How to Detect Gerrymandering:
Sampling the Space of Possible North Carolina Redistrictings

Luke Farrell, Jake Shulman

DUKE COMPUTER SCIENCE

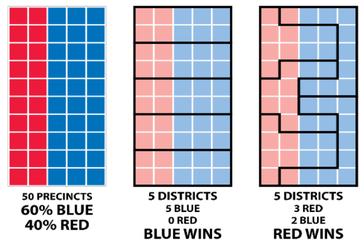
Duke UNIVERSITY | BASS CONNECTIONS

Duke UNIVERSITY MATH Research

Motivation

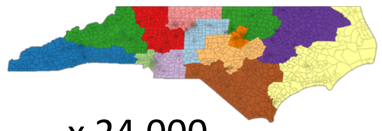
Partisan Gerrymandering is used to entrench political power by manipulating district maps. Voters of the opposing party are either "**packed**" into districts or "**cracked**" across districts to minimize the number of officials they can elect. While its results are severe, it is very difficult to rigorously detect.

Goal: Build a robust metric to detect gerrymandering.

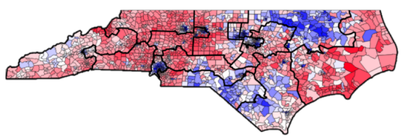


How Do We Detect A Gerrymander

1. Build an Ensemble:
Generate thousands of compliant maps without partisan data.

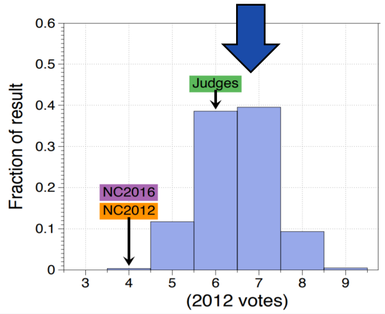


... x 24,000



2. Simulate Elections:
Feed real voting data through the ensemble of maps.

3. Check for outliers:
The distribution of election outcomes from the simulations is a non-partisan baseline. Outlier maps are suspicious.



Building the Ensemble

1. Build a score function as a heuristic for a map's legal compliance:

α_1= Contiguity
α_2= Compactness
α_3= Population Equality
α_4= County Splitting
α_5= Voting Rights Act

$$Score = \sum_i \alpha_i \omega_i$$

**The lower score = better the map*

2. Assign a probability of generating each map based on this score:

$$P(map) \propto e^{-\beta(Map\ Score)}$$

3. Since there are over 13^{300} possible maps car only take a sample with $P(map)$ as our guide.

Sampling the Space of Maps

Implement a Markov chain Monte Carlo [MCMC] process:

- Take "steps" to reach new maps in the space with probability $P(map)$, where each step represents a swap of two precincts between districts
- Take 100M+ steps to produce an ensemble of 10K+ maps

Problem: MCMC is prone to getting stuck because the space is huge, high dimensional, and rugged → Not exploring the whole space.

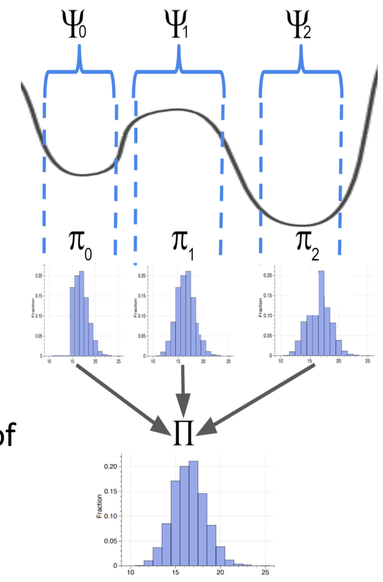
Designing Stratified Sampling

Solution: Enforce **exploration** by segmenting the space into strata and then **exploitation** by examining each individually.

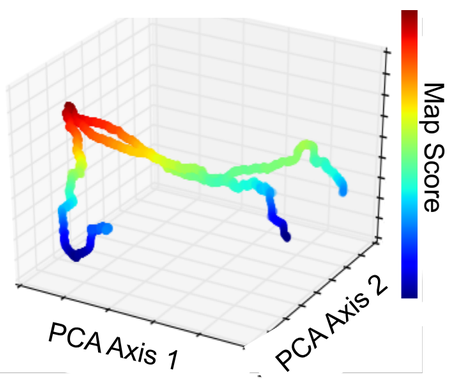
1. Learn strata:
Perform an MCMC walk at low β to define many strata.

2. Explore strata:
Within each of the strata boundaries perform MCMC at high β to find maps.

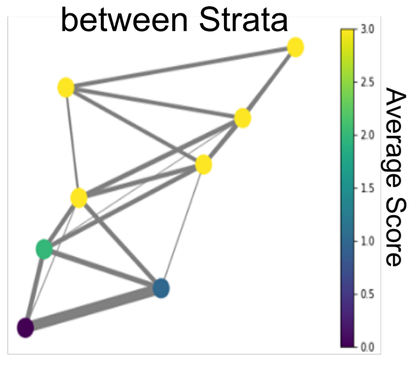
3. Merge strata:
Merge the local distributions of maps into a single global distribution using a method similar to the *Page Rank Alg*



MCMC Movement



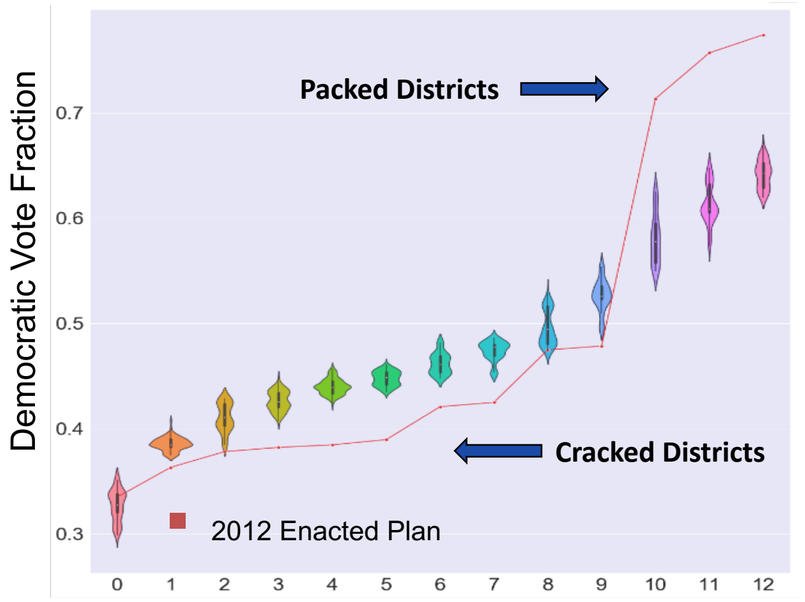
Movement between Strata



Results

*Each violin plot below shows the distribution of the democratic vote share our ensemble of maps generated for each of NC's 13 districts. This kind of evidence was presented to the **Supreme Court in Rucho v. Common Cause***

The Signature of Gerrymandering



North Carolina Voting Districts
Ordered Least to Most Democratic

Major Findings:

- Our method generated 50,000 unique redistrictings of North Carolina.
- This ensemble of maps is consistent with alternative sampling methods implemented, suggesting robustness.
- The 2012, 2016 Enacted maps are outliers and form characteristic S-shaped curve. We suggest this is ***Diagnostic of Partisan Gerrymandering in NC.***

Method Advantages:

- Highly parallelizable, allowing us to find double the number of maps of previous methods.
- Balances exploration and exploitation extremely well.

References

[1] B. Fifield, M. Higgins, K. Imai, and A. Tarr (2018) A New Automated Redistricting Simulator using Markov Chain Monte Carlo.
[2] G. Herschlag et. al (2018), Quantifying Gerrymandering in North Carolina.
[3] Warrington GS (2017) Quantifying gerrymandering using the vote distribution.
[4] Charles Matthews et. al (2016), Umbrella sampling: a powerful method to sample tails of distributions distribution
[5] Thiede E. H., et al. (1977) Journal of Chemical Physics, 145, 084115
Torrie G., Valleau J., 1977, Journal of Computational Physics, 23,187