Mathematics of Gerrymandering Part II

Mathematical Foundations of Democracy
AMS-IMD Engaged Pedagogy Series
April 25-27, 2023

Daryl DeFord

Department of Mathematics and Statistics
Washington State University

daryl.deford@wsu.edu
math.wsu.edu/faculty/ddeford/

AMERICAN MATHEMATICAL SOCIETY

INSTITUTE FOR MATHEMATICS AND DEMOCRACY

Introduction

Outline

1. Context
2. Applied Motivations
    - **Activity:** Tradeoffs and Metrics
3. Where is the math?
    - **Activity:** Operationalization
4. Discrete Sampling and MCMC
    - **Activity:** Cooperative War
5. Drawing maps with Districtr
    - **Activity:** All about Iowa
6. Applied Ensemble Analysis
    - **Activity:** Perfect Matchings
7. Game Theoretic Redistricting
    - **Activity:** ‘I cut, you freeze’ and Oversplitting
8. Group Discussion

Structure of this presentation

- Exposition
- Examples
- Activities
- Suggestions
- Resources

Possible Course Contexts

- Module in discrete mathematics or graph theory
- Module in statistical computing or probability
- Module in mathematical modeling
- Full course on Computational Redistricting
- ...

Possible Course Contexts

- Module in discrete mathematics or graph theory
- Module in statistical computing or probability
- Module in mathematical modeling
- Full course on Computational Redistricting
- ...
- As a relatively new field many of these topics push up against research frontiers and I'll give some examples throughout of potential student projects and motivations from recent papers. This also means that there are clear opportunities for community engagement and empirical work with real data.

Moral:

Computational Redistricting is not a solved problem!

☹ / ☺

Moral:

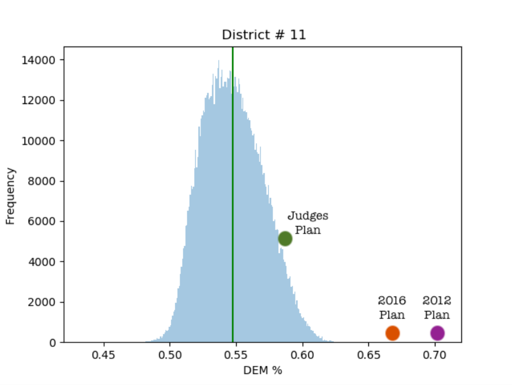
Computational Redistricting is not a solved problem!

☹ / ☺

Practical relevance and cultural understanding of mathematics.

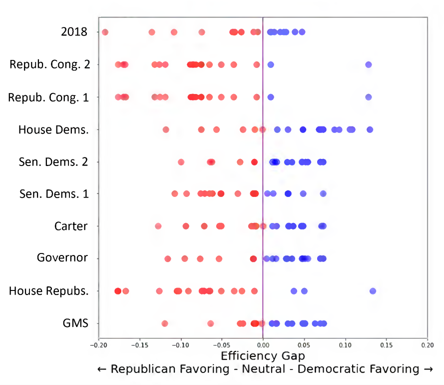
Applied Motivations

- **Court cases**
  - Detecting gerrymandering
  - Evaluating proposed remedies
- Reform Efforts
  - Establishing baselines
  - Modeling impact of changes
- Initial plan evaluation
  - Detecting unintentional gerrymandering [1]
  - Lawsuit prevention



[1] with apologies to J. Chen and J. Rodden, Unintentional Gerrymandering, Political Geography and Electoral Bias in Legislatures, Quarterly Journal of Political Science, 8, pp. 239–269, 2013.

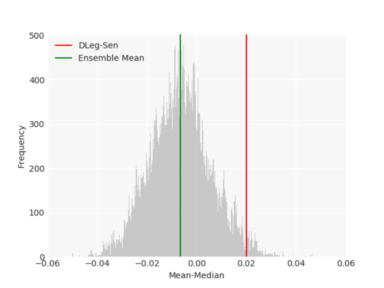
- **Court cases**
  - Detecting gerrymandering
  - Evaluating proposed remedies
- Reform Efforts
  - Establishing baselines
  - Modeling impact of changes
- Initial plan evaluation
  - Detecting unintentional gerrymandering [1]
  - Lawsuit prevention

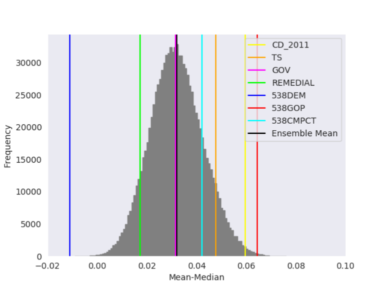


[1] with apologies to J. Chen and J. Rodden, Unintentional Gerrymandering, Political Geography and Electoral Bias in Legislatures, Quarterly Journal of Political Science, 8, pp. 239–269, 2013.

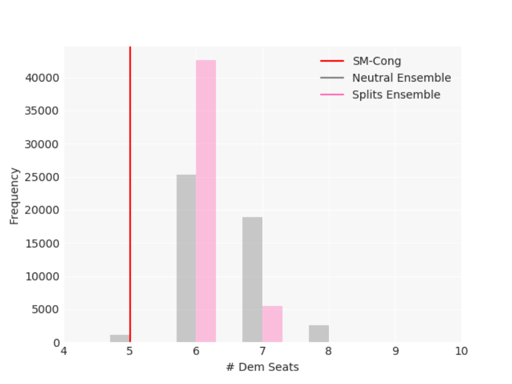
Mathematical Applications

- Court cases
  - Detecting gerrymandering
  - Evaluating proposed remedies
- **Reform Efforts**
  - Establishing baselines
  - Modeling impact of changes
- Initial plan evaluation
  - Supporting Commission Work
  - Exploring/Optimizing all plans

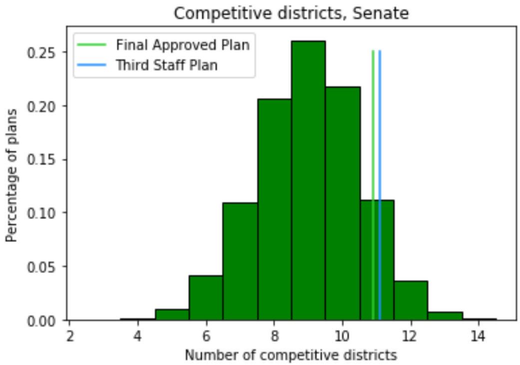




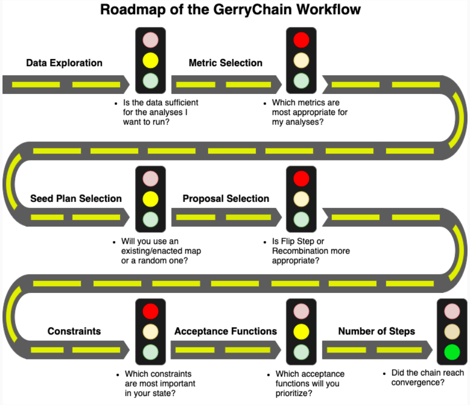
- Court cases
  - Detecting gerrymandering
  - Evaluating proposed remedies
- **Reform Efforts**
  - Establishing baselines
  - Modeling impact of changes
- Initial plan evaluation
  - Supporting Commission Work
  - Exploring/Optimizing all plans



- Court cases
  - Detecting gerrymandering
  - Evaluating proposed remedies
- Reform Efforts
  - Establishing baselines
  - Modeling impact of changes
- **Initial plan evaluation**
  - Supporting Commission Work
  - Exploring/Optimizing all plans



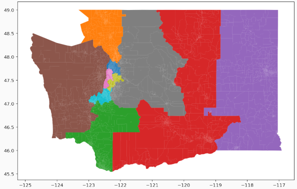
- Court cases
  - Detecting gerrymandering
  - Evaluating proposed remedies
- Reform Efforts
  - Establishing baselines
  - Modeling impact of changes
- **Initial plan evaluation**
  - Supporting Commission Work
  - Exploring/Optimizing all plans

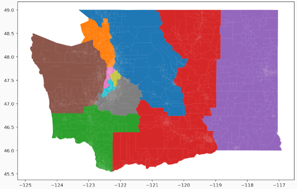


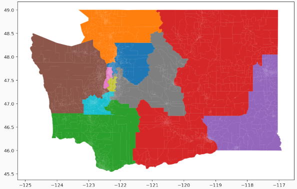
Activity: Tradeoffs and Metrics

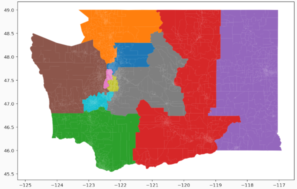
Motivating the Math

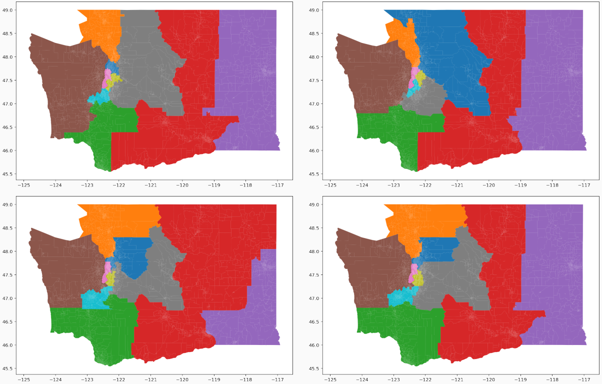
Proposed Washington Maps











Motivating Question: How can mathematics help us to evaluate districting plans?

Where do districts come from?

In many states, the text governing the redistricting process and procedures are outlined in the state constitution.

Where do districts come from?

In many states, the text governing the redistricting process and procedures are outlined in the state constitution.

- Who do the districts elect?
- How many districts are there?
- Who draws the lines?
- Who approves the map?
- Who can change the rules?

In many states, the text governing the redistricting process and procedures are outlined in the state constitution.

- Who do the districts elect?
- How many districts are there?
- Who draws the lines?
- Who approves the map?
- Who can change the rules?
- **What constraints are districts supposed to satisfy?**

Permissible Districting Plans

- Common Constraints
  - Contiguity
  - Population Balance
  - Compactness
  - VRA Compliance
  - Communities of Interest
  - Municipal Boundaries
  - Competitiveness/Responsiveness
  - Incumbency Protection
  - ...

Permissible Districting Plans

- Common Constraints
  - Contiguity
  - Population Balance
  - Compactness
  - VRA Compliance
  - Communities of Interest
  - Municipal Boundaries
  - Competitiveness/Responsiveness
  - Incumbency Protection
  - ...

- Modeling Concerns
  - Relative Ordering
  - Operationalization
  - Tradeoffs
  - Data Challenges
  - ...

Operationalization Examples

- Population Balance:
  - What does ± 1 mean?
  - What norm to measure deviation?
  - How much does variance matter?
  - Within-cycle malapportionment

[1] A. Becker and D. Gold: The gameability of redistricting criteria, Journal of Computational Social Science, 5, 1735–1777 (2022).

Operationalization Examples

- Population Balance:
  - What does ± 1 mean?
  - What norm to measure deviation?
  - How much does variance matter?
  - Within-cycle malapportionment
- Boundary Preservation[1]
  - Split Counties
  - Total Pieces
  - Pieces over required
  - Population?
  - (not) Nesting
  - Fracking
  - Discontiguous Units

[1] A. Becker and D. Gold: The gameability of redistricting criteria, Journal of Computational Social Science, 5, 1735–1777 (2022).

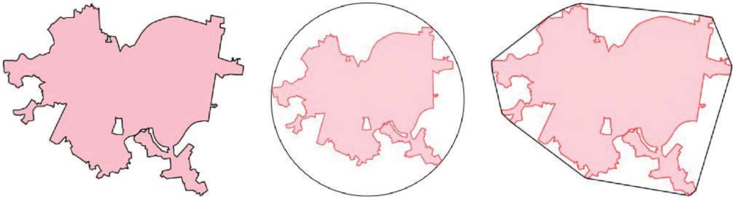
Competitiveness Operationalized

- WA: “provide fair and effective representation and to encourage electoral competition”
- NY: “districts not be drawn to discourage competition”
- AZ: “To the extent practicable, competitive districts should be favored where to do so would create no significant detriment to the other goals”
- CO: “maximize the number of politically competitive districts,” where competitive is defined as “having a reasonable potential for the party affiliation of the district’s representative to change at least once between federal decennial censuses”
- MO: Swing EG optimization
- NJ (proposed): Statewide average baseline
- Intuitive: ‘Bands’ around some fixed percentage[1]

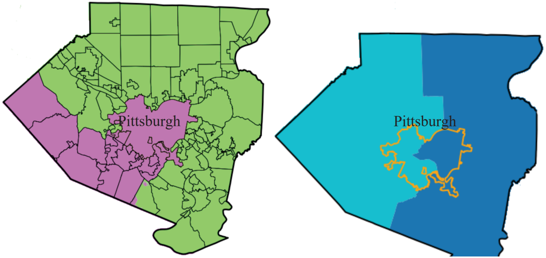
[1] D. DeFord, M. Duchin, and J. Solomon: A Computational Approach to Measuring Vote Elasticity and Competitiveness, Statistics and Public Policy, 7(1), 69-86, (2020).

Activity: Operationalization

Tradeoff: Compactness vs. City Preservation



(a) City Boundary (b) Bounding Circle (c) Convex Hull



Tradeoff: Compactness vs. Population Balance

- Toy Model: 50×50 grid where each node has population 1, split into two districts A and B
- Two constraints:
  - Population balance (difference between districts) $||A| - |B||$
  - Compactness (cut edges) $|\partial(A, B)|$
- Experiment: Use simulated annealing to sample from:

$$e^{-\beta(\gamma \text{pop deviation} + \text{cut edges})}$$

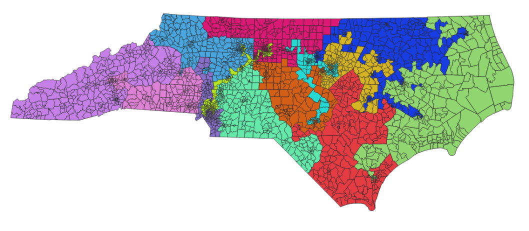
  as γ varies
- Phase transition? Rigidity?

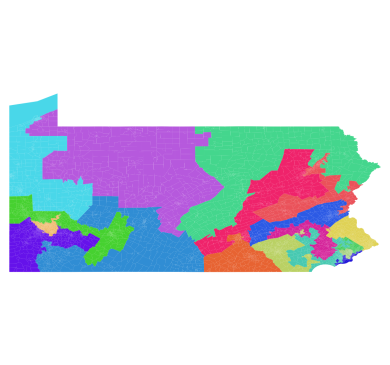
E. Najt, D. DeFord, and J. Solomon, Empirical Sampling of Connected Graph Partitions for Redistricting, Physical Review E, 064130, (2021).

More Metrics and Context


HAVERHILL
AMESBURY
SALSBURY
METHUEN
BRADFORD
N. PORT
NEWBURY
ANDOVER
ROXFORD
ROWLEY
MIDDLETON
TOPSFIELD
IPSWICH
LYNNFIELD
HAMILTON
WENHAM
DANVERS
BEVERLY
MANCHESTER
GLOCESTER
SALEM
MARBLE HEAD
LYNN
CHELSEA

Geometric Measurements





Problem (Barnes and Solomon 2018[1])

Geographic Compactness scores can be distorted by:

- *Data resolution*
- *Map projection*
- *State borders and coastline*
- *Topography*
- *...*

Theorem (Bar-Natan, Najt, and Schutzman 2019[2])

There is no local homeomorphism from the globe to the plane that preserves your favorite compactness measure.

[1] Gerrymandering and Compactness: Implementation Flexibility and Abuse, Political Analysis, 29(4), (2021).

[2] The Gerrymandering Jumble: Map Projections Permute Districts' Compactness Scores, Cartography and Geographic Information Science, 47(4), (2020).

Theorem (D., Lavenant, Schutzman, and Solomon 2019[1])

The total variation relaxation of the isoperimetric profile:

1. *satisfies an isoperimetric inequality,*
2. *is the lower convex envelope of the original profile,*
3. *and admits a distinguished family of efficiently computable solutions that recover the Cheeger set and take at most three values for every t.*

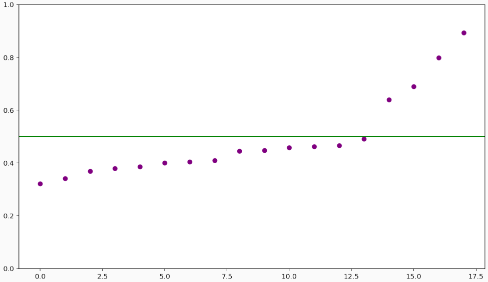
Theorem (Zhang, D., and Solomon 2020[2])

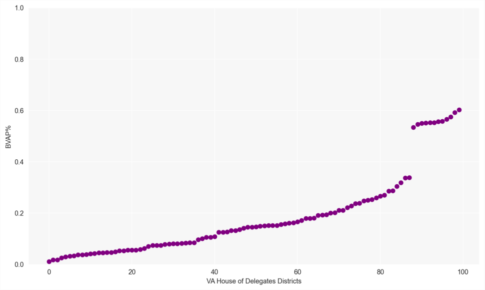
The greedy medial axis transform algorithm gives an upper bound on the isoperimetric profile that is tight on intervals containing both $t = 0$ and $t = \mathrm{Area}(\Omega)$.

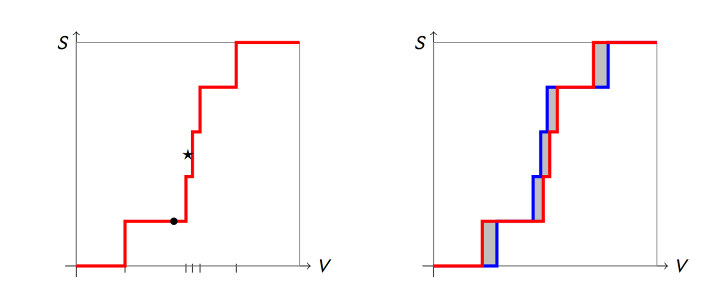
[1] Total Variation Isoperimetric Profiles, SIAM J. Appl. Algebra Geometry, 3(4), (2019).
[2] Medial Axis Isoperimetric Profiles, (with P. Zhang and J. Solomon) Symposium on Graphics Processing, Computer Graphics Forum, 39(5), (2020).

Theory Solidified in: Leonardi, G.P., Saracco, G. The isoperimetric problem in 2d domains without necks. Calc. Var. 61, 56, (2022).

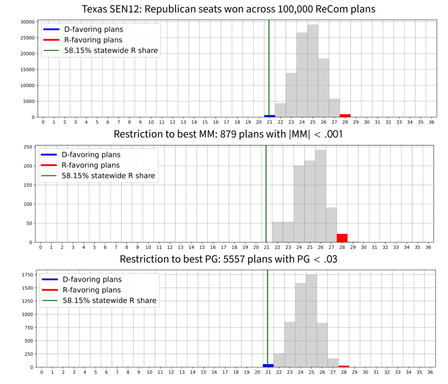
Packing and Cracking





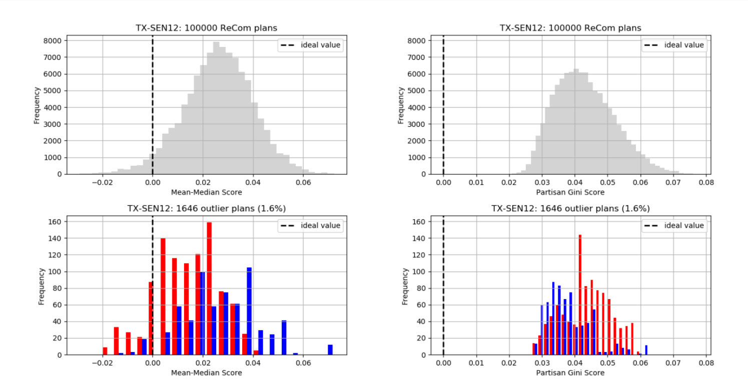

S
V
S
V

Exploitability



D. DeFord, N. Dhamankar, M. Duchin, V. Gupta, M. McPike, G. Schoenbach, K. W. Sim: Implementing partisan symmetry: Problems and paradoxes, Political Analysis, (2021).

Exploitability



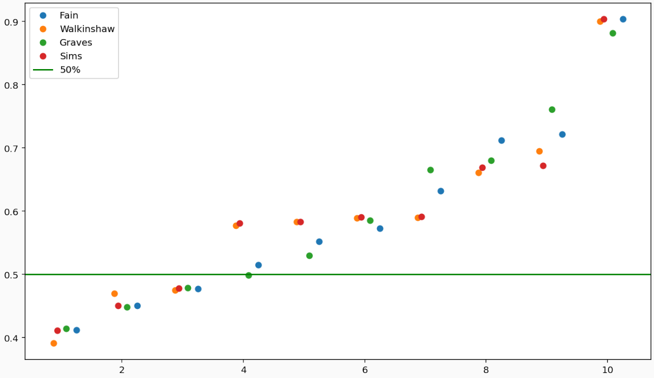
D. DeFord, N. Dhamankar, M. Duchin, V. Gupta, M. McPike, G. Schoenbach, K. W. Sim: Implementing partisan symmetry: Problems and paradoxes, Political Analysis, (2021).

- Problems
  - These examples higlights exploitability as a potential problem with individual geometry or partisan symmetry metrics (or any other single scalar measure)
  - We would also prefer targeted baselines, rather than generic bounds
  - Finally, incorporating not just the boundaries but where the voters live (political geography) is critical

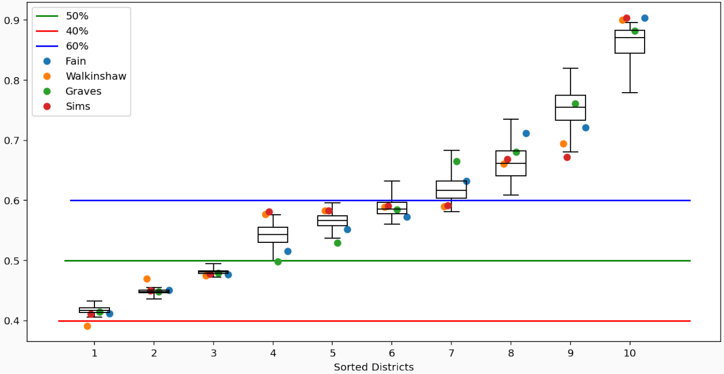
The Ensemble Method

- Problems
  - These examples higlights exploitability as a potential problem with individual geometry or partisan symmetry metrics (or any other single scalar measure)
  - We would also prefer targeted baselines, rather than generic bounds
  - Finally, incorporating not just the boundaries but where the voters live (political geography) is critical
- Ensembles
  - Randomly construct **many** feasible maps and compute metrics on those to understand possibilities
  - Incorporate actual census and voting data
  - Formulate relevant counterfactuals

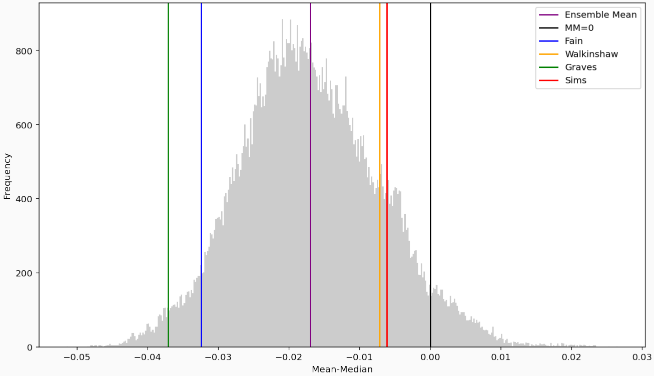
Additional WA Context



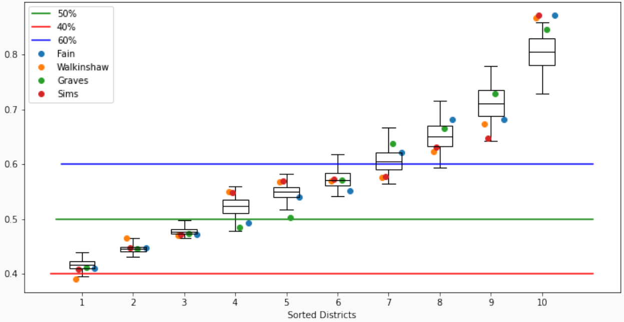
Additional WA Context

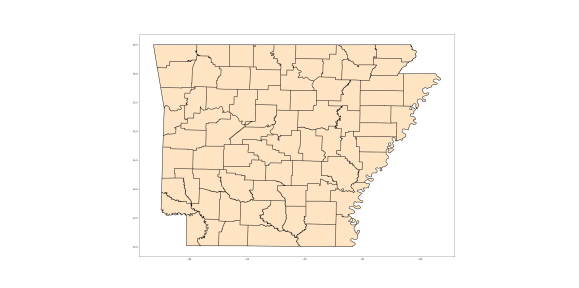


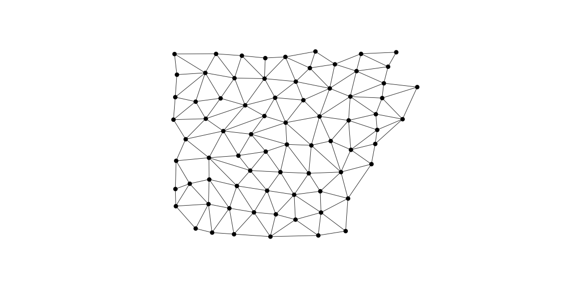
Additional WA Context

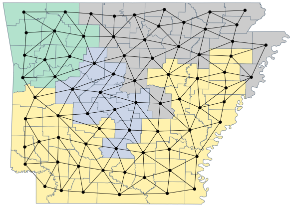


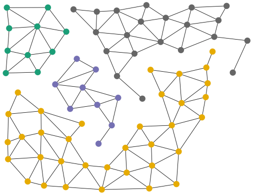
Additional WA Context





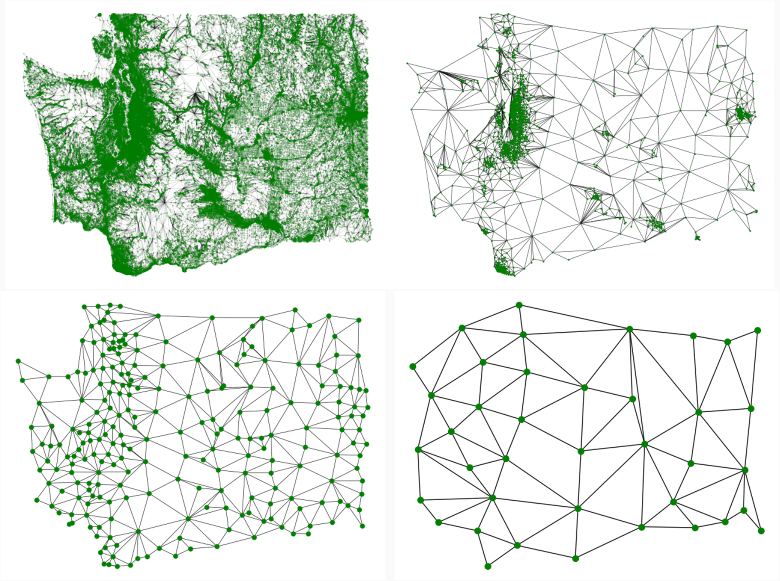




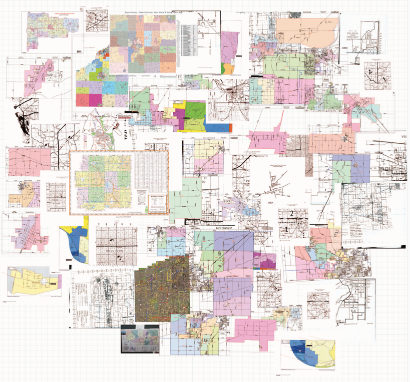


Geospatial Data Challenges

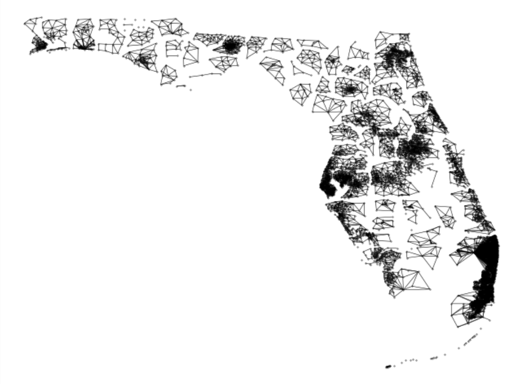
Census Data



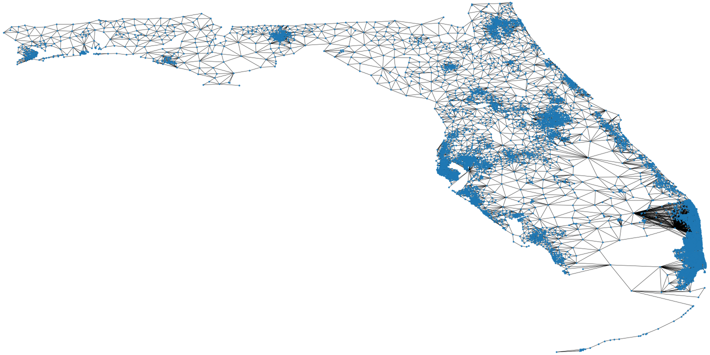
Voting Data (precincts)



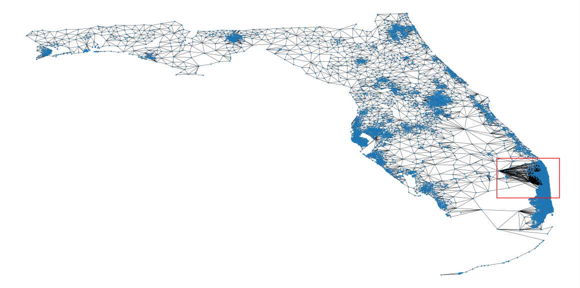
Florida



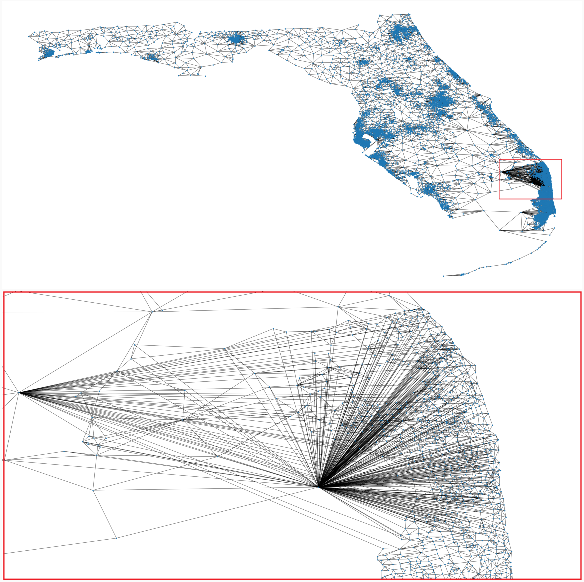
Florida



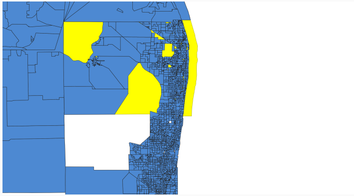
Florida

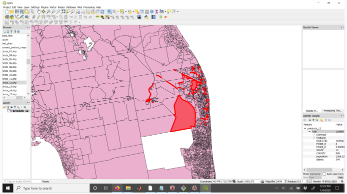


Florida



Florida





- For class use we are at the point where there are reasonable starting points:
  - redistrictingdatahub.org
  - github.com/mggg-states
  - NHGIS/Census API
- However, there are still usually some processing steps needed:
  - Merge multiple datasets
  - Prorate election data
  - Match precincts to current districts
  - Fix errors from matchings
  - Islands!

Activities: All about Iowa or Gridlandia

Sampling Graph Partitions

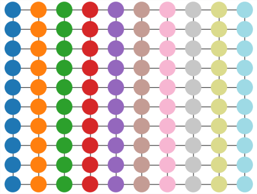
1. Set constraints to define the state space
2. Start with an initial plan
3. **Propose a modification**
4. Verify that the modification satisfies the constraints
5. Accept using MH criterion
6. Repeat

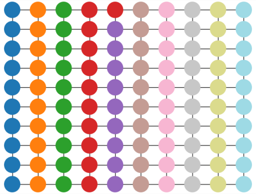
Why?

- Control over sampling distribution and input data
- Possibility of local sampling
- Ergodic Theorem

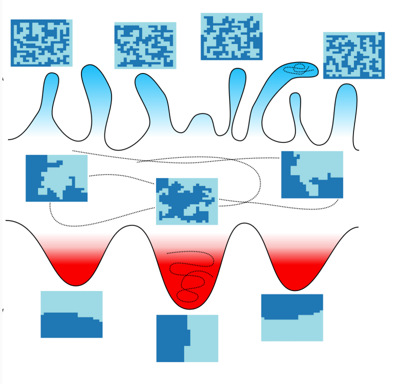
Single Edge Flip Proposals

1. Choose a boundary node
2. Change its assignment





- Herschlag et al., Quantifying Gerrymandering in North Carolina, Statistics and Public Policy, 7(1), (2020). Court cases in NC and WI.
- Pegden et al. Assessing significance in a Markov chain without mixing, PNAS, (2017). Court cases in NC and PA.



E. Najt, D. DeFord, and J. Solomon, Empirical Sampling of Connected Graph Partitions for Redistricting, Physical Review E, 064130, (2021).

Theorem (Najt, D., and Solomon)

Suppose that C is the class of connected planar graphs and $k \geq 2$. If there is a polynomial time algorithm to sample uniformly from:

- *the connected k-partitions of graphs in C,*
- *or the connected, 0-balanced k-partitions of graphs in C.*

then $RP = NP$.

Theorem (Najt, D., and Solomon)

Suppose that C is the class of connected planar graphs and $k \geq 2$. If there is a polynomial time algorithm to sample uniformly from:

- *the connected k-partitions of graphs in C,*
- *or the connected, 0-balanced k-partitions of graphs in C.*

then $RP = NP$.

Better news:

- E. A. Autry, D. Carter, G. Herschlag, Z. Hunter, and J. Mattingly, Metropolized Multiscale Forest Recombination for Redistricting, SIAM MMS, 19 (4), (2021).
- A. Frieze and W. Pegden, Subexponential mixing for partition chains on grid-like graphs SODA '22, (2022).
- A. D. Procaccia and J. Tucker-Fultz, Compact Redistricting Plans Have Many Spanning Trees, SODA '22, (2022).

Probability Terminology: Example

- **Random Variable:** A variable whose value depends on the result of a random experiment **Result of flipping two coins**

Probability Terminology: Example

- **Random Variable:** A variable whose value depends on the result of a random experiment **Result of flipping two coins**
- **Outcome:** A possible result of the experiment **HT**

Probability Terminology: Example

- **Random Variable:** A variable whose value depends on the result of a random experiment **Result of flipping two coins**
- **Outcome:** A possible result of the experiment **HT**
- **Sample Space:** The full collection of outcomes **{HH,TH,HT,TT}**

Probability Terminology: Example

- **Random Variable:** A variable whose value depends on the result of a random experiment **Result of flipping two coins**
- **Outcome:** A possible result of the experiment **HT**
- **Sample Space:** The full collection of outcomes **{HH,TH,HT,TT}**
- **Event:** A set of outcomes **{HH,TH,HT}**

Probability Terminology: Example

- **Random Variable:** A variable whose value depends on the result of a random experiment **Result of flipping two coins**
- **Outcome:** A possible result of the experiment **HT**
- **Sample Space:** The full collection of outcomes **{HH,TH,HT,TT}**
- **Event:** A set of outcomes **{HH,TH,HT}**
- **Probability:** The likelihood of an event **.75**

Probability Terminology: Example

- **Random Variable:** A variable whose value depends on the result of a random experiment **Result of flipping two coins**
- **Outcome:** A possible result of the experiment **HT**
- **Sample Space:** The full collection of outcomes **{HH,TH,HT,TT}**
- **Event:** A set of outcomes **{HH,TH,HT}**
- **Probability:** The likelihood of an event **.75**
- **Probability Distribution:** A function that associates probabilities to all outcomes.
  - **HH** $\mapsto$ .25
  - **HT** $\mapsto$ .25
  - **TH** $\mapsto$ .25
  - **TT** $\mapsto$ .25

- Example: Rolling a die
  - Equal probability of all outcomes
  - **State space** (1,2,3,4,5,6)
  - $\mathbb{P}(2) = \frac{1}{6}$
  - $\mathbb{P}(1, 3, 5) = \frac{1}{2}$
  - Probabilities sum to 1!
- Example: Drawing a red marble out of a jar with three blue marbles and 7 red marbles
  - **State space:** Success: red or Failure: blue
  - **Probabilities** 70% or 30%
  - Probabilities sum to 1!

Discrete Distributions: Probability Mass Function

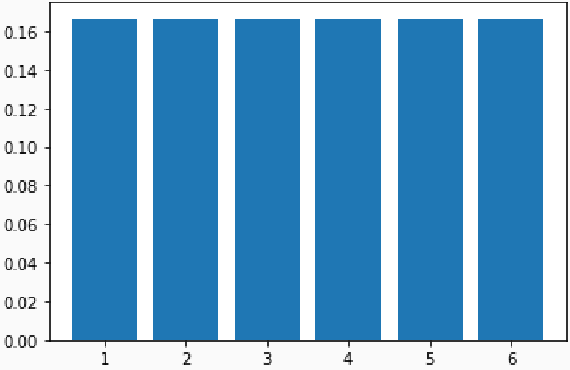


Figure 1: Probability Mass Function single die roll

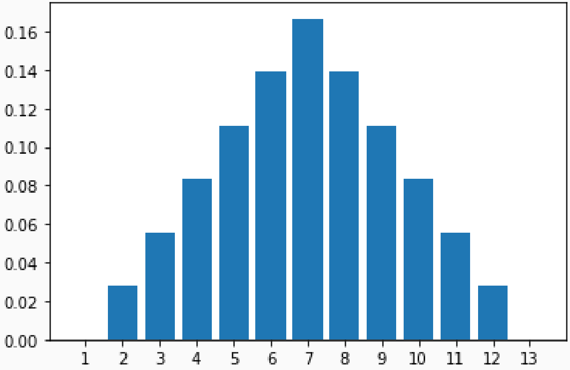


Figure 2: Probability Mass Function for the sum of two die rolls

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M	N
Frequency	9	2	2	4	12	2	3	2	9	1	1	4	2	6
Score	1	3	3	2	1	4	2	4	1	8	5	1	3	1

Letter	O	P	Q	R	S	T	U	V	W	X	Y	Z	' '	
Frequency	8	2	1	6	4	6	4	2	2	1	2	1	2	
Score	1	3	10	1	1	1	1	4	4	8	4	10	0	

Table 1: Frequencies and point values of Scrabble tiles.

- **Question:** What is the expected score if we draw a tile from the Scrabble bag randomly?

[1] or computing expected values over

- **Question:** What is the expected score if we draw a tile from the Scrabble bag randomly?
- Often, we are interested in drawing samples from[1] a complicated distribution over a discrete set, where computing the normalizing constant in infeasible.
- However, if we can draw samples from a controllable distribution over our state space, and evaluate
- This is the key mathematical insight of Markov Chain Monte Carlo (MCMC) sampling, which is one of the most important algorithmic developments of the 20th century.
- **Goal:** Turn samples from a known 'easy' distribution into the correct proportions for a different 'hard' distribution.

[1] or computing expected values over

Rules:

1. Nominate a dealer in your group of 3 players
2. Deal 2 cards to each player
3. Each round begins with the dealer playing the highest value card from their hand.
4. The round continues counterclockwise with each player playing the highest card in their hand **only if** it is higher than the previously played card. If not, skip that player and move on to the next.
5. The round ends once each player has had an opportunity to play a card.
6. You (collectively) win if all players have played all of their cards at the end of the second round.

Cooperative War Example 1

- Hands:
  - Dealer $= \{J\spadesuit, 9\clubsuit\}$, Player 2 $=\{2\spadesuit, 3\clubsuit\}$, Player 3 $= \{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer:
  - Player 2:
  - Player 3:
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 1

- Hands:
  - Dealer = $\{9\clubsuit\}$, Player 2 =$\{2\spadesuit, 3\clubsuit\}$, Player 3 = $\{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2:
  - Player 3:
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 1

- Hands:
  - Dealer = $\{9\clubsuit\}$, Player 2 =$\{2\spadesuit, 3\clubsuit\}$, Player 3 = $\{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2:☹
  - Player 3:
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer = $\{J\spadesuit, 9\clubsuit\}$, Player 2 =$\{Q\spadesuit, K\clubsuit\}$, Player 3 = $\{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer:
  - Player 2:
  - Player 3:
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer $= \{9\clubsuit\}$, Player 2 $=\{Q\spadesuit, K\clubsuit\}$, Player 3 $= \{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2:
  - Player 3:
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer = $\{9\clubsuit\}$, Player 2 =$\{Q\spadesuit\}$, Player 3 = $\{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2: $K\clubsuit$
  - Player 3:
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer = $\{9\clubsuit\}$, Player 2 =$\{Q\spadesuit\}$, Player 3 = $\{3\clubsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2: $K\clubsuit$
  - Player 3: $K\heartsuit$
- Round 2:
  - Dealer:
  - Player 2:
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer = {}, Player 2 =$\{Q\spadesuit\}$, Player 3 = $\{3\clubsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2: $K\clubsuit$
  - Player 3: $K\heartsuit$
- Round 2:
  - Dealer: $9\clubsuit$
  - Player 2:
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer = $\{\}$, Player 2 =$\{\}$, Player 3 = $\{3\clubsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2: $K\clubsuit$
  - Player 3: $K\heartsuit$
- Round 2:
  - Dealer: $9\clubsuit$
  - Player 2: $Q\spadesuit$
  - Player 3:

Cooperative War Example 2

- Hands:
  - Dealer = {}, Player 2 ={}, Player 3 = {3♣}
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2: $K\clubsuit$
  - Player 3: $K\heartsuit$
- Round 2:
  - Dealer: 9♣
  - Player 2: $Q\spadesuit$
  - Player 3: ☹

Cooperative War Example 2

- Hands:
  - Dealer = $\{J\spadesuit, 9\clubsuit\}$, Player 2 =$\{Q\spadesuit, K\clubsuit\}$, Player 3 = $\{3\clubsuit, K\heartsuit\}$
- Round 1:
  - Dealer: $J\spadesuit$
  - Player 2: $K\clubsuit$
  - Player 3: $K\heartsuit$
- Round 2:
  - Dealer: $9\clubsuit$
  - Player 2: $Q\spadesuit$
  - Player 3: ☹

Question

What is the probability that you win, given a randomly shuffled deck?

Question

What is the probability that you win, given a randomly shuffled deck?

Answer

Try it out!

Question

What is the probability that you win, given a randomly shuffled deck?

Answer

Try it out!

Answer

Simulate! http://math.wsu.edu/faculty/ddeford/solitaire.html

Interlude: How many shuffles?

- In the previous example we made implicit use of the fact that we can tell when a deck is 'randomly' shuffled.
- A commonly repeated fact is that it takes 7 riffle shuffles to sufficiently randomize a 52 card deck but there are many different types of shuffles
- In general:
  - The top to random shuffle takes about $n\log(n)$ (90 for a 52 card deck)
  - The random transposition shuffle takes: $\frac{1}{2}n\log(n)$ (45)
  - The riffle shuffle takes about $\frac{3}{2}\log_2(n)$ (9)
- The proofs of these results rely on exactly the MCMC method that we are going to discuss next.

Question

What is the expected distance between two random points on $[0,1]$*?*

Question

What is the expected distance between two random points on $[0,1]$*?*

Answer

$$\int_0^1 \int_0^1 |x-y|dxdy = \frac{1}{3}$$

Question

What is the expected distance between two random points on $[0, 1]$*?*

Answer

$$\int_0^1 \int_0^1 |x - y| dxdy = \frac{1}{3}$$

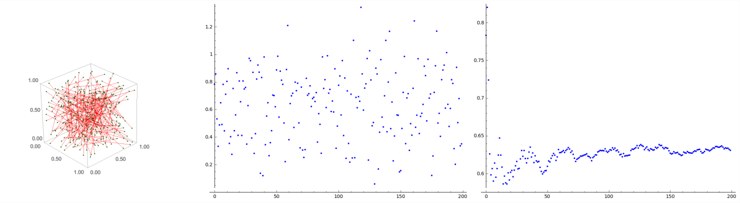
Question

What is the expected distance between two random points on $[0, 1]^n$*?*

Answer

$$\int_0^1 \cdots \int_0^1 \sqrt{\sum_{j=1}^{n} (x_j - y_j)^2} dx_1 \cdots dx_n dy_1 \cdots dy_n = \quad ☹$$

Try it out: Cube Distances



Properties of Monte Carlo Methods

- Draw (independent) samples from a random distribution
- Compute some measure for each draw
- Repeat lots and lots of times
- Average/aggregate the derived data

- Draw (independent) samples from a random distribution
- Compute some measure for each draw
- Repeat lots and lots of times
- Average/aggregate the derived data

Today, we tend to take access to random numbers for granted but some of the first applications of Monte Carlo were physical systems for generating random numbers. (Our card game example is also historically relevant!)

Definition (Markov Chain)

A sequence of random variables $X_1, X_2, \ldots,$ is called a Markov Chain if

$$\mathbb{P}(X_n = x_n : X_1 = x_1, \ldots, X_{n-1} = x_{n-1}) = \mathbb{P}(X_n = x_n : X_{n-1} = x_{n-1}).$$

What is a Markov chain?

Definition (Markov Chain)

A sequence of random variables $X_1, X_2, \ldots,$ is called a Markov Chain if

$$\mathbb{P}(X_n = x_n : X_1 = x_1, \ldots, X_{n-1} = x_{n-1}) = \mathbb{P}(X_n = x_n : X_{n-1} = x_{n-1}).$$

Definition (Transition Probability)

Given a finite state space $X = x_1, x_2, \ldots, x_n$ we can specify a Markov chain over X with transition probabilities $p_{i,j} = \mathbb{P}(X_m = i : X_{m-1} = j)$ and associated transition matrix $P = [p_{i,j}]$.

Definition (Random Walk)

We can also view a Markov chain as random walk on a directed, weighted graph, with weights given by the $p_{i,j}$. (mggg.org/metagraph)

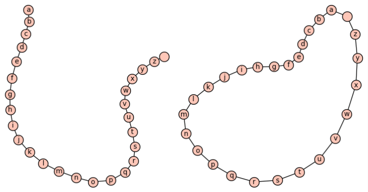
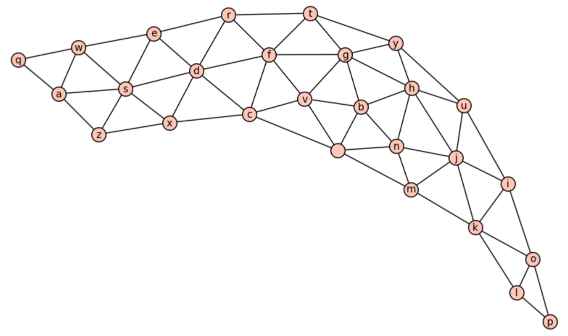


Figure 3: The alphabet path and cycle graphs.



(a) Keyboard Adjacency

Ant on a Keyboard

Letter Walk Examples

- Path
- Cycle
- Keyboard
- i.i.d. Scrabble
- i.i.d. Uniform

Try it out: Markov Chain Samples

Try it Out!

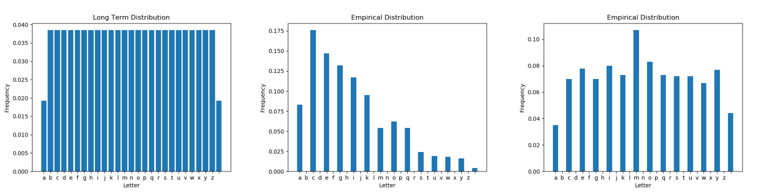
Example

Which of the 5 Markov chains are closest to their stationary distribution after 50 steps? 1000 steps?

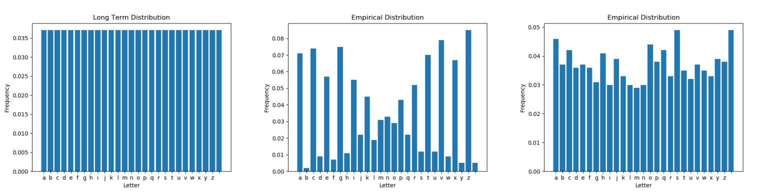
Example

Which combination of walk and starting letter leads to the most attractive distribution plot?

Path Walk

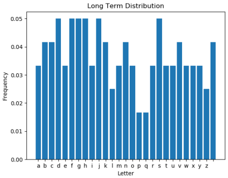

Long Term Distribution
Empirical Distribution
Empirical Distribution
Frequency
Letter

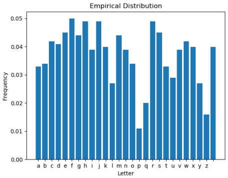
Cycle Walk

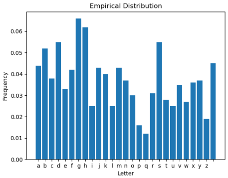


Long Term Distribution
Frequency
Letter
Empirical Distribution
Frequency
Letter
Empirical Distribution
Frequency
Letter

Keyboard Walk







Desirable Adjectives

- Irreducible: A chain is irreducible if each state is (eventually) reachable from every other state.
- Aperiodic: A chain is aperiodic if for each state, the GCD of the lengths of the loops, starting and ending at that state is equal to 1.
- Steady State Distribution: A distribution π is said to be a steady state of the chain if $\pi = \pi P$. For simple random walks on graphs this is proportional to the degree of each node.

Key Theorem

If the chain is **irreducible** and **aperiodic** then $\lim_{m \to \infty} P^m = 1\pi$ for a unique π. Even better, if f is any function defined on the state space $y_1, y_2, \ldots, y_m$ are samples from π then,

$$\lim_{m \to \infty} \frac{1}{m} \sum_{i=1}^{m} f(y_i) = \mathbb{E}[f]$$

The key idea of MCMC is to create an irreducible, aperiodic Markov chain whose steady state distribution π is the distribution we are trying to sample from.

What is MCMC?

In our Monte Carlo methods we just required that we sample from our space uniformly but this isn't always easy to do. MCMC gives us a way to sample from a desired pre–defined distribution by forming a related Markov chain (or walk) over our state space, with transition probabilities determined by a multiple of the distribution that we are trying to sample from.

A common way this arises is when we have a score function or a ranking on our state space and want to draw proportionally to these scores. Given a score $s : X \to \mathbb{R}$ we want to sample from the distribution where the states appear proportional to s. That is, element $y \in X$ should appear with probability

$$\mathbb{P}(y) = \frac{s(y)}{\sum_{x \in X} s(x)}.$$

When $|X|$ is enormous, we don't want to/can't compute the denominator.

Notice that we can compute ratios of probabilities, since the denominators cancel:

$$\frac{\mathbb{P}(z)}{\mathbb{P}(y)} = \frac{\dfrac{s(z)}{\sum_{x \in X} s(x)}}{\dfrac{s(y)}{\sum_{x \in X} s(x)}} = \frac{s(z)}{s(y)}.$$

This is the trick that turns out to allow us to draw samples according to s without having to compute the denominator directly.

Score: A function $s : X \to \mathbb{R}_{\geq 0}$ that determines our target distribution.

Score: A function $s : X \to \mathbb{R}_{\geq 0}$ that determines our target distribution.

Proposal Distribution: A Markov chain Q over X with the property that $Q(x_j : x_i)$.

Terminology

Score: A function $s : X \to \mathbb{R}_{\geq 0}$ that determines our target distribution.

Proposal Distribution: A Markov chain Q over X with the property that $Q(x_j : x_i)$.

Metric: Another function $f : X \to \mathbb{R}$ that is our quantity of interest for the distribution.

Metropolis Procedure

Given that we have a given score function, proposal distribution, metric, and initial state X_0, at each step of the Metropolis–Hastings chain $X_1, X_2, \ldots$ we follow this sequence of steps, assuming that we are currently at state $X_k = y$:

1. Generating a proposed state $\hat{y}$ according to $Q_{y,\hat{y}}$.
2. Compute the acceptance probability:
$$\alpha = \min\left(1, \frac{s(\hat{y})}{s(y)} \frac{Q_{\hat{y},y}}{Q_{y,\hat{y}}}\right)$$
3. Pick a number β uniformly on $[0, 1]$
4. Set
$$X_{k+1} = \begin{cases} \hat{y} & \text{if } \beta < \alpha \\ y & \text{otherwise.} \end{cases}$$

- Uniform
- Scrabble Points
- Scrabble Counts
- Alphabetical (a=1, b=2, etc.)
- Vowels (consonants = 1, vowels = 100, y = 50)

Try it out: Compute Expected Values

Try it out!

Example

Using the Keyboard walk, which score function converges to the right expected value fastest.

Example

For the path walk and alphabetical score, which starting point converges most slowly?

Example

For the keyboard walk and vowel score, which starting point converges fastest?

Example

Are there any pure states that converge faster than uniform states?

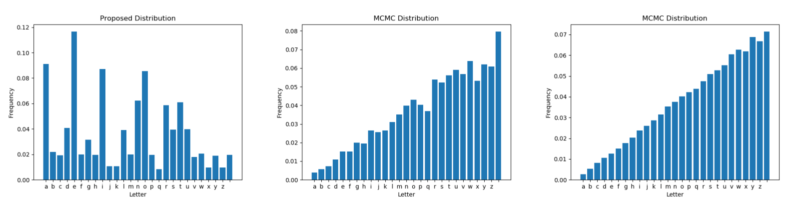
Example

Can you find a random state that requires more than 5 steps to get below $\frac{1}{4}$?

Example

Using the path walk, which starting letter mixes fastest to the vowel distribution?

Letter Example



Try it out: here

First Step

1. We uniformly pick a key on the keyboard next to 'a': 'q'.
2. We next need to compute some numbers in order to compute the acceptance probability:
   - $s(a) = 1$
   - $s(q) = 10$
   - $Q_{q,a} = \frac{1}{2}$ since 'q' has two neighbors
   - $Q_{a,q} = \frac{1}{3}$ since 'a' has three neighbors

   These let us compute:

$$\alpha = \min(1, \frac{10}{1}\frac{\frac{1}{2}}{\frac{1}{3}}) = \min(1, 15) = 1$$

3. Uniformly pick $\beta = .188256$
4. Set the next state to be 'q' since $\beta < \alpha$.

This makes sense with our interpretation, since 'q' has a higher score than 'a' we should always accept this transition.

Second Step

Now we will try to take another step from ‘q’:

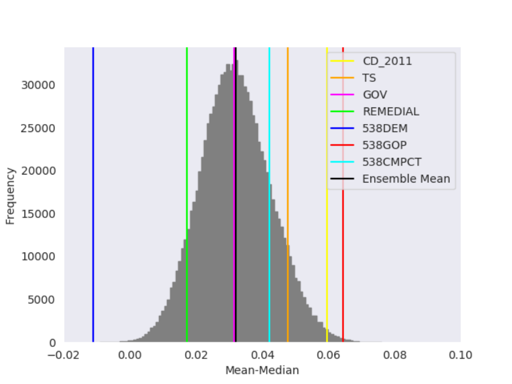
1. We uniformly pick a neighbor of ‘q’ and get ‘w’.
2. We again need to compute some numbers:
    - $s(q) = 10$
    - $s(w) = 4$
    - $Q_{q,w} = \frac{1}{2}$ since ‘q’ has two neighbors
    - $Q_{w,q} = \frac{1}{3}$ since ‘w’ has four neighbors

    These let us compute:

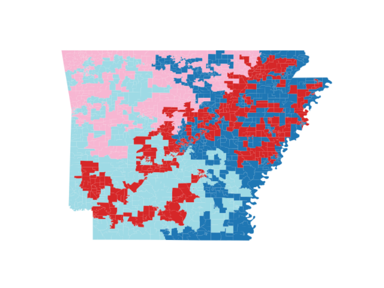
$$\alpha = \min(1, \frac{4}{10}\frac{\frac{1}{4}}{\frac{1}{2}}) = \min(1, .2) = .2$$

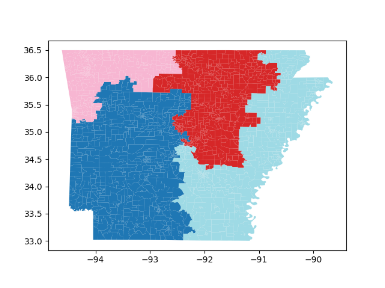
3. Uniformly pick $\beta = .7593544$
4. Set the next state to be ‘q’ since $\beta > \alpha$.

Ensemble Example



Which Ensembles?

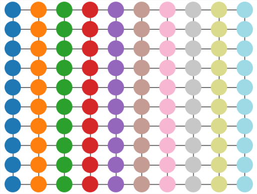


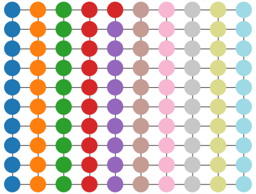


Activity: Metagraph Exploration

Single Edge Flip Proposals

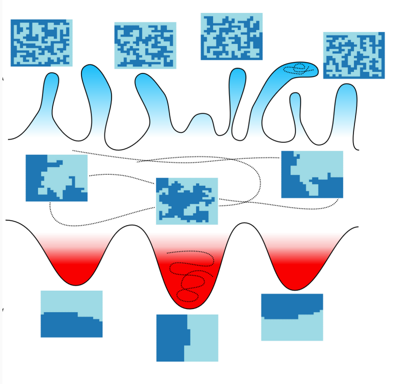
1. Choose a boundary node
2. Change its assignment





- Herschlag et al., Quantifying Gerrymandering in North Carolina, Statistics and Public Policy, 7(1), (2020). Court cases in NC and WI.
- Pegden et al. Assessing significance in a Markov chain without mixing, PNAS, (2017). Court cases in NC and PA.

Configuration Space of Partitions



L. Najt, D. DeFord, and J. Solomon, Empirical Sampling of Connected Graph Partitions for Redistricting, Physical Review E, 064130, (2021).

Theorem (Najt, D., and Solomon 2019)

Suppose that C is the class of connected planar graphs and $k \geq 2$. If there is a polynomial time algorithm to sample uniformly from:

- *the connected k-partitions of graphs in C,*
- *or the connected, 0-balanced k-partitions of graphs in C.*

then $RP = NP$.

Theorem (Najt, D., and Solomon 2019)

Suppose that C is the class of connected planar graphs and $k \geq 2$. If there is a polynomial time algorithm to sample uniformly from:

- *the connected k-partitions of graphs in C,*
- *or the connected, 0-balanced k-partitions of graphs in C.*

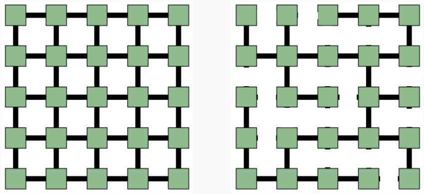
then $RP = NP$.

Theorem (Najt, D., and Solomon 2019)

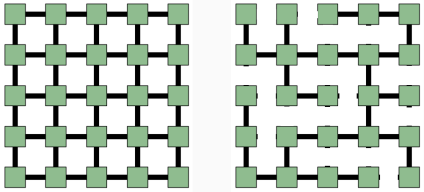
Let G be any connected graph. Then let $G^{(d)}$ be the graph obtained by replacing each edge by a doubled d-star. Then the flip walk on partitions of family of graphs $G^{(d)}_{d \geq 1}$ is slowly mixing, in the sense the Cheeger constant is decaying exponentially fast. More specifically:

$$H(\mathit{Partition\ Graph}(G^{(d)}) = O(2^{-d})$$

Proposal Method: Spanning Trees



[1] ReCombination: A family of Markov chains for redistricting, with M. Duchin and J. Solomon, Harvard Data Science Review, (2021).

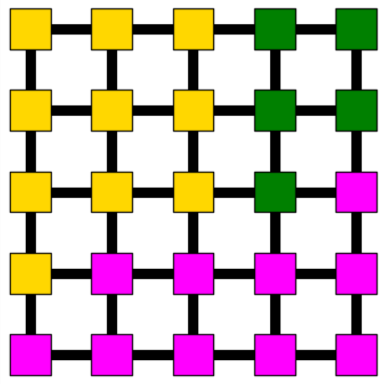


ReCombination [1]

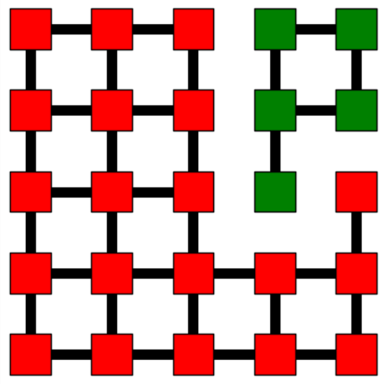
1. At each step, select two adjacent **districts**
2. Merge the subunits of those two districts
3. Draw a spanning tree for the new super–district
4. Delete an edge leaving two population balanced districts

[1] ReCombination: A family of Markov chains for redistricting, with M. Duchin and J. Solomon, Harvard Data Science Review, (2021).

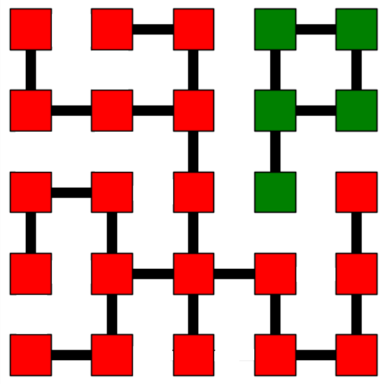
ReCombination Example



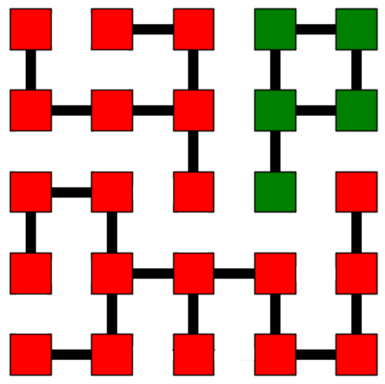
ReCombination Example



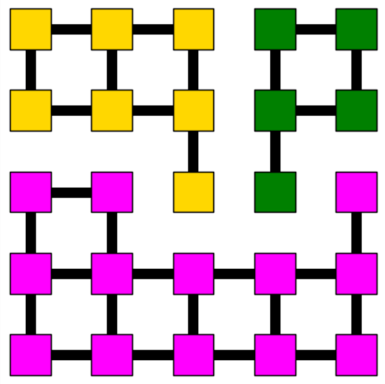
ReCombination Example



ReCombination Example



ReCombination Example



Examples

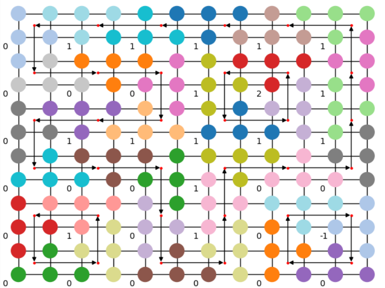
Tree Partitioning Questions

- Characterizing the distribution on partitions defined by cutting trees!
- How bad is the best cut?
- Criteria for determining when a tree is ε cuttable?
- Criteria for determining when all spanning trees of a graph are ε cuttable?
- How hard is it to find the mininum ε for which a cut exists?
- As a function of ε what proportion of spanning trees are cuttable?
- As a function of ε what proportion of edges in a given tree are cuttable?
- What is the fastest way to sample uniformly from $k-1$ balanced cut edges?

- **Tree Methods**
  - A Merge-Split Proposal for Reversible Monte Carlo Markov Chain Sampling of Redistricting Plans: D. Carter, G. Herschlag, Z. Hunter, and J. Mattingly
  - Multi-Scale Merge-Split Markov Chain Monte Carlo for Redistricting: E. Autry, D. Carter, G. Herschlag, Z. Hunter, and J. Mattingly
  - Reversible Recom: S. Cannon, M. Duchin, D. Randall and P. Rule
  - Cycle Basis Walks: D. DeFord, E. Najt, and J. Solomon
- **Other Samplers**
  - Non-reversible Markov chain Monte Carlo for sampling of districting maps, G. Herschlag: J. Mattingly, M. Sachs, and E. Wyse
  - Sequential Monte Carlo for Sampling Balanced and Compact Redistricting Plans: C. McCartan and K. Imai C. Haas, L. Hachadoorian, S. Kimbrough , P. Miller, F. Murphy

Other Sources of Rigidity

- Number of parts in the partition
- Complexity of the underlying dual graph (treewidth)
- Population tolerance
- Target distribution
- ...



Let G and H be graphs with $\pi : H \mapsto G$ a projection and M a Markov chain (proposal) on H. We define the lifted chain on G by $\hat{M} = \pi M \pi^{-1}$

$$\begin{tikzcd} H \arrow[r, "M"] & H \arrow[d, "\pi"] \\ G \arrow[u, "\pi^{-1}"] \arrow[r, "\hat{M}"] & G \end{tikzcd}$$

and if we choose H, π, and M cleverly we can guarantee both that the stationary distribution of $\hat{M}$ is equal to that of our desired walk on G and favorably bound the mixing time.

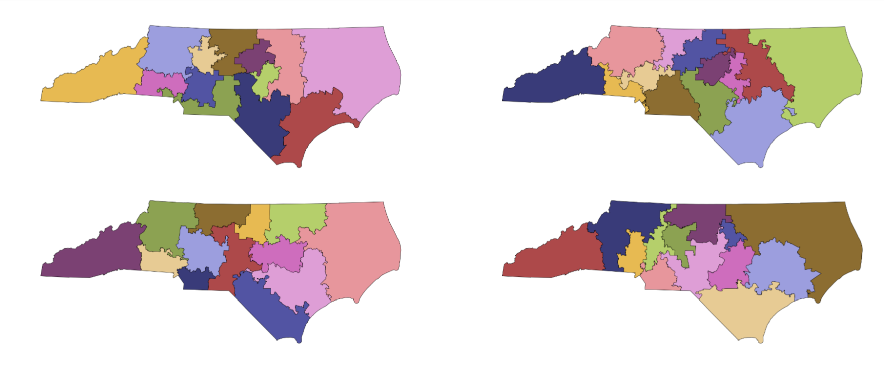
- Lots of choices!
- How to project
- How to cut
- Whether to preserve cut edges (work on marked trees)
- How many steps to take
- Whether to enforce the constraints on G or H
- ...

Interlude: Other Partitioning Problems in GerryChain

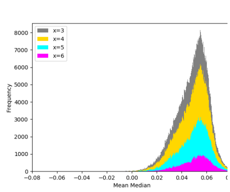
- Spin Glasses
- Height Functions and Tilings
- Compartmental Epidemiology
- Cellular Automata
- Evolutionary Game Theory
- ...

Activity: MCMC Intro Widgets and Metagraph Exploration

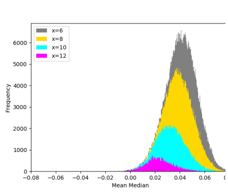
Partisan Symmetry Paradoxes



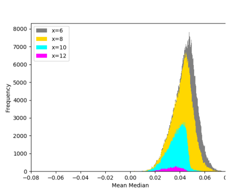
Implementing partisan symmetry: Problems and paradoxes, with N. Dhamankar, M. Duchin, V. Gupta, M. McPike, G. Schoenbach, and K. W. Sim, Political Analysis, (2021).



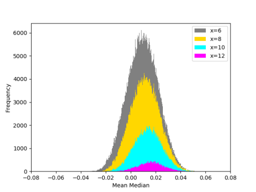
(a) UT



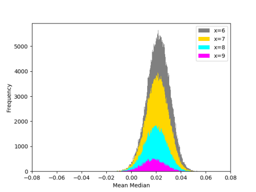
(b) GA



(c) WI



(d) VA



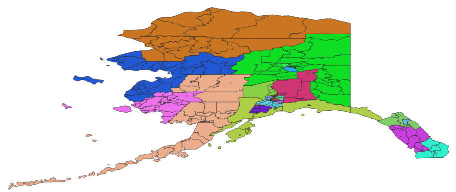
(e) MA

A Computational Approach to Measuring Vote Elasticity and Competitiveness, with M. Duchin and J. Solomon, Statistics and Public Policy, 7(1), (2020).

Colorado Counties

Colorado in Context: Congressional Redistricting and Competing Fairness Criteria in Colorado, with J. Clelland, H. Colgate, B. Malmskog, and F. Sancier-Barbosa, Journal of Computational Social Science, 2021.

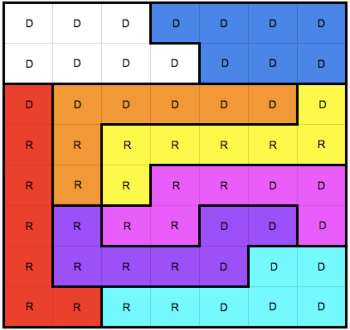
Nested Legislative Districts

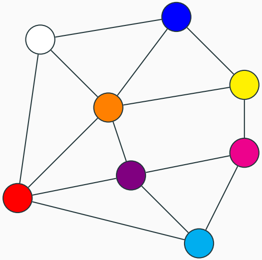


Alaska	40 House → 20 Sen	Illinois	118 House → 59 Sen
Iowa	100 House → 50 Sen	Minnesota	134 House → 67 Sen
Montana	100 House → 50 Sen	Nevada	42 House → 21 Sen
Oregon	60 House → 30 Sen	Wyoming	60 House → 30 Sen
Ohio	99 House → 33 Sen	Wisconsin	99 House → 33 Sen

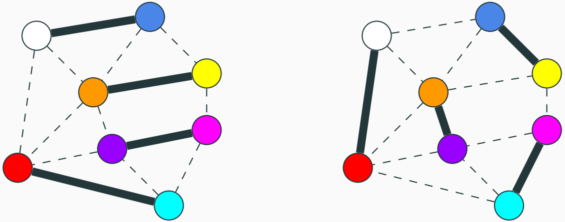
S. Caldera, D. DeFord, M. Duchin, S. Gutenkust, and C. Nix: Mathematics of Nested Districts: The case of Alaska, Statistics and Public Policy, 7(1), (2020).

Matching Example





Example Matchings



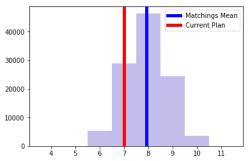
Problem Scope

	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

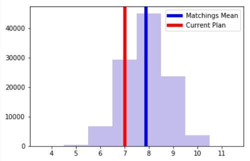
	Montana	Nevada	Oregon	Wyoming
House districts	100	42	60	60
Dual edges	269	111	158	143
Matchings	11,629,786,967,358	313,698	229,968,613	920,864

- Is the enacted matching an outlier?
    - Compute all House matchings of the enacted plan
    - Compare partisan statistics across distribution to enacted plan
- Was the House plan itself an outlier?
    - Generate 100k House plans with MCMC
    - In addition to partisan statistics, look at number of matchings
- What are the expected properties of Senate plans without matchings?
    - Generate 100k Senate plans
    - Compare partisan statistics to enacted plan

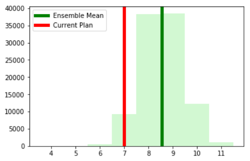
Alaska Results - Paired vs. Sampled



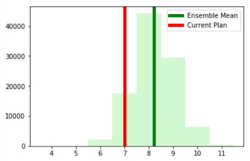
Cong18-A D Senate seats



Gov18-A D Senate seats



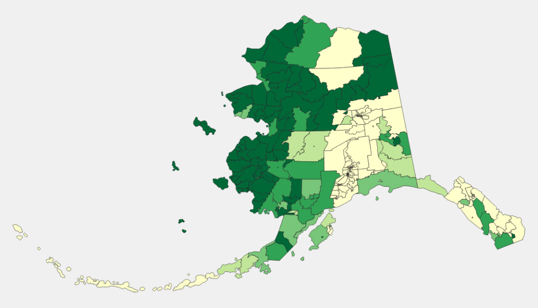
Cong18-A D Senate seats

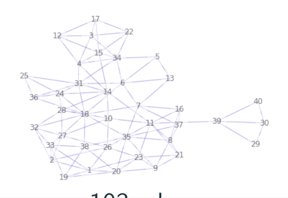


Gov18-A D Senate seats

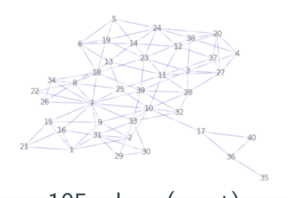
Native Majority Districts

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

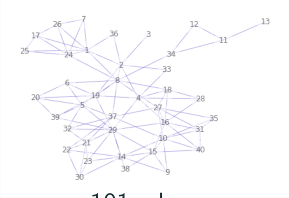




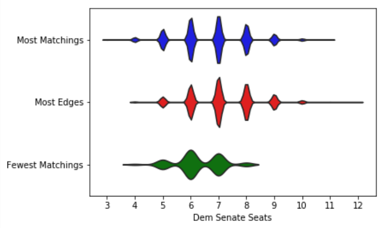
103 edges
961,176 matchings (most)

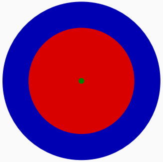


105 edges (most)
436,220 matchings

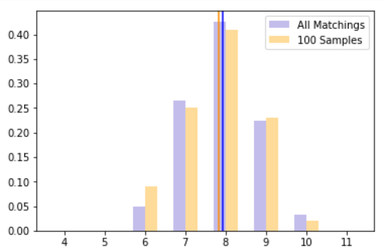


101 edges
852 matchings (fewest)

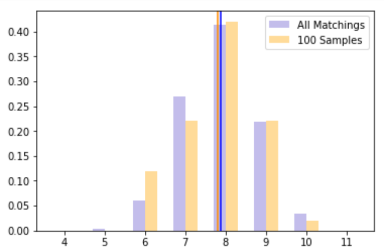




Violin plot scale



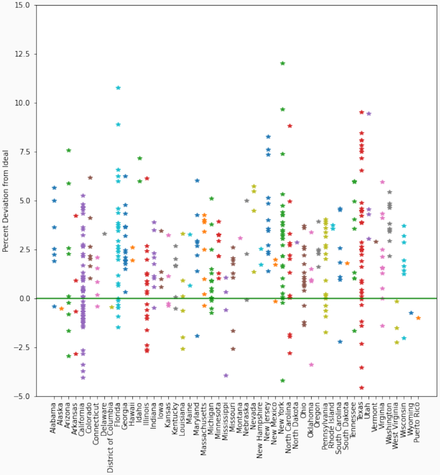
Cong18-A D Senate seats



Gov 18-A D Senate seats

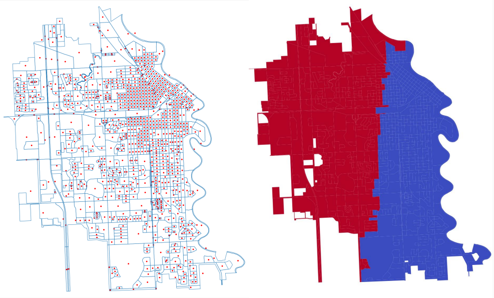
abs. error	Tight	Restricted	Permissive
Cong18-A	0.0082	0.0234	0.0868
Gov18-A	0.0203	0.0361	0.0814

Multi-Balanced Redistricting



D. DeFord, E. Kimsey, and R. Zerr: Multi-Balanced Redistricting, Preprint, (2022).

North Dakota Example

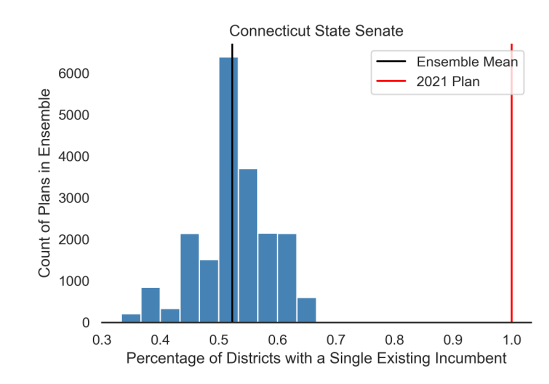


D. DeFord, E. Kimsey, and R. Zerr: Multi-Balanced Redistricting, Preprint, (2022).

New District Placement

R. Ahmed: Data Science for Social Good Case Study (2021).

Incumbency Preservation



K. Evans and R. Chang: Connecticut Redistricting Analysis, arXiv:2209.00076, (2022).

Activity: Game Theory Examples

Group Discussion

Questions?