Quantifying Gerrymandering: Revealing Geopolitical Structure through Sampling

Turing Institute
London, UK

Jonathan Mattingly, Greg Herschlag
+ The Team @ Duke Math
Impact of Duke Team’s work

**Common Cause v. Rucho** (*N.C. Congressional*):
- 3 judge conditional panel. Direct appeal to SCOTUS. Nov 2017
- Provided expert testimony and report in lawsuit
- Heavily cited in court judgment

**Gill v. Whitford** (*WI State Assembly*):
- Oral argument held in Supreme Court (SCOTUS) October, 2017
- Provide report supporting Amicus Brief by Eric S. Lander

**North Carolina v. Covington** (*N.C. State Assembly*):
- 3 judge panel rule racial gerrymander. Affirmed by SCOTUS in June
- Provide expert testimony on new maps produces at courts order
- Preparing for partisan gerrymander

[sites.duke.edu/quantifyinggerrymandering](sites.duke.edu/quantifyinggerrymandering)

As seen in …

- Undergraduate research 2014, 2015, 2016, 2018
- nature International journal of science
- The Atlantic
- The New York Times
- The Herald-Sun
packed into overwhelmingly Democratic districts at the top of the “S” or dispersed within safe Republican districts at the bottom of the “S.” App.102-103.

On January 9, 2018, the District Court issued an opinion holding unanimously that Appellants had standing to challenge the 2016 Plan on a statewide and district-by-district basis. The court unanimously found that the Plan violates the Equal Protection Clause and Art. I, §§ 2 and 4. A two-judge majority also held that the Plan violates the First Amendment. On January 18, 2018, this Court stayed the judgment pending appeal.

Dr. Mattingly, meanwhile, generated over 24,000 alternative maps using only nonpartisan districting criteria. Fewer than 0.7% of them resulted in a Republican advantage as lopsided as 10-3. App.101.

Dr. Mattingly’s simulations also confirmed that the 2016 Plan both packed and cracked Democratic voters. As he explained, this can be shown by plotting the Democratic vote share of each district on a graph, with the most Republican districts on the left and the most Democratic on the right. As the diagram below reflects, with no packing or cracking, the median map yields a straight line. The actual results for the 2016 Plan (shown in blue) are quite different. They resemble an “S” curve, with Democratic voters either...
Gerrymander

- Manipulate district boundaries to favor one party (partisan) or class (racial)

- Change the outcome of an election
  "gerrymander the results"

- Boston Gazette — 26 March, 1812

racial vs partisan gerrymander
North Carolina

13 Congressional Representatives
Gerrymander

- **Manipulate** district boundaries to favor one party (partisan) or class (racial)

- **Change** the outcome of an election
  "gerrymander the results"

- Boston Gazette — 26 March, 1812

**racial vs partisan gerrymander**
NC 2012
US Congressional Districts

Is Gerrymandering Oddly Shape Districts?
Which Doesn’t Belong?
Beyond Gerrymandering Judges

NC 2012: Same
NC 2016: Same
Different
Gerrymandering as “Startling”

Election Results

NC : US House 2012

<table>
<thead>
<tr>
<th>Vote</th>
<th>Seats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Democratic</td>
<td>50.65%</td>
</tr>
<tr>
<td>Republican</td>
<td>48.80%</td>
</tr>
</tbody>
</table>

WI : Gen Assembly 2014

<table>
<thead>
<tr>
<th>Vote</th>
<th>Seats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Democratic</td>
<td>51.28%</td>
</tr>
<tr>
<td>Republican</td>
<td>48.72%</td>
</tr>
</tbody>
</table>

MD : US House 2012

<table>
<thead>
<tr>
<th>Vote</th>
<th>Seats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Democratic</td>
<td>63%</td>
</tr>
<tr>
<td>Republican</td>
<td>33%</td>
</tr>
</tbody>
</table>

USA : US House 2012

<table>
<thead>
<tr>
<th>Vote</th>
<th>Seats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Democratic</td>
<td>50.65%</td>
</tr>
<tr>
<td>Republican</td>
<td>48.80%</td>
</tr>
</tbody>
</table>

- The most Democratic district had 79.63% Democratic votes
- The most Republican district had 63.11% Republican votes.
How to quantify how gerrymandered or unrepresentative a redistricting is?

U.S. Not a Proportional Representation System

Geographically Localize Representation
Geographically Diverse

Population Density

Presidential Election 2016

Ideal Congressional District (1/13 of population): 733,499 people
Charlotte Area: Charlotte-Gastonia-Salisbury- population 2,402,623
The Triangle: Raleigh-Durham-Cary-Chapel Hill- population 1,749,525
The Piedmont Triad: Greensboro—Winston-Salem—High Point- population 1,589,200
How to quantify gerrymandering?

maybe we should understand this as …

How to reveal a state’s natural geopolitical structure?
60% Red
40% Blue

Red wins 3
Blue wins 2

Red wins 5
Blue wins 0

Red wins 2
Blue wins 3

Wikipedia; image by Steven Nass
How to quantify how gerrymandered or unrepresentative a redistricting is?

When is a map fair?

When is a map typical?
What if we drew the districts randomly?

with no regard for party registration or most demographics

Look for the “likely” behavior of an ensemble of districting plans

create a null-hypothesis without partisan bias
Groups using algorithmic generated maps to benchmark

- Jowei Chen (Michigan), Jonathan Rodden (Stanford)
- Wendy Cho (UIUC)
- Kosuke Imai, Benjamin Fifield (Princeton)
- Alan Frieze, Wesley Pegden, Maria Chikina (CMU,Pitt)

All generating alternative maps. Some sampling a defined distribution. Some using actual surrogate districts.

Focus on our group at Duke is based on principled, explicit distribution on redistricting plans.
The Recipe

1. Determine a **compliant random** redistricting plan (equal population, compact, VRA compliant, communities of interest kept intact)

2. Count number of Democratic and Republican **votes** in each of the **new districts** using **actual votes**

3. Determine **winner** in each district of the random redistricting plan

4. Return to step 1

   **Use Markov Chain Monte Carlo to sample a distribution on redistricting plans**
Criteria for Sampling
non-partisan design criteria (HB 92)

1. districts have equal population
2. the districts are connected and compact,
3. splitting counties is minimized, and
4. African American voters are sufficiently concentrated in 2 districts to affect the winner.
N.C. HOUSE BILL 92
REDISTRICTING STANDARDS

• Districts within 0.1% of equal population
• Districts shall be reasonably compact
• Contiguous territory, attempting not to split cities or counties
• Comply with the Voting Rights Act of 1965
• Ignore: Incumbency, party affiliation, demographics
• Districts within 0.1% of equal population (we get close)

• Districts shall be reasonably compact

• Contiguous territory, attempting not to split cities or counties

• Comply with the Voting Rights Act of 1965

• Ignore: Incumbency, party affiliation, demographics
Use **Markov Chain Monte Carlo** to sample from redistricting with good scores.

Sample: $(\text{density}) \propto e^{-\beta(\text{score of redistricting})}$

Know what distribution we are sampling from.

Not just generating a large number of alternatives.
N.C. Precincts: around 3,000

\(\xi: \{\text{Precincts}\} \mapsto \{1, \ldots, 13\}\)

\[ P(\xi) = \frac{1}{\mathcal{Z}} e^{-\beta J(\xi)} \]

(a 13 color Potts Model with an unusual energy)
Score function

\[ J(\xi) = w_p J_{\text{pop}}(\xi) + w_I J_{\text{compact}}(\xi) + w_c J_{\text{county}}(\xi) + w_m J_{\text{mino}}(\xi) \]

\[ \xi: \{\text{Precincts}\} \mapsto \{1, \ldots, 13\} \quad P(\xi) = \frac{1}{\mathcal{Z}} e^{-\beta J(\xi)} \]

(a 13 color Potts Model with an unusual energy)
Population Score

Sum of square deviation from ideal district population

\[
\sum_{n=1}^{13} \left[ \text{Ideal} - \text{(Pop in district } n) \right]^2
\]

\[
\text{Ideal} = \frac{\text{Population of N.C.}}{13} \approx 733,499
\]
Compactness score

Minimized for a circle

\[ \frac{(\text{Perimeter})^2}{\text{Area}} \geq 4\pi \approx 12.5 \]

Also considered the ratio of district's area to the smallest circumscribing rectangle
Also include score terms for Voting Rights Act and Preserving County Boundaries

Soft penalization:
- for number of split counties of different sizes
- redistricting plans without two districts meeting minimal voting age black population.
Use **Markov Chain Monte Carlo** to sample

Sample: \((\text{density}) \propto e^{-\beta(\text{score of redistricting})}\)
One Step of MCMC Proposal

Then accept/reject according to score function
N.C. Precincts around 3,000
Ensemble of ~24,000 NC redistricting plans

![Bar chart showing the distribution of Democrats elected in 2012 and 2016 congressional votes. The x-axis represents the number of Democrats elected, ranging from 3 to 9, and the y-axis represents the fraction of results. The chart shows a comparison between 2012 and 2016 congressional votes.]
Situate maps in ensemble of 24,000 redistricting plans
Across many elections

Democrats Elected

Democratic Vote Fraction

Statewide Democratic Vote Fraction

0 2 4 6 8 10

0.44 0.46 0.48 0.50 0.52

NC2012
NC2016
Judges

GOV12
USH12
USH16
GOV16
PRE12
PRE16
US14
NCSS16
NC 2012

Atypical

NC 2016

Atypical

Panel of Judges

Typical
Gerrymandering can occur in the absence of oddly shaped districts
What about “Startling” results?
Order the districts by the Blue vote fraction

Percentage of Blue from lowest to highest

<table>
<thead>
<tr>
<th>Most Red</th>
<th>Most Blue</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>10%</td>
<td>60%</td>
</tr>
<tr>
<td>60%</td>
<td>60%</td>
</tr>
<tr>
<td>60%</td>
<td>60%</td>
</tr>
</tbody>
</table>
Order the districts by the Blue vote fraction

Percentage of Blue from lowest to highest

<table>
<thead>
<tr>
<th></th>
<th>Most Red</th>
<th></th>
<th></th>
<th>Most Blue</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>10%</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
</tr>
<tr>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>40%</td>
<td>40%</td>
<td>40%</td>
<td>40%</td>
<td>40%</td>
<td>40%</td>
</tr>
<tr>
<td>20%</td>
<td>30%</td>
<td>40%</td>
<td>50%</td>
<td>60%</td>
<td></td>
</tr>
</tbody>
</table>
Judges

Comp/County

Compact

a 538 plan
Are we sampling the space in a reasonable way?

![Graph showing Democratic vote fraction against congressional votes](image)

- Judges
- 538 - Comp/Cnty
- 538 - Compact
- Medians

Most Republican To Most Democratic Districts

Democratic vote fraction

(2012 congressional votes)

Most Republican To Most Democratic Districts

(2016 congressional votes)
Gerrymandering Index

Probability distribution

(2012 congressional votes)

(2016 congressional votes)
Most Republican To Most Democratic Districts

Democratic vote fraction

(2012 congressional votes)

(2016 congressional votes)

Most Republican To Most Democratic Districts
Gerrymandering Index

“Outlier analysis” Eric Lander's Amicus Brief in Gill v. Whitford
Signature of Gerrymandering

Most Republican To Most Democratic Districts

Democratic vote fraction

Medians

538 - GOP
538 - Dem
NC2012
NC2016
Judges

(2012 votes)
Most Republican To Most Democratic Districts

(2016 votes)
Signature of Gerrymandering

Two principle plots presented in Common Cause v. Rucho

Identify Cracked and Packed districts
The Signature of Gerrymandering and its effects
Wisconsin General Assembly

Ensemble of around 19,000 districting plans
Wisconsin historical elections

Fraction of Republican vote vs. Number of Republican seats

- USH12
- GOV12
- GOV14
- WSA16
- WSA14
- WSA12
- PRE16

Number of Republican seats:
- 45
- 50
- 55
- 60
- 65
- 70
- 75

Fraction of Republican vote:
- 0.48
- 0.49
- 0.50
- 0.51
- 0.52
- 0.53
- 0.54
Wisconsin historical elections

Fraction of Republican vote vs. Number of Republican seats

- WSA16
- GOV12
- GOV14
- USH14
- WSA14
- PRE16
- USH12
- WSA12
Wisconsin historical elections

Fraction of Republican vote

Number of Republican seats

Firewall
Stagnating NC election results due to Gerrymandering

Statewide Democratic Vote Fraction

Democrats Elected

Democratic vote fraction
Back to WI...

Structural advantage exists; sampling decouples geopolitical effects from Gerrymandered effects
Where is Gerrymandering occurring?

Localized analysis
Precinct Level Analysis

vote fraction at predict level

is this precinct gerrymandered?
Precinct Level Analysis

Vote fraction at predict level

Pick a districting plan
Precinct Level Analysis

vote fraction at predict level

the district has a partisan vote fraction
Precinct Level Analysis

Example District Map

vote fraction at predict level

Democratic vote fraction

Probability

0.3 0.4 0.5 0.6 0.7

0 5 10
NC 2012

Red = more Republican than expected
Blue = more Democratic than expected

vote fraction at predict level
average (signed) log likelihood of NC2016 district level results relative to ensemble

vote fraction at predict level
NC 2016 - Triangle

average (signed) log likelihood of NC2016 district level results relative to ensemble

vote fraction at predict level
NC Beyond Gerrymandering Judges

average (signed) log likelihood of Judges district level results relative to ensemble

vote fraction at predict level
Local analysis can detect which districts have been Gerrymandered
Stability of Conclusions
Common Metrics
Common Metrics

- **Efficiency Gap: (McGhee & Stephanopoulos)**

  \[
  \text{Waste} = \sum_{\text{districts won}} (\text{vote fraction} - 0.5) + \sum_{\text{districts lost}} \text{vote fraction}
  \]

  \[
  \text{EG} = \text{Waste(Democrat)} - \text{Waste(Republican)} \\
  \approx [\text{Vote(Dem)} - \text{Vote(Rep)}] - \frac{1}{2} [\text{Seats(Dem)} - \text{Seats(Rep)}]
  \]

  Bernstein & Duchin ‘16

- **Partisan Bias: (Gelman and King)**

  difference in seat fraction won by the Republicans if they receive 55% of the vote and the seat fraction won by the Democrats if they receive 55% of the vote (under partisan swing assumption).
Efficiency Gap

More Wasted R Votes  More Wasted D Votes

Probability distribution

Efficiency gap

More Wasted R Votes  More Wasted D Votes

NC2012  NC2016

Judges

Probability distribution

Efficiency gap

More Wasted R Votes  More Wasted D Votes

NC2012  NC2016

Judges
Efficiency Gap

Efficiency gaps over all elections and plans

Statewide Republican Vote Fraction

0.48 0.50 0.52 0.54 0.56 0.58

Efficiency Gap

-0.2 0 0.2

GOV12 USH16 USS16 PRE16 USS14 PRE16 GOV16 ATT16 USH12 NCSS16
Partisan Bias

Fraction w/ result

Dem. Bias  
Rep. Bias

Partisan bias

(2012 votes)

(2016 votes)
Partisan Bias

Statewide Republican Vote Fraction

Partisan Bias over all elections and plans

Percent of result
The Team

Jonathan Mattingly  Christy Graves  Sachet Bangia  Sophie Guo  Bridget Dou

Justin Luo  Hansung Kang  Robert Ravier  Greg Herschlag  Michael Bell


Duke MATH  Data+  IID  Duke  POLIS

https://sites.duke.edu/quantifyinggerrymandering/
Conclusions

• Compact districts do not preclude gerrymandering
• Sampling techniques can detect gerrymandering
• Provide a Null model against which to compare
• Local analyses indicate which districts have been gerrymandered
Math Questions ?
Assume the population is uniform

model a random distribution of political parties

**Q:** Find null distribution of order statistics for district make up
Q: Give some form of stability of plots over a class of energy functions which have certain marginal statistics.
Q: Characterize the structure of the energy landscape

Even with just population and compactness

some evidence of phase transitions, and shattering of phase space
Accelerate the sampling

- parallel tempering
- accelerated sampling
- hierarchical sampling
- parallel algorithms
One Step of MCMC Proposal

Then accept/reject according to score function
More details on Wisconsin
Shift the global percentages

Under Uniform Partisan Swing Assumption
seats vs global vote (Wisconsin)
Measuring Representativeness

\[ \ell(\text{map}) = - \log \text{Prob(outcome map produces)} \]

Average \( \ell(\text{map}) \) over shift
Measuring Representativeness

The Wisconsin plans are clearly an outlier for the average log likelihood over shifts 45%-55%
Engineered?

results should be stable under small changes to districts

sample near by districts and observe changes
Judge’s districts resemble near by districts

NC 2012 and NC 2016 do not
Gerrymander Index

Local Perturbations

Fraction w/ worse index

Gerrymandering index (2012 votes)

NC2012

NC2016

Judges