IS THIS A GERRYMANDER? EVALUATING MICHIGAN REDISTRICTING PROPOSALS

HENRY FLEISCHMANN

1. Introduction

In the United States, our State House, Senate, and U.S. House legislators represent particular districts of our states. Every ten years, these districts are redrawn to reflect changes in population from the census. The goal is to guarantee that each vote continues to count equally. Historically, governors and legislatures have been responsible for redrawing these districts. Unfortunately, far too often they take this social responsibility as an opportunity to manipulate the districts to benefit their party politically, endangering the credibility of the political process as a whole. Such politically motivated redistricting is called *gerrymandering*.

It may not be immediately obvious how one can even draw "unfair" voting districts. In the below redistricting plan, the red houses represent voters for the red party, and similarly for the blue houses. The houses are partitioned into five districts of three houses each. Observe the way that red voters are "packed" into supermajority districts and "cracked" into districts with slim blue majorities. Even though there are nine red houses and only six blue houses, the blue party wins a majority of the districts.

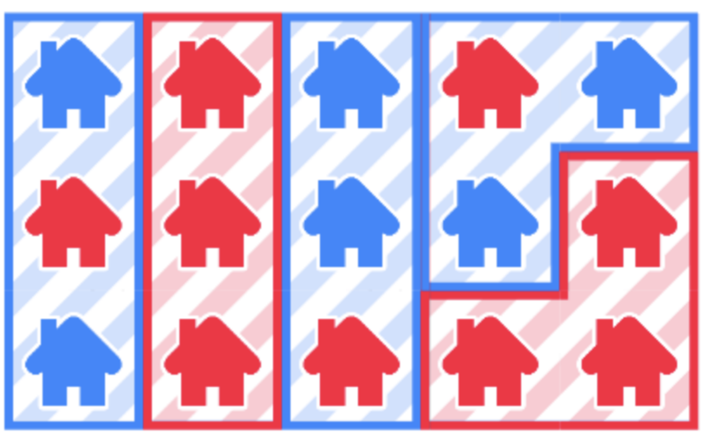


Figure 1. Partisan Gerrymander in Favor of Blue [5]

Nowadays, with election data available at the level of census blocks and the accelerating power of computation, it is more possible than ever for partisan actors to sculpt unfair districting plans.

As such, in order to foster fair elections, it is crucial to be able to detect gerrymandering. Traditionally, gerrymandered districts have been characterized as strange looking, with tendril-like components required to accommodate their unfair design. This intuition is the basis for many geometric measures of gerrymandering, such as comparing the area of districts to their perimeter, a measure called *compactness*. Related measures instead weigh the relative *convexity* of districts, the probability that straight lines between points in the district remain in the district. However, both measures are flawed and easily gameable. Measures of compactness fail to accurately describe coastline districts and other districts with fractal-like boundaries. For example, Maryland's district 1 is not necessarily clearly gerrymandered, but will always fail this compactness measure. Convexity

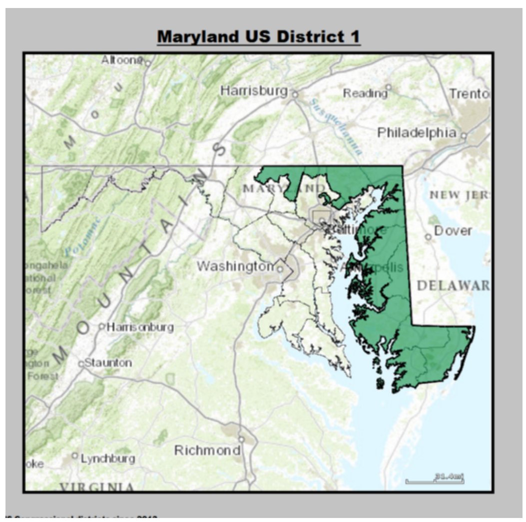


FIGURE 2. Maryland District 1, 1 Million Scale Project

measures incorrectly treat extended convex shapes, such as long rectangles, as ideal district shapes.

Crucially, modern districting plans can be gerrymandered without obvious geometric evidence of nefarious behavior. As such, it is not enough to only consider the geometry of districting plans. The 2011–2020 Wisconsin congressional districts are a clear example of this. They are not "visibly gerrymandered," whatever that truly means, and, although the vote is reasonably balanced between the Democratic and Republican parties in Wisconsin, the Republicans still hold a super-majority in the State Assembly. Clearly there is a disparity in the "efficiency" of the votes from members of each party.

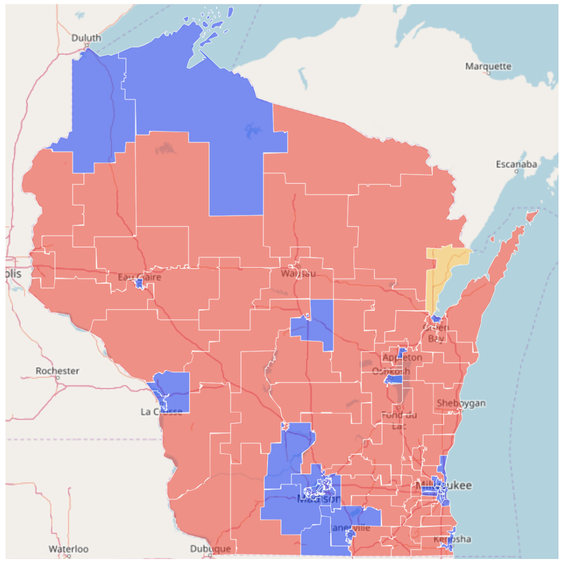


Figure 3. Wisconsin 2020 Assembly Voting Results.

This observation about vote efficiency motivates several political science measures of gerrymandering. The *efficiency gap* is the difference in the sum of "wasted" votes for each party. In this case, we consider a vote "wasted" if it is given to a losing candidate or it is extra on top of those necessary to elect a candidate. However, this measure is only sometimes effective: in a state with the minority party evenly distributed across all districts, such as Republicans in Massachusetts, it is impossible for that party to win in any plan. The efficiency gap would thus flag any districting plan as a gerrymander. Many historical measures are detailed in Moon Duchin's article on gerrymandering metrics [4].

Fortunately, recently developed tools, applications of the widely used Markov Chain Monte Carlo methods, appear to be much more reliable methods of identifying gerrymandering. The core of these methods is the assumption that, if you could simulate the same election on all possible districting plans and collect the partisan outcomes, then the current plan's partisan outcome should be reasonably close to the average outcome of all possible plans; i.e., if the current plan is fair, it should not lead to an outlier in election results. For example, if you expect Republicans to win four of thirteen seats and they instead win seven, that could be suggestive of gerrymandering. As nice as this would be, in practice we cannot simulate elections on all possible plans. There are more possible plans than the number of atoms in the universe.

The magic of the recent developments of Markov Chain Monte Carlo methods, such as Recombination [3] and Merge-Split[2] [1], is that these algorithms are designed to generate *representative samples* of all possible districting plans. By generating representative samples of all plans, these algorithms allow us to generate collections of plans which demonstrate similar results as those of all plans. The average election results on this sample of plans should be a reasonable approximation of the average election result on all possible fair plans.

A major complication of these probabilistic methods is how to simulate elections on different voting plans. If we model districts as collections of precincts and use past voting data, we have to assume that voters vote along the same party lines to gather mock election results. Although the assumption that voters vote along party lines regardless of district conformation or candidate is undesirable, we can ideally moderate the effects of candidate bias by incorporating data from multiple elections.

These positive mathematical developments prompted this project. Given that practical applications of these mathematical tools are inevitably entangled with state-specific legal requirements and data, our work was primarily focused for the nuances of the state of Michigan.

2. Recombination

Our work primarily used the Recombination algorithm, which is colloquially known as Recom. Although the mathematical basis for the algorithm is fairly involved, the fundamental ideas are quite approachable. We focus here on the steps involved in applying this algorithm, specifically in the context of Michigan.

First, we must prepare our data so that we can generate the representative ensembles of possible redistricting plans we desire. To do so, we first simplify the problem of drawing redistricting plans. Since precincts must be contained entirely within single districts, it suffices to think only at the granularity of precincts. So, the map of Michigan can be considered as a network of points (precincts) with lines connecting points when they correspond to geographically adjacent precincts. Below is an example with Iowa:

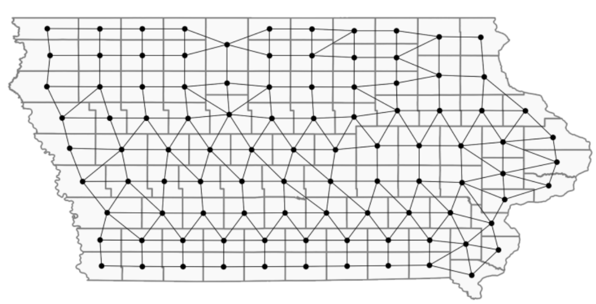


Figure 4. Network for Iowa[3]

However, in order for Recom to work, we must make a connected network—every point has to have a series of connections leading to every other point. In the case of Michigan, this means we had to manually add edges connecting the upper and lower peninsulas and the various islands in the lakes to the mainland.

Recom is an iterative process. We begin with some initial districting plan, say the existing U.S. House districts in Michigan, and, at each step, repeat the following process:

(1) Randomly select two geographically adjacent districts.
(2) Consider the network given by merging the nodes in the two adjacent districts.
(3) In this network, randomly connect the nodes (this uses an algorithm known as Wilson's algorithm).
(4) Use this connection of the nodes to partition the nodes into two connected parts such that they each have equal population (by tallying the population corresponding to each node/precinct in the merged districts).
(5) This partition of the nodes in the merged districts gives a new districting plan with these two districts transformed.

For example, we might begin with a districting plan such as the current U.S. House districts for Michigan:

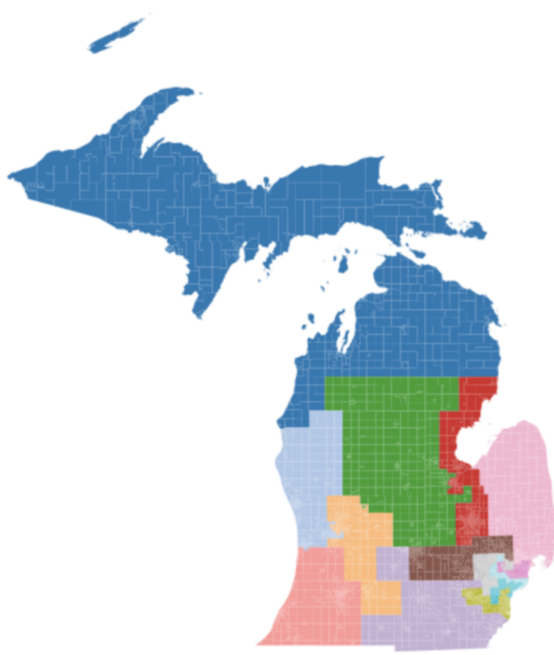


FIGURE 5. Michigan U.S. House Districts 2016

Then, after random selecting two geographically adjacent districts (colored green and orange here) we consider the precinct nodes contained in them and do the following:

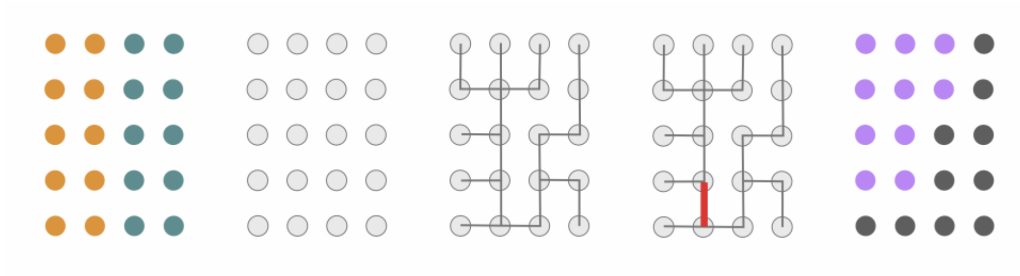


FIGURE 6. Recombination in Action [3]

The first step merges the nodes in the two districts. The second step applies Wilson's algorithm to find a random way to minimally connect all the nodes. The third step finds an edge such that removing it separates the nodes into two parts of equal population, leaving the new purple and gray districts in the last step.

After 5,000 iterations, we might yield a district which looks something like this:

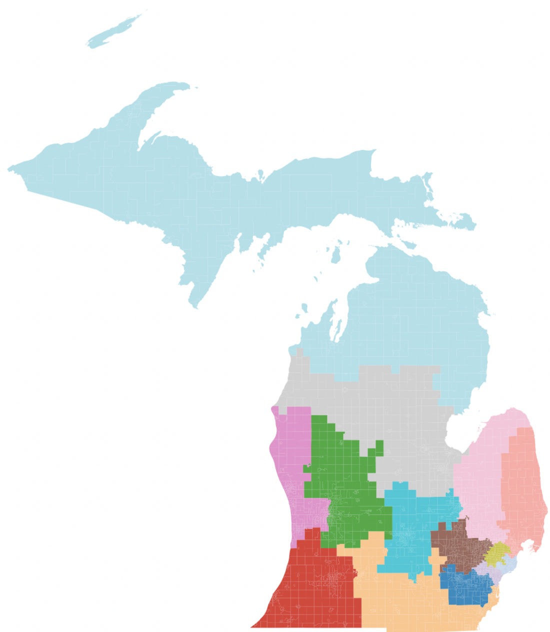


FIGURE 7. Districts After 5000 Iterations

This iterative process has many desirable properties including:

(1) There are the same number of districts in each step of the algorithm.
(2) The districts remain roughly equal population with each step.
(3) The process of merging and then splitting into new districts naturally yields reasonably compact districts.

There is lots of room for customization in generating these ensembles. Certain bounds can be placed on how much population is allowed to vary, the compactness of the districting plans, and more involved metrics. Our specific design decisions are described in more detail in the next section.

Crucially, it does not take a prohibitively large number of iterations for the samples generated by running this algorithm to provide a representative sample of all possible valid redistricting plans [3]. As a result, for a sufficiently large number of iterations, we have an ensemble ready to run mock elections on.

This is quite simple. First, we pick an election to use the voter data from. Then, we assume that each precinct will vote the exact same way they did in that election—the same number of votes for each party. By tallying the votes for each party in each precinct in each district, we have a mock election result!

For example, when we run a mock election on the voting districts in Figure 5 and Figure 7 using 2016 presidential data, the Republicans win 9/14 seats each time, even though the districting plans are visibly different.

3. Our Work

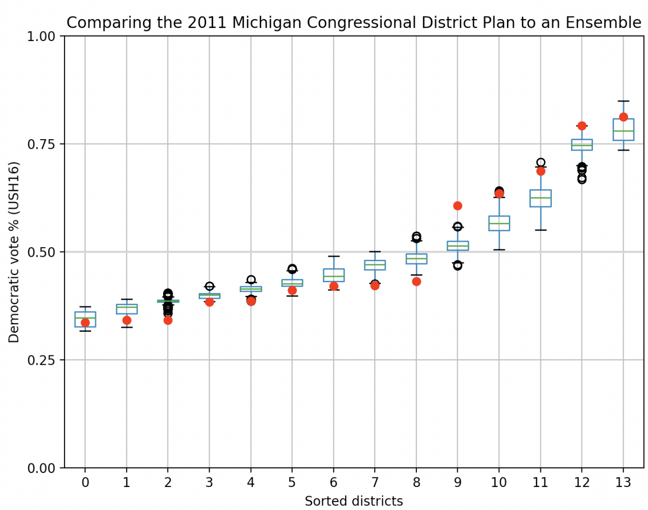
Our work centered on creating a tool customized to Michigan election law. As of the 2018 constitutional amendment, the Michigan redistricting plans should take into account the following, in order of ranked importance:

(1) Making districts of equal population
(2) Complying with the Voter Rights Act of 1965
(3) Making contiguous Districts
(4) Maintaining "communities of interest"
(5) Providing no partisan advantage
(6) Providing no candidate advantage
(7) Maintaining existing boundaries
(8) Having reasonably compact districts

Recom naturally checks off the 1st, 3rd, 5th, 6th, and 8th requirements with little to no addditional effort. The Voting Rights Act requires that redistricting plans do not discriminate on the basis of race. Oftentimes this vague requirement is interpreted as having a certain number of majority minority voter districts. In the case of Michigan, there have historically been two, so having at least two in a redistricting plan is a reasonable constraint to add. It is easy to incorporate this into the algorithm—after we generate the ensemble, we can ignore all plans that do not meet this requirement.

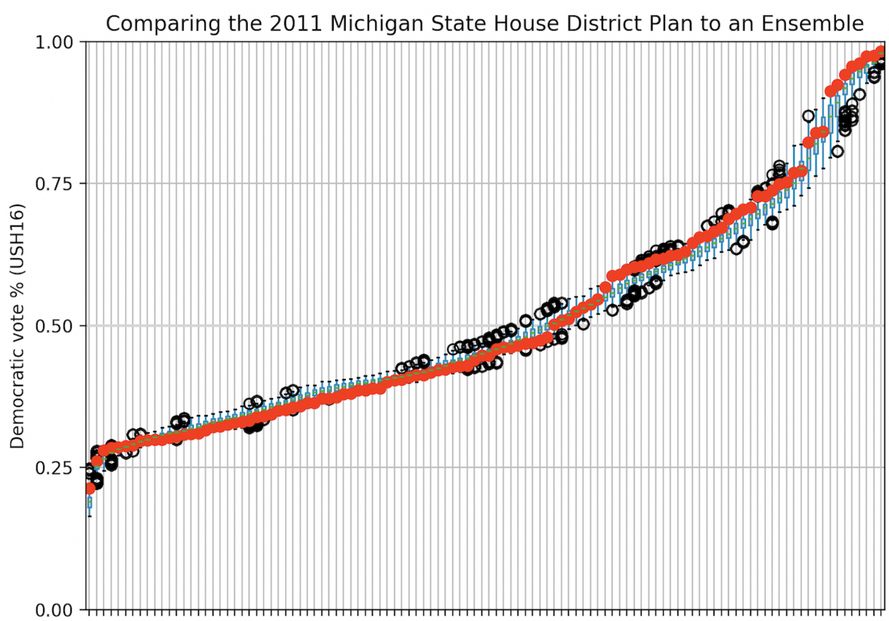
The other two conditions, maintaining "communities of interest" and maintaining existing boundaries, are very related. The former is an umbrella term describing any group of people linked by some apolitical shared interest, whether that be a church community, affiliation with a neighboring city, or a attendees of a local art museum. Such communities will be established with the Michigan Independent Redistricting Committee (established alongside the 2018 constitutional amendment) after a series of community meetings. This condition refers to keeping such communities within individual districts to facilitate their voices being heard. Similarly, condition seven refers to keeping existing boundaries such as towns or counties within the same district as much as possible. In order to deal with these two conditions, our tool allows the user to specify a list of communities/existing boundaries and preferences plans that increase the number in the same districts when generating the sample.

With these specifications in mind, we studied Michigan's current voting districts. The question you are probably wondering: "is Michigan gerrymandered?" Based on 2016 U.S. House election data, the outcomes from the current U.S. Congressional Districts are not entirely unreasonable but have some suspicious features. Running our algorithm, we found that the average number of Democratic U.S. House seats was 5.0 ± 0.65 and the current plan would result in 5 Democratic seats with this data, well within the realm of possibility. The efficiency gap was also reasonable. Nonetheless, by sorting the Democratic vote percentages in the districts and comparing that distribution with the initial distribution, we see an odd trend:



This plot shows the sorted Democratic vote percentages in the 14 districts in every plan in our ensemble compared to the sorted Democratic vote percentages in the current plan (which are marked by the red dots). This is notably not the Democratic vote percentages in the districts of the numbers in the x-axis of the plot. The box plots on the columns give the reasonable expected range for the vote percentages for each district. Crucially, every Democratically leaning district is far above the median for its Democratic vote percentage, with all but the most Democratic district being either entirely or nearly outside of the error bounds. Meanwhile, districts with Republican median voting percentages close to 50% are consistently more Republican than the median, at or past the error bounds, making the districts less affected by swing votes. The most Republican are within the error bounds. By itself this trend is not an indicator of an effective gerrymander, especially since the election results are well within the expected bounds. Nonetheless, this trend is almost certainly not accidental and is suggestive of deliberate map creation and possibly a past gerrymander.

Running our tool on the Michigan State House with 2016 Michigan State House data, we find a similar result. On average, generated plans yielded 49 ± 1.8 Democratic seats. The initial plan yielded 47 Democratic seats—on the low end of range of the expected number of seats but still within the margin of error. The efficiency gap was similarly more in favor of Republicans than average but not so far from the average as to be statistically improbable (9.9% vs. 7.3% with a standard deviation of 1.7%). We also consider the analogous boxplot for this election:



As with the U.S. House districts in the plot above, nearly every Democratic-leaning district was more Democratic than its median value. Additionally, there are jumps in the distribution of vote percentage near the 50% Democratic line which are not reflected in the medians from the sample. This could be the result of past packings. As a whole, this plot suggests that the plan could likely have been designed as a gerrymander but no longer effectively functions as one.

In addition to evaluating the existing Michigan districting plans, we also designed our tool to efficiently evaluate proposals considered by the Michigan Citizen's Independent Redistricting Commission to aid them in drawing fair voting maps for the decade to come. Their initial practice was on Ohio maps and we applied our tool to Ohio to give detailed feedback on the fairness of their maps. We hope to similarly apply it to Michigan proposals as the process proceeds.

If there is one essential idea to take away from this it is the following: gerrymandering is not simple. You cannot look at a district and know it is gerrymandered. Most mathematical attempts at quantifying gerrymandering are fundamentally flawed. However, when used carefully, mathematics can still be a useful tool in detecting unfair districting plans and promoting fair elections.

References

[1] Eric A. Autry, Daniel Carter, Gregory Herschlag, Zach Hunter, and Jonathan C. Mattingly. Multi-scale merge-split markov chain monte carlo for redistricting. August 2020.

[2] Daniel Carter, Gregory Herschlag, Zach Hunter, and Jonathan Mattingly. A Merge-Split Proposal for Reversible Monte Carlo Markov Chain Sampling of Redistricting Plans. *arXiv:1911.01503 [cs, math]*, October 2019. arXiv: 1911.01503.

[3] Daryl DeFord, Moon Duchin, and Justin Solomon. Recombination: A family of markov chains for redistricting. *Harvard Data Science Review*, 01 2021.

[4] Moon Duchin. Gerrymandering metrics: How to measure? What's the baseline? *arXiv:1801.02064 [physics]*, January 2018. arXiv: 1801.02064.

[5] GameTheory. GerryMander Game Screenshot. `http://gametheorytest.com/gerry/game/`, 2021. [Online; accessed February 2, 2021].