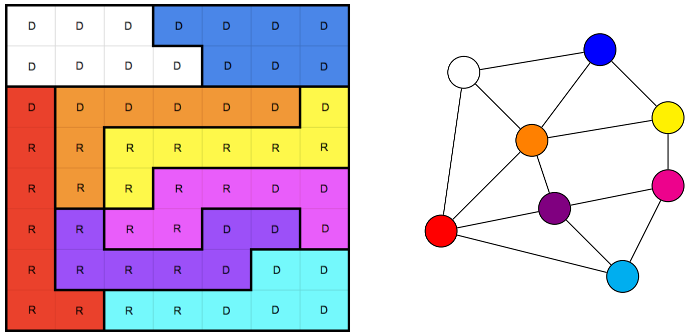


Figure 1: At left, an illustrative map of 56 voters in eight equally-sized House districts to be paired into four Senate districts. At right, the *dual graph* that encodes districts adjacency.

By contrast, there are 2,332,394,150 ways to create four (contiguous, equal-size) Senate districts from these 56 units without that restriction.[2] If we name the colors White, Blue, Red, Orange, Yellow, Magenta, Cyan, and Violet, we can represent the matchings in the table below.

Matching	**Results**	**# D**
WB/RC/OY/MV	D/R/R/R	1
WB/RO/YM/CV	D/R/R/D	2
WB/RV/CM/OY	D/R/D/R	2
WB/RC/OV/MY	D/R/D/R	2

Matching	**Results**	**# D**
WO/BY/RC/VM	D/D/R/R	2
WO/BY/RV/CM	D/D/R/D	3
WR/BO/YM/CV	D/D/R/D	3
WR/BY/OV/CM	D/D/D/D	4

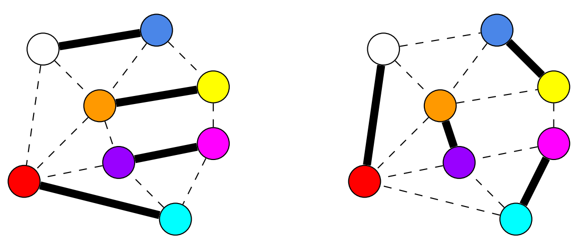


Figure 2: The districts can be matched the eight different ways listed here, leading to the Democratic party getting anywhere from 25% to 100% of the Senate seats. The two perfect matchings corresponding to the extreme outcomes are shown here.

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.

[2]This is the number of partitions of a 7×8 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.

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].[3]

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

[3] This 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,

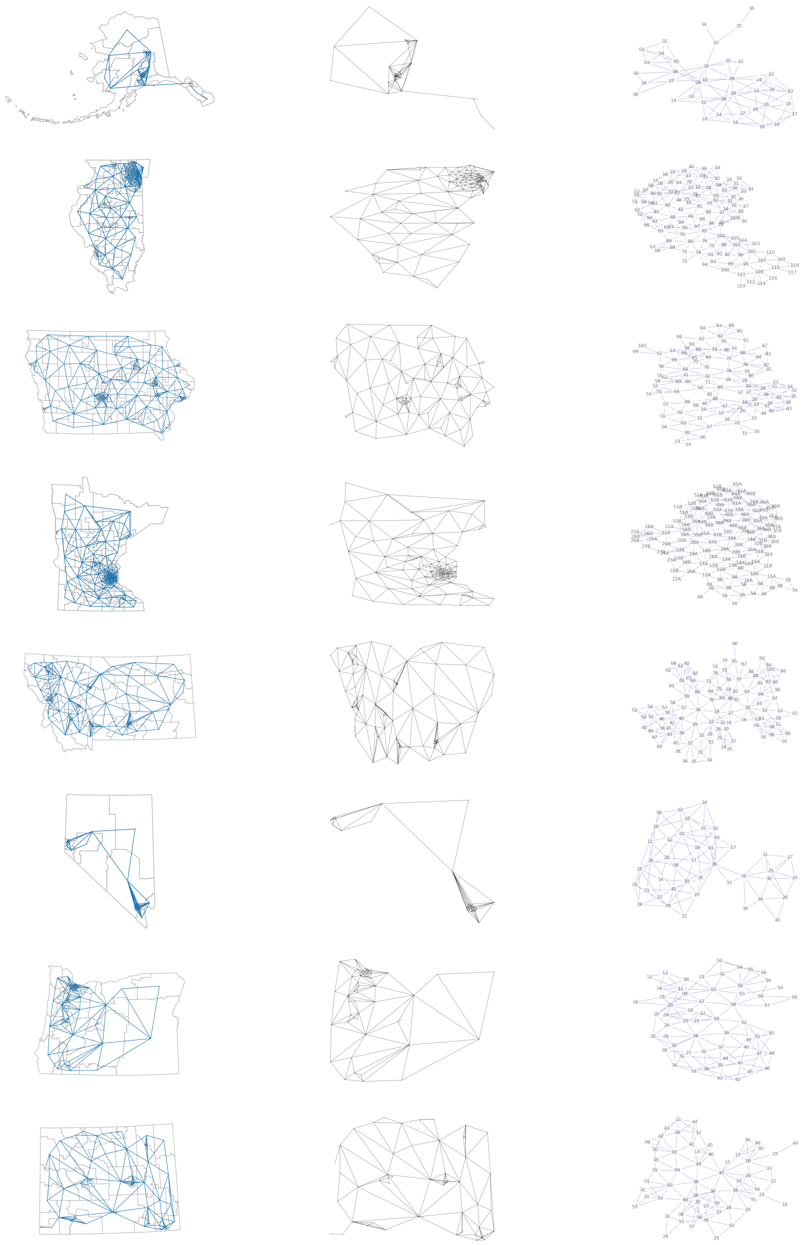


Figure 3: Dual graphs for the states that require two-to-one nesting. The left-hand column shows the House districts, with the dual graph overlaid. The center column abstracts the graph, and the right-hand column shows a nearly-planar embedding with accurate district labels. Given these House districts, 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.[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.

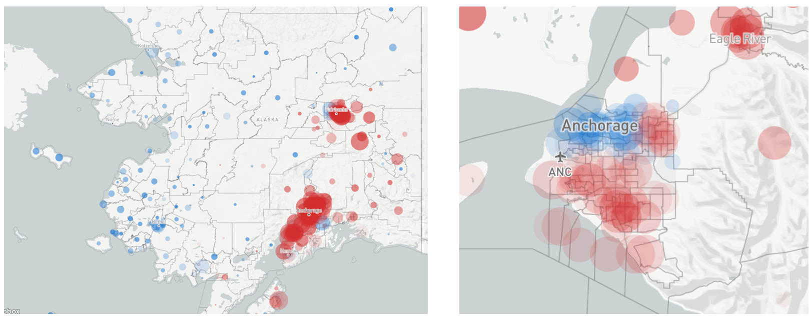


Figure 4: Trump share of major-party Presidential vote per precinct, sized by number of votes cast.

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 stronget 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.[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

[4] See http://www.elections.alaska.gov/Core/politicalgroups.php.

[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].[6] 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.

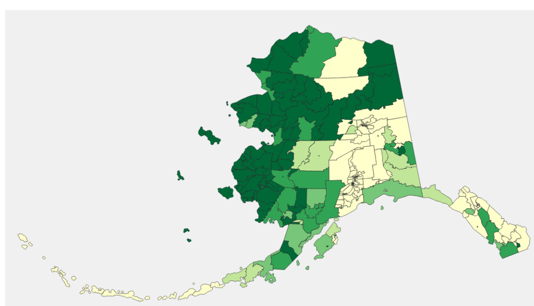


Figure 5: Proportion of Alaska Native or other Native American residents across Alaska. The color scale is in equal intervals of 20%; the darkest shade marks precincts that are 80-100% Native.

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

[6]Twenty-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.

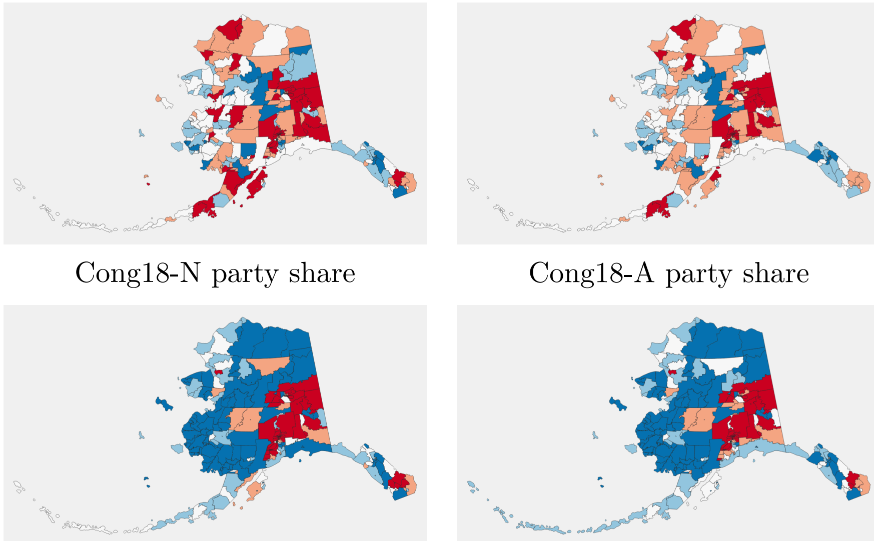
We consider only those elections which occurred after the implementation of new maps in July 2013. (A map approved for temporary use in 2012 was replaced after litigation.) Seven statewide races occurred in this time period:

Gov14	**Walker** (I/D)	[48.10%]	Parnell (R)	[45.88%]
Cong14	**Young** (R)	[50.97%]	Dunbar (D)	[40.97%]
			McDermott (L)	[7.61%]
Sen14	**Sullivan** (R)	[47.96%]	Begich (D)	[45.83%]
Cong16	**Young** (R)	[50.32%]	Lindbeck (D)	[36.02%]
			McDermott (L)	[10.31%]
Sen16	**Murkowski** (R)	[44.36%]	Miller (L)	[29.16%]
			Stock (I)	[13.23%]
			Metcalfe (D)	[11.62%]
Cong18	**Young** (R)	[53.08%]	Galvin (D)	[46.50%]
Gov18	**Dunleavy** (R)	[51.44%]	Begich (D)	[44.41%]

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

[10] Walker 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.



Cong18-N party share　　Cong18-A party share

Gov18-N party share　　Gov18-A party share

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.[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

[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 [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.

Distr		Cong18-N	Cong18-A	Gov18-N	Gov18-A	Cong18-N		Cong18-A		Gov18-N		Gov18-A	
		D%	D%	D%	D%	D	R	D	R	D	R	D	R
A	1	0.5059	0.5266	0.4974	0.5187	1853	1810	2933	2636	1751	1769	2768	2568
	2	0.3901	0.4108	0.3742	0.3916	1006	1573	1542	2211	930	1555	1406	2184
B	3	0.2543	0.2774	0.2079	0.2338	1128	3307	1644	4281	895	3410	1341	4395
	4	0.5313	0.5665	0.5264	0.5619	3067	2706	4913	3759	2942	2647	4692	3658
C	5	0.5139	0.5324	0.5099	0.5278	2386	2257	3690	3240	2296	2207	3532	3160
	6	0.3727	0.3908	0.3840	0.3968	1794	3019	2754	4293	1789	2870	2687	4085
D	7	0.2843	0.2906	0.2392	0.2435	1213	3053	2107	5142	1005	3196	1735	5391
	8	0.2461	0.2520	0.1963	0.2040	1245	3814	1857	5511	974	3989	1472	5743
E	9	0.3040	0.3242	0.2829	0.2996	1675	3834	2613	5445	1521	3855	2346	5484
	10	0.3052	0.3146	0.2545	0.2658	1601	3645	2583	5628	1310	3838	2142	5916
F	11	0.3258	0.3380	0.2869	0.3021	1830	3787	2847	5575	1574	3913	2482	5733
	12	0.2849	0.2979	0.2606	0.2739	1599	4014	2514	5923	1443	4094	2270	6019
G	13	0.3592	0.3628	0.3240	0.3332	1244	2219	1803	3166	1091	2276	1605	3212
	14	0.3873	0.3964	0.3625	0.3753	2288	3620	3413	5196	2088	3672	3142	5229
H	15	0.5109	0.5198	0.4788	0.5019	1173	1123	1858	1716	1064	1158	1729	1716
	16	0.5275	0.5484	0.5223	0.5426	2247	2013	3428	2823	2170	1985	3298	2780
I	17	0.5599	0.5890	0.5624	0.5913	2075	1631	3411	2380	2020	1572	3309	2287
	18	0.5990	0.6224	0.6011	0.6267	2191	1467	4007	2431	2131	1414	3898	2322
J	19	0.6076	0.6327	0.6126	0.6404	1773	1145	2603	1511	1741	1101	2556	1435
	20	0.6492	0.6673	0.6509	0.6745	2250	1216	4156	2072	2191	1175	4051	1955
K	21	0.5707	0.5931	0.5641	0.5830	2978	2240	4603	3157	2875	2222	4398	3146
	22	0.4682	0.4756	0.4539	0.4646	2541	2886	3655	4030	2412	2902	3469	3998
L	23	0.4999	0.5106	0.4786	0.4925	2073	2074	3036	2910	1939	2112	2841	2927
	24	0.4382	0.4545	0.4472	0.4572	2572	3297	3852	4622	2569	3176	3772	4478
M	25	0.4940	0.5165	0.4866	0.5081	2537	2599	3799	3556	2442	2577	3638	3522
	26	0.4272	0.4426	0.4266	0.4407	2350	3151	3612	4549	2301	3093	3509	4453
N	27	0.4955	0.5097	0.4978	0.5079	2732	2782	4006	3853	2688	2712	3900	3779
	28	0.4672	0.4871	0.4689	0.4824	3150	3592	5125	5397	3085	3494	4930	5290
O	29	0.3305	0.3372	0.2894	0.3031	1778	3601	2837	5575	1514	3717	2472	5684
	30	0.3153	0.3261	0.2789	0.2872	1542	3349	2537	5243	1326	3429	2169	5383
P	31	0.4367	0.4793	0.4028	0.4539	2631	3394	4649	5051	2351	3486	4251	5114
	32	0.4637	0.4840	0.5255	0.5485	1974	2283	3055	3257	2120	1914	3280	2700
Q	33	0.6810	0.7077	0.6568	0.6962	3365	1576	6665	2752	3127	1634	6301	2750
	34	0.5285	0.5694	0.4909	0.5386	2420	2159	5091	3850	2178	2259	4643	3977
R	35	0.5496	0.5608	0.5396	0.5566	2876	2357	4488	3514	2684	2290	4212	3355
	36	0.4472	0.4522	0.4065	0.4174	2329	2879	3287	3982	2033	2968	2902	4050
S	37	0.4837	0.4848	0.6618	0.6556	1711	1826	2276	2418	2205	1127	2893	1520
	38	0.5850	0.5846	0.7358	0.7313	2620	1859	3090	2195	3067	1101	3600	1323
T	39	0.5047	0.5155	0.7189	0.7193	2398	2353	2909	2734	3135	1226	3739	1459
	40	0.4666	0.4725	0.6151	0.6193	1640	1875	1960	2188	1977	1237	2359	1450

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.

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.

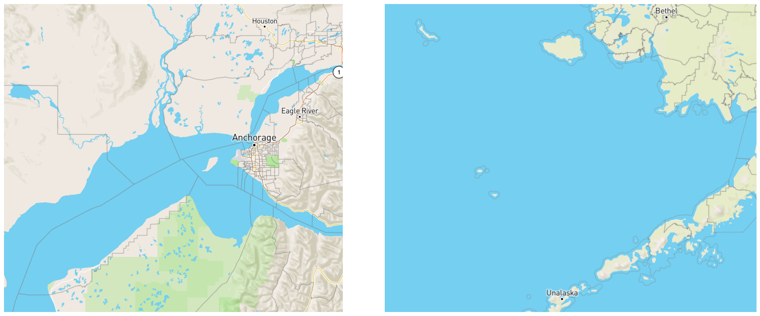


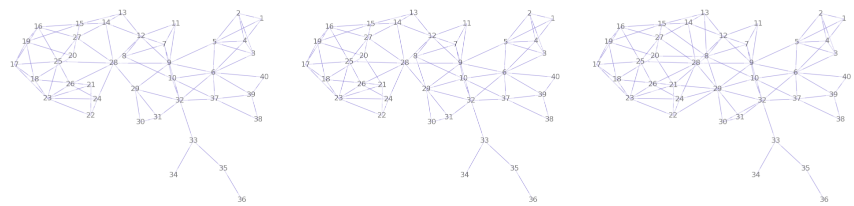
Figure 7: The Cook Inlet is a body of water stretching up from the Gulf of Alaska; its Knik Arm and Turnagain Arm surround the densest part of Anchorage, separating it from rural precincts to the north and south. Following precinct adjacencies provided by the state would allow districts to jump across 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



Tight (89 edges) | Restricted (92 edges) | Permissive (100 edges)

Figure 8: The three Alaska dual graphs.

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.[12]

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 [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 sufficient 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

[12] The Pfaffian is a general matrix operation that agrees with $\sqrt{\det A}$ for skew-symmetric matrices. See [23] for more mathematical details.

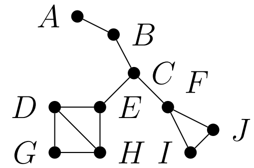
a matching. If any connected component has an odd number of nodes, then it cannot be perfectly matched, so the whole graph also fails to have a perfect matching. If all connected components have even numbers of nodes, then we proceed.

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.

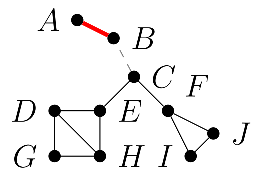
Example

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

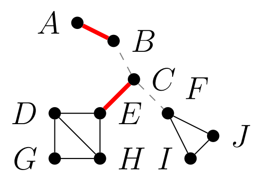
Initiate.



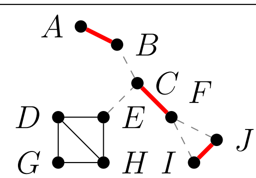
A has degree one, so we record AB as a forced pair. Remove A, B from the graph. Test for parity. There is one remaining component with eight nodes, so we pass the parity check.



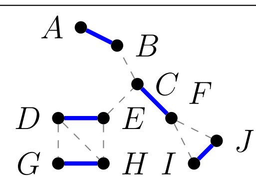
C is lowest-indexed node of degree 2, so try to pair with E. However, one complementary component (FIJ) has an odd number of nodes, so we abandon this branch of the decision tree.

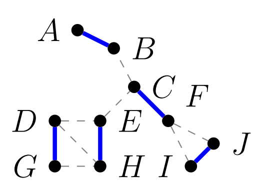


Pairing C instead with F and removing both from the graph forces the IJ pairing.



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





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].[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 disributions) 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.

ALASKA	Tight	Restricted	Permissive
Dual edges	89	92	100
Matchings	14,446	29,289	108,765
FKT runtime	0.022 sec	0.022 sec	0.027 sec
Prune–and–choose runtime	14.2 sec	28.5 sec	105.1 sec

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.[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

[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.

	Alaska	Illinois	Iowa	Minnesota
House districts	40	118	100	134
Dual edges	100	326	251	260
Matchings	108,765	9,380,573,911	1,494,354,140,511	6,156,723,718,225,577,984
FKT runtime	0.027 sec	0.39 sec	0.21 sec	0.53 sec

	Montana	Nevada	Oregon	Wyoming
House districts	100	42	60	60
Dual edges	269	111	158	143
Matchings	11,629,786,967,358	313,698	229,968,613	920,864
FKT runtime	0.24 sec	0.038 sec	0.079 sec	0.056 sec

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.

partly due to lower edge density and several forced matchings.[15] The number of matchings varies greatly across the other states, with Minnesota having the most at 6.1×10^{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.[16]

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 10×10 grid with 100 nodes is 258,584,046,368 (fewer than Iowa, which has 100 House districts) and a 12×12 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 10×10 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.

	D Senate districts			
	Enacted	Tight	Restricted	Permissive
Cong18-N	6	6.90	6.78	6.52
Cong18-A	7	8.30	8.18	7.93
Gov18-N	7	7.69	7.57	7.33
Gov18-A	7	8.24	8.13	7.88

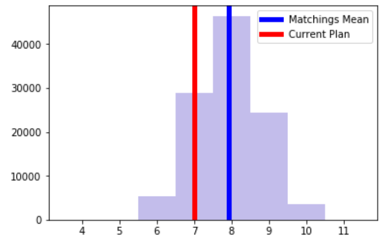
	Competitive districts			
	Enacted	Tight	Restricted	Permissive
Cong18-N	13	12.12	11.90	11.91
Cong18-A	12	12.33	12.26	12.54
Gov18-N	11	10.54	10.55	10.21
Gov18-A	10	9.78	9.68	9.68

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.

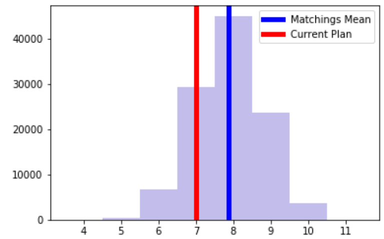
We find that the pairings in the current enacted plan do indeed exhibit a moderate Republican tilt, falling about one seat to the Republican side of the typical matching.[17] (At the same time, we can observe the substantial effect of discarding absentee ballots: it shifts outcomes towards Republicans 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.



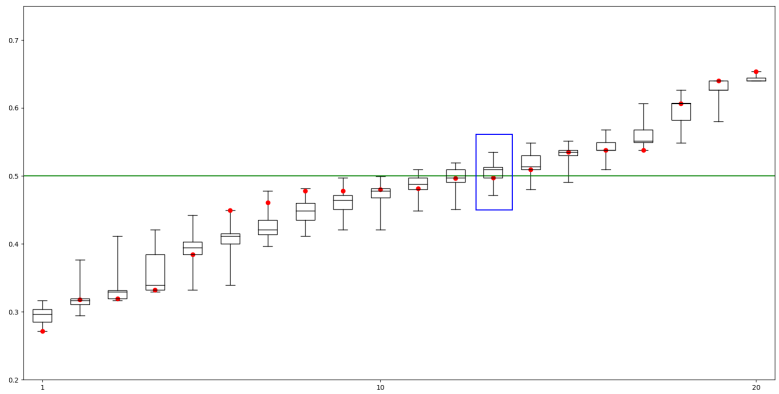
Cong18-A D Senate seats



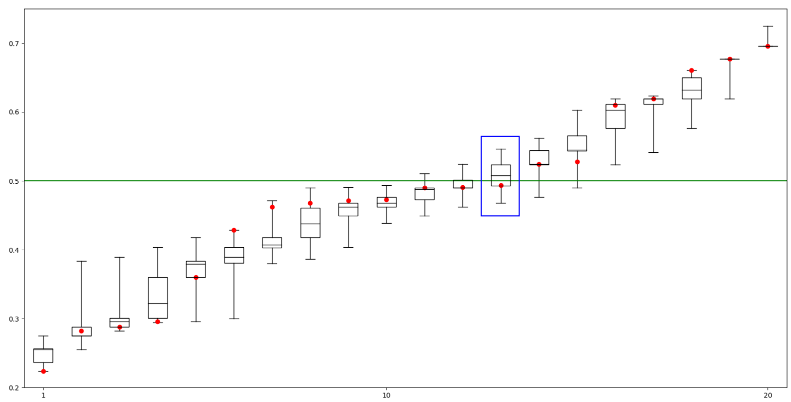
Gov18-A D Senate seats

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].



Cong18-A Dem share by Senate district



Gov18-A Dem share by Senate district

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

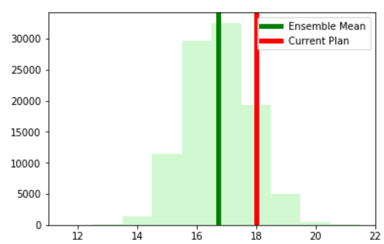
Using the Markov chain ensembles of 100,000 plans each as a neutral counterfactual for drawing districts, we first report the number of House districts out of 40 with more Democratic than Republican votes.

	D House districts			
	Enacted	Tight	Restricted	Permissive
Cong18-N	15	14.10	13.93	14.28
Cong18-A	18	16.35	16.33	16.73
Gov18-N	15	14.73	14.55	14.44
Gov18-A	20	17.86	17.63	17.65

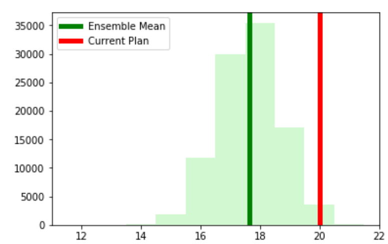
	Competitive districts			
	Enacted	Tight	Restricted	Permissive
Cong18-N	24	25.51	25.01	25.23
Cong18-A	24	24.85	24.62	24.88
Gov18-N	19	21.03	20.60	20.59
Gov18-A	19	20.83	20.57	20.52

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.



Cong18-A D House seats

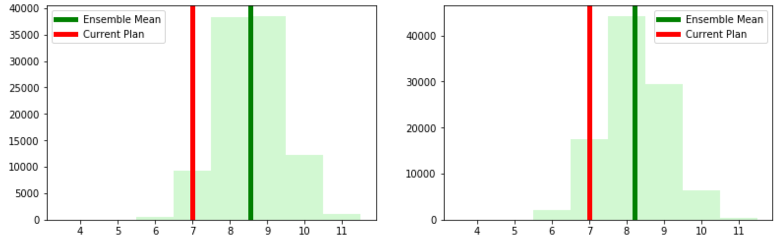


Gov18-A D House seats

Figure 11: The number of districts with a D majority in the indicated election, over the ensemble of (permissive) House plans.

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



Cong18-A D Senate seats   Gov18-A D Senate seats

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.

Number of majority-Native House districts	2	3	4
Permissive ensemble	444	53,596	45,960
Restricted ensemble	1,053	97,069	1,878
Tight ensemble	1,135	97,507	1,358

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.[18]

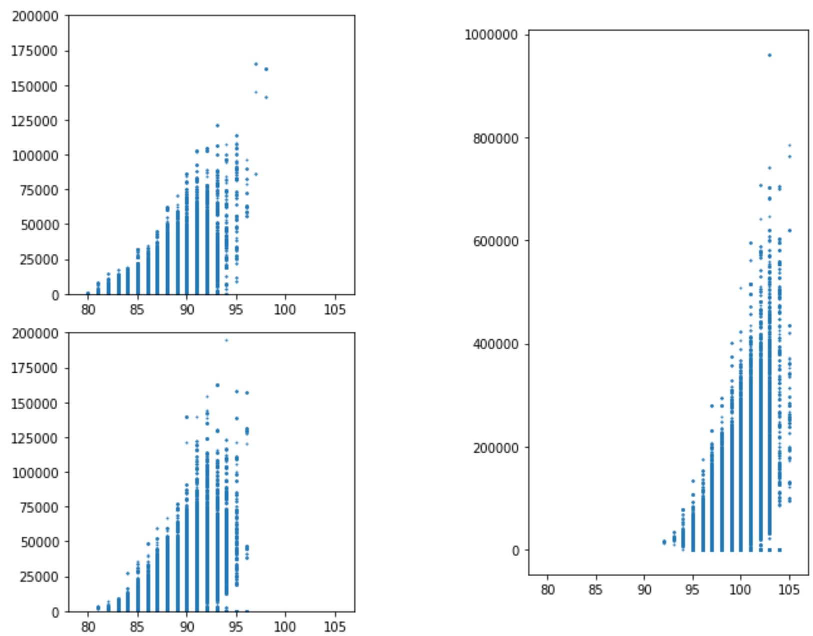


Figure 13: The relationship between the number of edges (x axis) and the number of matchings (y axis). Each point is a House plan, varying over the tight (top left), restricted (lower left), and permissive (right) ensembles.

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

[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.

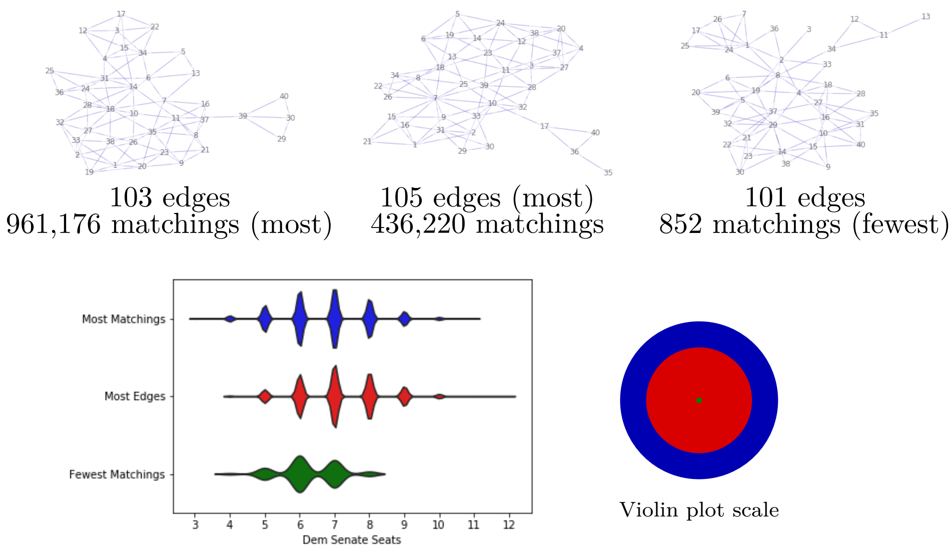


Figure 14: A selection of three House plans from our ensembles whose dual graphs have various extremal properties. The violin plot shows the number of Democratic districts with respect to Gov18-A vote data, and the colored regions show the relative sizes of the matching sets.

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 FindMatchings$(G, \emptyset)$.

Algorithm 1 Pseudo-code for Prune-and-Choose Algorithm to Find All Perfect Matchings in a Graph G

procedure FindMatchings(G, M) ▷ Input a graph G and the current set of matched edges M
  if G is connected and has exactly two vertices u, v **then**
    return $G \setminus \{u, v\}$, $M \cup (u, v)$
  else if G has any vertex with exactly one neighbor **then**
    prune: let u be the first degree-one vertex; let v be its neighbor
    return FindMatchings$(G \setminus \{u, v\}, M \cup (u, v))$ ▷ Pair forced vertices and recurse.
  else if G contains a component with an odd number of vertices **then**
    break ▷ There are no perfect matchings in G
  else
    let u be first vertex with a minimum number of neighbors $v_1, ..., v_k$
    for $1 \leq i \leq k$, let $G_i = G \setminus \{u, v_i\}$ and $M_i = M \cup (u, v_i)$
    return $\bigcup_{i=1}^{k}$FindMatchings$(G_i,\ M_i)$ ▷ Recurse to find all perfect matchings with each pair

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.

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 $v_1, ..., v_k$. 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.

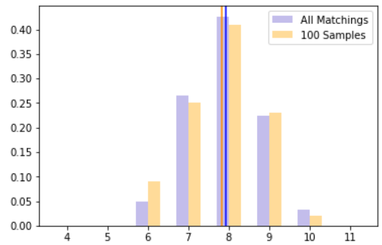
Next, we discuss how our techniques can be adapted using sampling to those settings in which listing all perfect matchings is computationally infeasible.

B Sampling and extremization over matchings

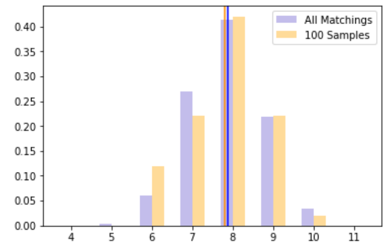
For Minnesota's 6.1 quintillion matchings, it would be prohibitively inefficient to list them all, no matter the algorithmic design. On the other hand, we can construct uniform samples of the full set of matchings by making use of the *self-reducible structure* in the perfect matching problem [17] as 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 $\frac{n}{2}$ 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.



Cong18-A D Senate seats



Gov 18-A D Senate seats

abs. error	Tight	Restricted	Permissive
Cong18-A	0.0082	0.0234	0.0868
Gov18-A	0.0203	0.0361	0.0814

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 δ, taking $\max\left(\frac{4}{\varepsilon^2}, \frac{4\ln(\frac{1}{\delta})}{\varepsilon^2}\right)$ samples suffices to estimate the probability of each individual outcome to within ε (i.e., an L^∞ 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

[1] Alaska Division of Elections. Research information.

[2] Alaska Senate Majority Profile. Senator Lyman Hoffman.

[3] Associated Press. LeDoux stripped of assignments after break with caucus, May 2019.

[4] Ballotpedia. Alaska House of Representatives elections, 2018, 2019.

[5] Sachet Bangia, Christy Vaughn Graves, Gregory Herschlag, Han Sung Kang, Justin Luo, Jonathan C. Mattingly, and Robert Ravier. Redistricting: Drawing the Line. *arXiv:1704.03360 [stat]*, April 2017. arXiv: 1704.03360.

[6] Robin E. Best, Shawn J. Donahue, Jonathan Krasno, Daniel B. Magleby, and Michael D. McDonald. Authors' Response– Values and Validations: Proper Criteria for Comparing Standards for Packing Gerrymanders. *Election Law Journal: Rules, Politics, and Policy*, 17(1):82–84, November 2017.

[7] Robin E. Best, Shawn J. Donahue, Jonathan Krasno, Daniel B. Magleby, and Michael D. McDonald. Considering the Prospects for Establishing a Packing Gerrymandering Standard. *Election Law Journal: Rules, Politics, and Policy*, 17(1):1–20, September 2017.

[8] Anders Björner and Richard P Stanley. *A combinatorial miscellany.* L'Enseignement mathématique, 2010.

[9] Robert Burton and Robin Pemantle. Local Characteristics, Entropy and Limit Theorems for Spanning Trees and Domino Tilings Via Transfer-Impedances. *The Annals of Probability*, 21(3):1329–1371, July 1993.

[10] Susanna Caldwell. Voting Rights Act: What does ruling mean for Alaskans? . *Anchorage Daily News*, 2013.

[11] Maria Chikina, Alan Frieze, and Wesley Pegden. Assessing significance in a Markov chain without mixing. *Proceedings of the National Academy of Sciences*, 114(11):2860–2864, March 2017.

[12] Daryl DeFord and Moon Duchin. Redistricting reform in Virginia: Districting criteria in context. *Preprint*, pages 1–22, 2019.

[13] Daryl DeFord, Moon Duchin, and Justin Solomon. Comparison of districting plans for the Virginia House of Delegates. *MGGG Technical Report*, pages 1–26, 2018.

[14] Jack Edmonds. Maximum matching and a polyhedron with 0, 1-vertices. *Journal of research of the National Bureau of Standards B*, 69(125-130):55–56, 1965.

[15] Jack Edmonds. Paths, trees, and flowers. *Canadian Journal of mathematics*, 17:449–467, 1965.

[16] Patti Epler. Alaska redistricting board gets to work. *Anchorage Daily News*, 2011.

[17] Mark R. Jerrum, Leslie G. Valiant, and Vijay V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43:169 – 188, 1986.

[18] P. W. Kasteleyn. The statistics of dimers on a lattice : I. The number of dimer arrangements on a quadratic lattice. *Physica*, 27:1209–1225, December 1961.

[19] Richard Kenyon. The asymptotic determinant of the discrete Laplacian. *Acta Mathematica*, 185(2):239–286, September 2000.

[20] Richard Kenyon and Andrei Okounkov. What is a dimer? *Notices of the AMS*, 52:342–343, 2005.

[21] Richard Kenyon, James Propp, and David Wilson. Trees and matchings. *The Electronic Journal of Combinatorics*, 7(1):Research paper R25, 34 p., 2000.

[22] Justin Levitt. A Citizen's Guide to Redistricting. *Brennan Center Report*, 2008.

[23] Nicholas Loehr. *Bijective Combinatorics*. CRC Press, 2011.

[24] László Lovász and Michael D Plummer. *Matching theory*, volume 367. American Mathematical Soc., 2009.

[25] Richard Mauer. Judge scolds Alaska redistricting board. *Anchorage Daily News*, 2013.

[26] Eric McGhee. Rejoinder to "Considering the Prospects for Establishing a Packing Gerrymandering Standard". *Election Law Journal: Rules, Politics, and Policy*, 17(1):73–82, October 2017.

[27] Metric Geometry and Gerrymandering Group. Mggg-States. *GitHub repository*, 2018.

[28] Alexander Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer Science & Business Media, 2003.

[29] State Constitution of Alaska. Article 6 - Legislative Apportionment.

[30] H. N. V. Temperley. Enumeration of graphs on a large periodic lattice. *Combinatorics (Proc. British Combinatorial Conf., Univ. Coll. Wales, Aberystwyth, 1973)*, pages 155–159. London Math. Soc. Lecture Note Ser., No. 13, 1974.

[31] H. N. V. Temperley and Michael E. Fisher. Dimer problem in statistical mechanics-an exact result. *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics*, 6(68):1061–1063, August 1961.

[32] U.S. Department of Justice. Preclearance, 2015.

[33] L. G. Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8(2):189–201, January 1979.

[34] Voting Rights Data Institute. GerryChain. *GitHub repository*, 2018.

[35] Voting Rights Data Institute. Maup. *GitHub repository*, 2019.

[36] Voting Rights Data Institute. Replication code. *GitHub repository*, 2019.

[37] Bo Waggoner. Lp testing and learning of discrete distributions. In *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science*, ITCS '15, pages 347–356, New York, NY, USA, 2015. ACM.