# Èdouard Lucas:

*The theory of recurrent sequences is an inexhaustible mine which contains all the properties of numbers; by calculating the successive terms of such sequences, decomposing them into their prime factors and seeking out by experimentation the laws of appearance and reproduction of the prime numbers, one can advance in a systematic manner the study of the properties of numbers and their application to all branches of mathematics.*

# Preview

# Graphs, Geometry, and Gerrymanding

Daryl DeFord

MIT – CSAIL
Geometric Data Processing Group

Math ~~Monday~~ Friday
Kenyon College
September 19, 2019

# Outline

# Collaborators

- Prof. Moon Duchin **Tufts Math**
- Prof. Justin Solomon **MIT CSAIL**
- Lorenzo Najt **Wisconsin Math**
- VRDI Students

# Collaborators

- Prof. Moon Duchin **Tufts Math**
- Prof. Justin Solomon **MIT CSAIL**
- Lorenzo Najt **Wisconsin Math**
- VRDI Students

- *Complexity and Geometry of Sampling Connected Graph Partitions* (with L. Najt and J. Solomon), arXiv: 1908.08881.
- *ReCombination: A family of Markov chains for redistricting* (with M. Duchin and J. Solomon), preprint.
- *Competitiveness Measures for Evaluating Districting Plans* (with M. Duchin and J. Solomon), preprint.
- Redistricting Reform in Virginia: Districting Criteria in Context (with M. Duchin), Virginia Policy Review, 12(2), 120-146, (2019).

# Additional Materials

- Computational Approaches for Political Redistricting
- Interactive Notes on Discrete MCMC (with Scrabble)
- MGGG widgets
- VRDI materials

# MORAL #1:

# MORAL #1:

Computational Redistricting is NOT a solved problem!

☹

# MORAL #2:

# MORAL #2:

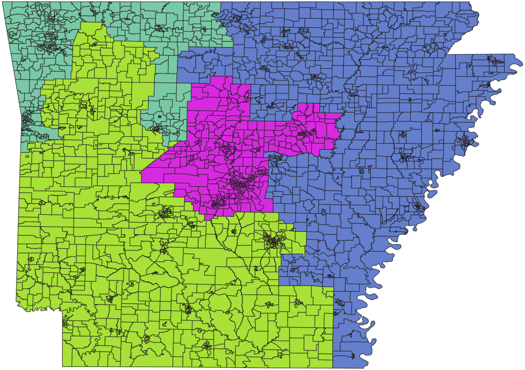
Computational Redistricting is NOT a solved problem!

# MORAL #2:

Computational Redistricting is NOT a solved problem!

☺

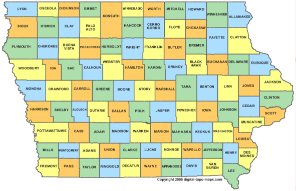
# Arkansas Congressional Districts

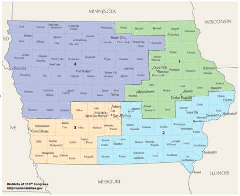


# Permissible Districting Plans

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

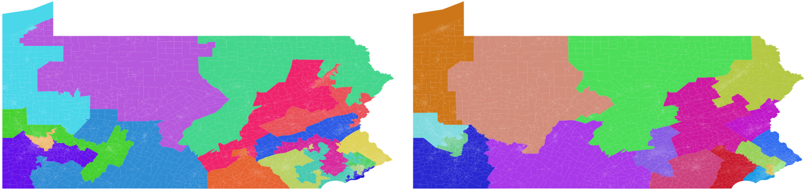
# Example: Iowa



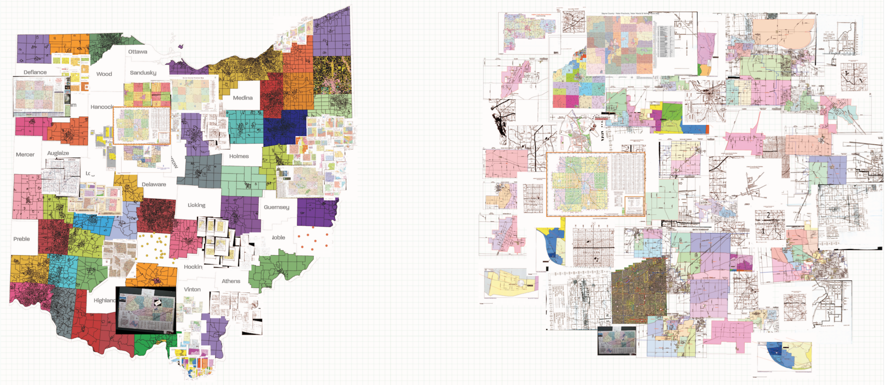

- 4 Congressional Districts, 100 House Districts, 50 Senate Districts
- House districts nest into Senate districts
- Congressional districts made out of counties
- Independent committee with legislative approval
- No partisan data allowed

# Example: Pennsylvania



- 18 Congressional Districts, 203 House Districts, 50 Senate Districts
- Zero–balanced population
- Legislature draws congressional districts - subcommittee draws legislative districts
- Partisan considerations allowed

# Data Availability

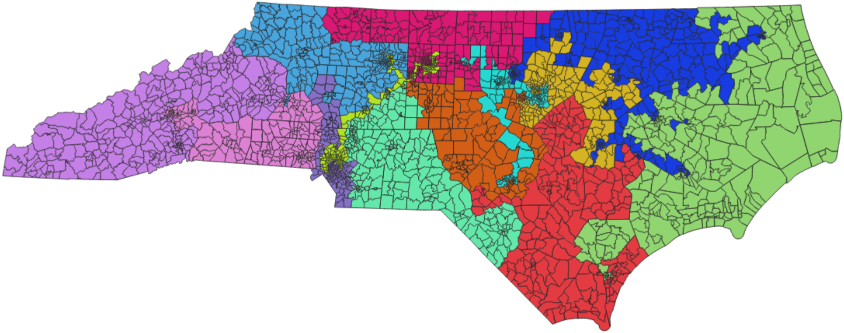


# Why analyze?

- Court cases
  - Detecting gerrymandering
  - Evaluating proposed remedies
- Reform Efforts
  - Establishing baselines
  - Potential impacts of new rules
- Commissions and plan evaluation
  - Unintentional gerrymandering
  - Full space of plans

# Gerrymandering

# Ugly Shapes



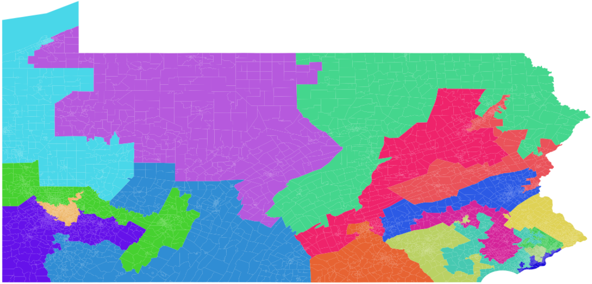
# Ugly Shapes

NC12 #1

NC12 #2

NC12 #12

# Ugly Shapes



# Measurement Problems

**Theorem (Bar-Natan, Najt, and Schutzman 2019[1])**

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

**Problem (Barnes and Solomon 2018[2])**

*Geographic Compactness scores can be distorted by:*

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

[1] The Gerrymandering Jumble: Map Projections Permute Districts' Compactness Scores, arXiv:1905.03173

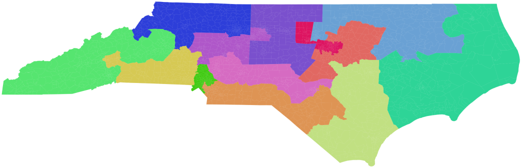
[2] Gerrymandering and Compactness: Implementation Flexibility and Abuse, arXiv:1803.02857

# Isoperimetric Profiles

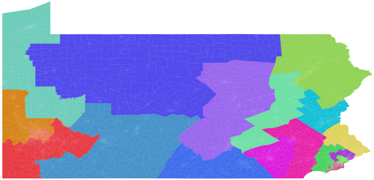
Total Variation Isoperimetric Profiles (with H. Lavenant, Z. Schutzman, and J. Solomon),

SIAM Journal on Applied Algebra and Geometry, to appear (2019).

# Partisan Imbalance



NC16

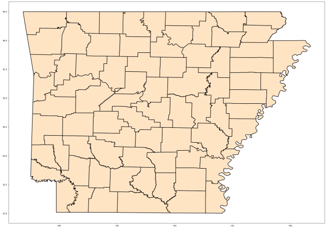


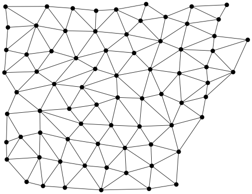
PA TS-Proposed

# MORAL:

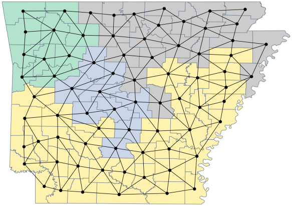
Computational Redistricting is **NOT** a solved problem!

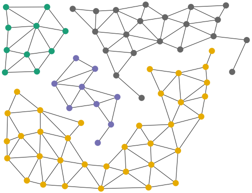
# Discrete Partitioning



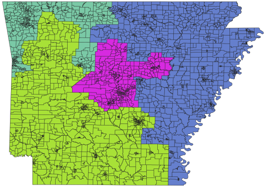


# Discrete Partitioning

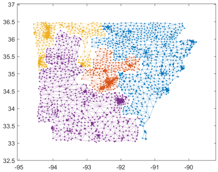




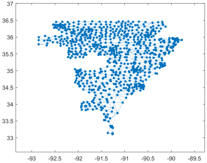
# Arkansas Congressional Districts



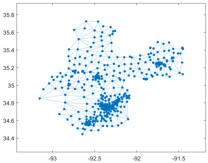
Geography



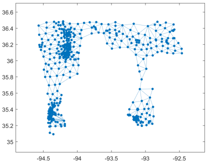

Dual Graph



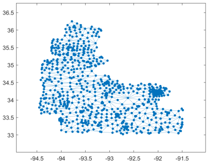

District # 1




District # 2




District # 3




District # 4

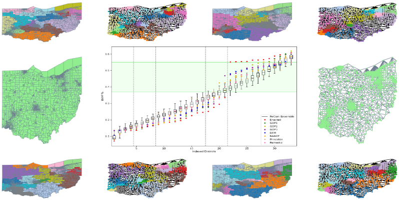
# Permissible Districting Plans

We want to partition a given geography (graph), at a given scale, into $k$ pieces, satisfying some constraints:

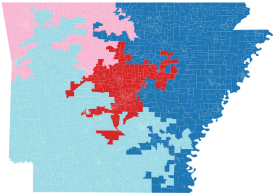
- **Contiguity**
- **Population Balance**
- Compactness
- Communities of Interest
- Municipal Boundaries
- Competitiveness/Responsiveness
- Incumbency Protection
- ...

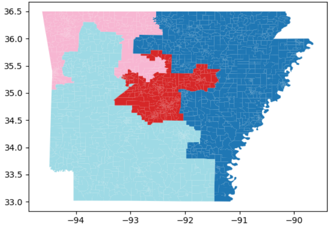
# Ensemble Analysis

- The wide variety in rules applied to districting problems (even in the same state) means that any single measure of gerrymandering will be insufficient/exploitable
- Instead we want to compare to large ensembles of other feasible plans.
- This allows us to understand the impacts of the underlying political and demographic geography on a wide collection of metrics.



# Which ensembles?





# Ensembles in Practice

- The appeal of an ensemble method is that you get to control the input data very carefully
- However, just because a particular type of data was not considered doesn't mean that the outcome is necessarily "fair"
- There are lots of "random" methods for constructing districting plans
- Most don't offer any control over the distribution that you are drawing from

# Markov Chain Monte Carlo

- A methodology for sampling (and evaluating) complex state spaces
- Developed during the Manhattan project for statistical physics computations
- One of the "top 10" algorithms of the 20th century

Introduction to Discrete MCMC for Redistricting

Daryl DeFord

June 16, 2019

**Contents**

# Geometric Probability

**Question**

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

# Geometric Probability

## Question

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

## Answer

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

# Geometric Probability

### Question

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

### Answer

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

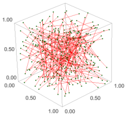
### Question

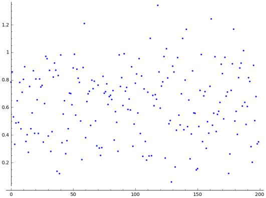
*What is the expected distance between two random points on* $[0,1]^n$*?*

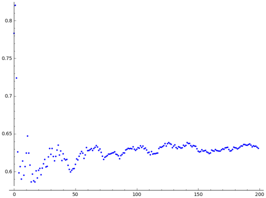
### Answer

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

# Random Points in the Cube



# 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

# 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

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.

# What is a Markov chain?

**Definition (Markov Chain)**

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

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

# 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 $P = [p_{i,j}]$.

# Ant on a Keyboard

# 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 $\pi$ 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 $\pi$. Even better, if $f$ is any function defined on the state space $y_1, y_2, \ldots, y_m$ are samples from $\pi$ 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 $\pi$ is the distribution we are trying to sample from.

# Proportional to a distribution!?!

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.

# How does it work?

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

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

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

# 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 $\beta$ uniformly on $[0, 1]$
4. Set

$$X_{k+1} = \begin{cases} \hat{y} & \text{if } \beta < \alpha \\ y & \text{otherwise.} \end{cases}$$

# MCMC on 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

# MCMC on 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?

# MCMC on 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

# MCMC on partitions

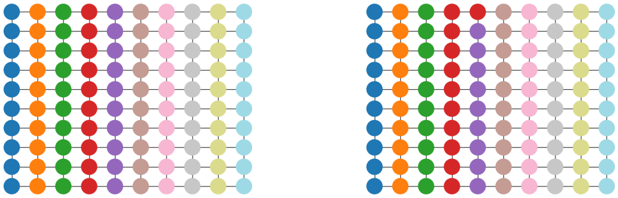
1. Set constraints to define the state space
2. Start with an initial plan
3. **Propose a modification**
4. Verify that the modification satisfies the constraints
5. Accept using MH criterion
6. Repeat

## Why?

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

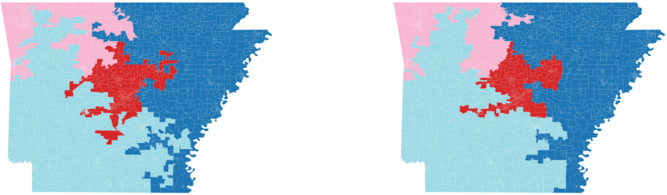
# Single Edge Flip Proposals

1. Uniformly choose a cut edge
2. Change one of the incident node assignments to the other

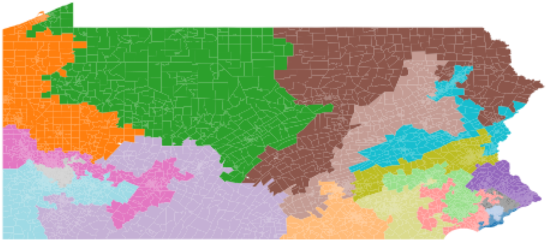


- Mattingly et al. (2017, 2018) Court cases in NC and WI.
- Pegden et al. Assessing significance in a Markov chain without mixing, PNAS, (2017). Court case in PA.

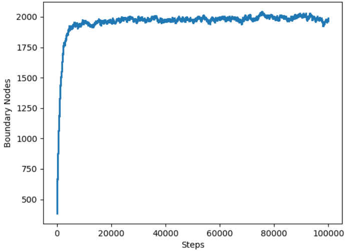
# Single Edge Ensembles

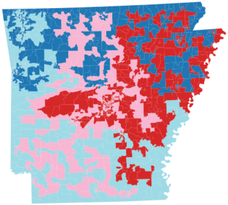


# PA Single Edge Flip

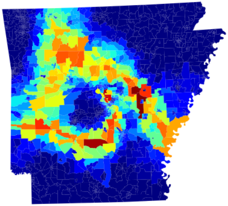


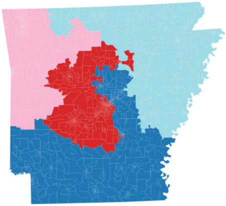
# Unconstrained Flip



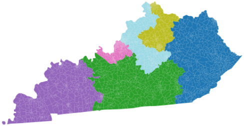



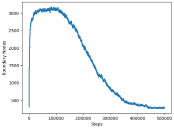
# Constrained Flip

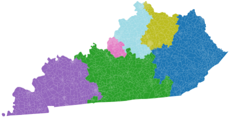




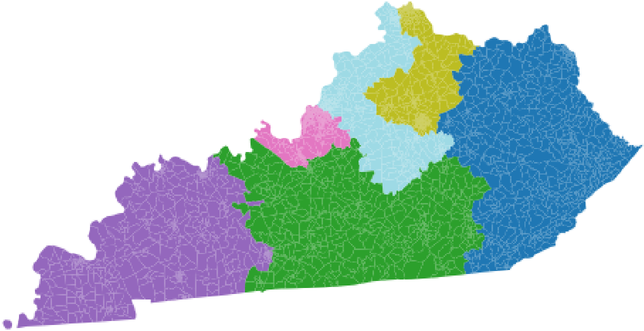
# Annealing



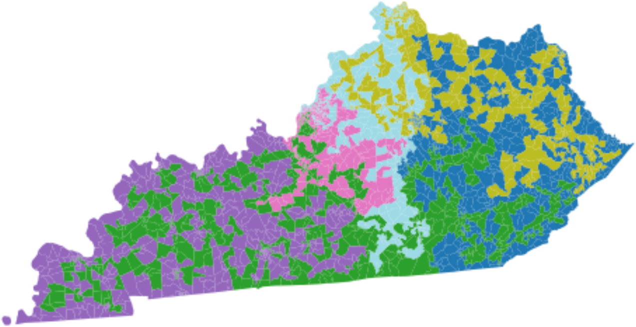

Boundary Nodes
Steps
3000
2500
2000
1500
1000
500
0
100000
200000
300000
400000
500000



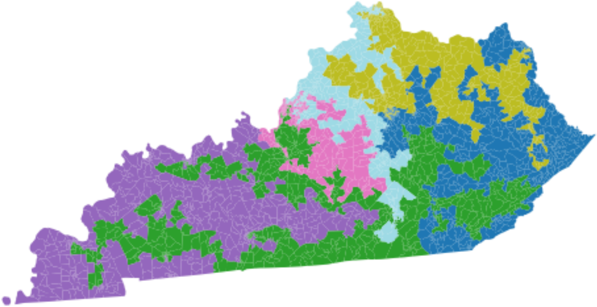
# Annealing



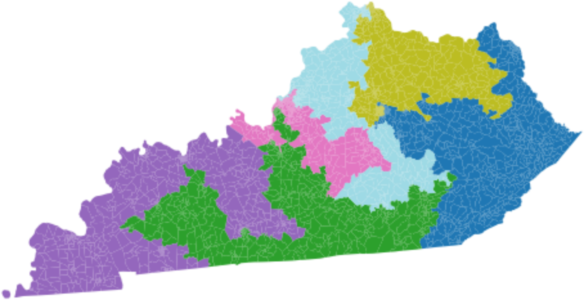
# Annealing



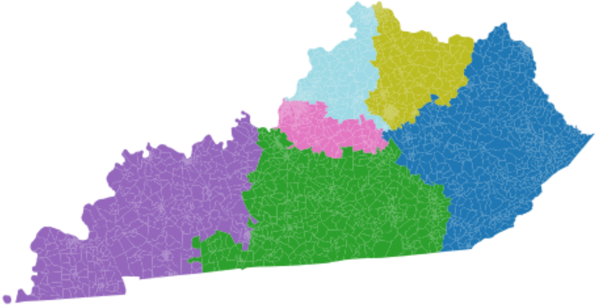
# Annealing



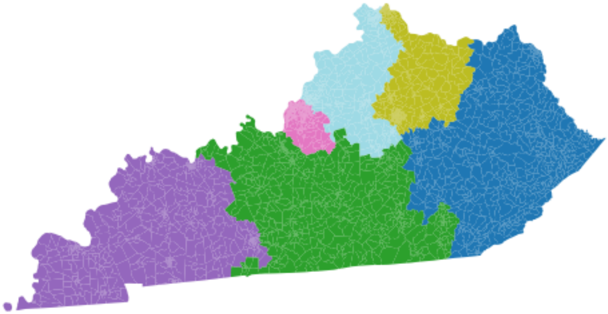
# Annealing



# Annealing



# Annealing



# Uniform Sampling of Contiguous Partitions

**Theorem (Najt, D., and Solomon 2019)**

*Suppose that $\mathscr{C}$ 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 $\mathscr{C}$,*
- *or the connected, 0-balanced k-partitions of graphs in $\mathscr{C}$.*

*then $RP = NP$.*

# Uniform Sampling of Contiguous Partitions

**Theorem (Najt, D., and Solomon 2019)**

*Suppose that $\mathscr{C}$ 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 $\mathscr{C}$,*
- *or the connected, 0-balanced k-partitions of graphs in $\mathscr{C}$.*

*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:*

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

# Proof Outline Sketch

**Following technique of Jerrum, Valiant, and Vazirani[1].**

1. Show that uniformly sampling simple cycles is hard on some class $\mathscr{C}$
   1. Choose a gadget that respects $\mathscr{C}$ and allows us to concentrate probability on long cycles
   2. Count the proportion of cycles as a function of length
   3. Reduce to Hamiltonian path on the graph class
2. Show closure of class under planar dual
3. Identify partitions with cut edges $\mapsto$ simple cycles (via planar duality)
4. Conclude that sampling partitions would allow you to sample from cycles which would allow you to find Hamiltonian cycles

□

[1] M. Jerrum, L. Valiant, and V. Vazirani, Random generation of combinatorial structures from a uniform distribution, Theoretical Computer Science, 43 (1986), 169–188.

# Proof Sketch – Planar 2–Partitions

**Still following technique of Jerrum, Valiant, and Vazirani.**

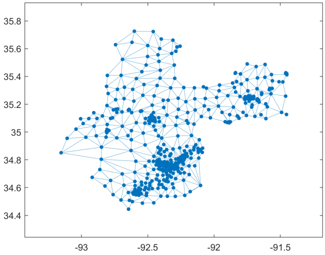
1. Let $\mathscr{C}$ be the planar connected graphs
   1. Replace the edges with chains of dipoles
   2. Hamiltonian hardness for $\mathscr{C}$ given by [1]
2. $\mathscr{C}$ closed under planar duals
3. Identify partitions with cut edges (via planar duality)

$\square$

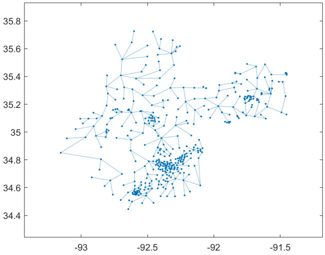
[1] M. Garey, D. Johnson, and R. Tarjan, The Planar Hamiltonian Circuit Problem is NP-Complete, SIAM Journal on Computing, 5, (1976), 704–714.

# Slow Mixing Example

# Tree based methods

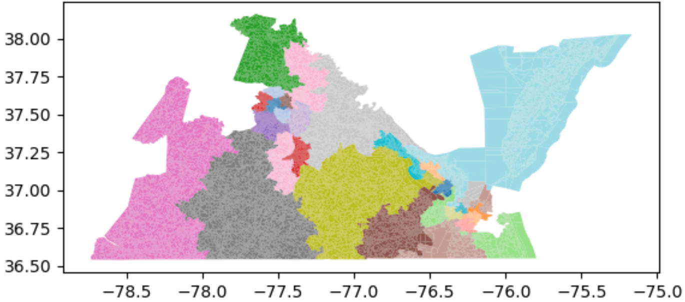



District




Spanning Tree

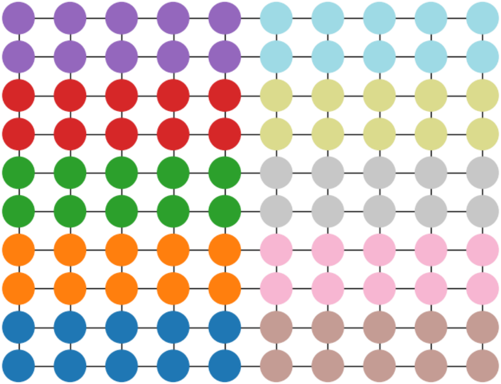
# Tree Seeds Ensemble



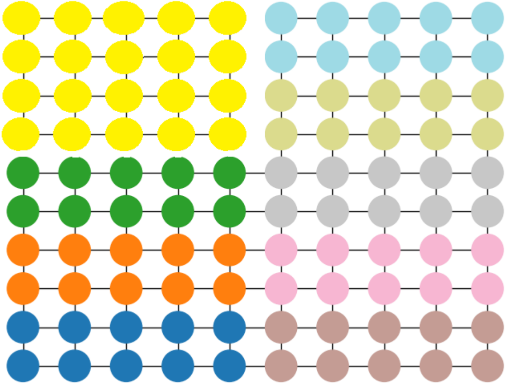
# Recombination Steps

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
5. Repeat

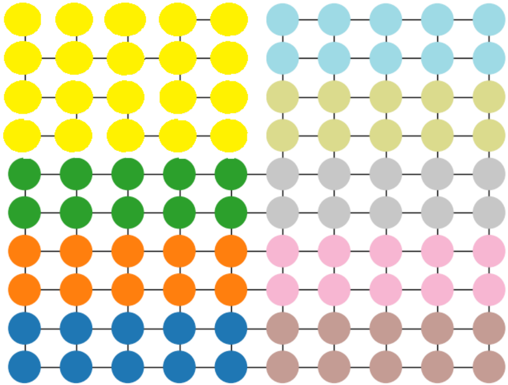
# Recombination Step Example



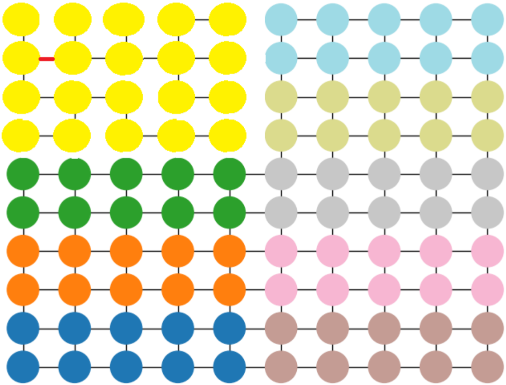
# Recombination Step Example



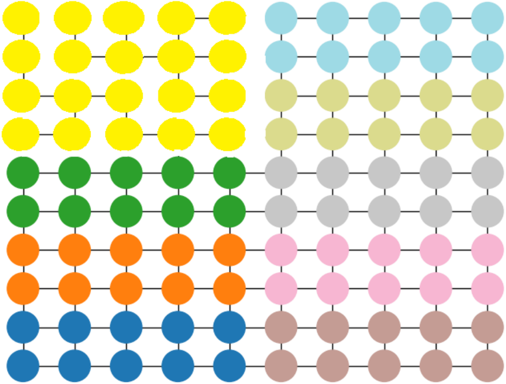
# Recombination Step Example



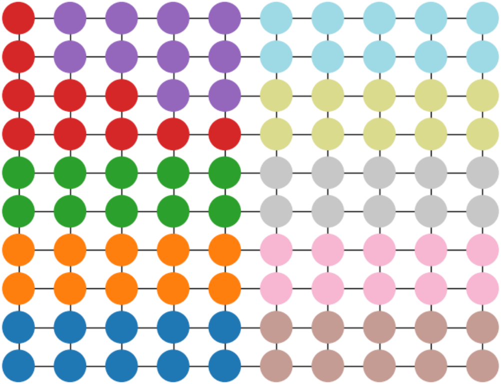
# Recombination Step Example



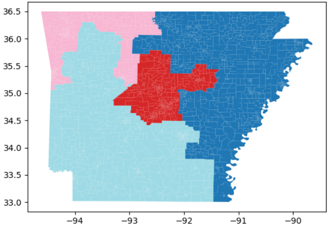
# Recombination Step Example

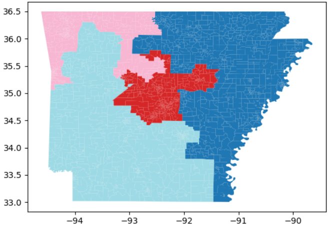


# Recombination Step Example



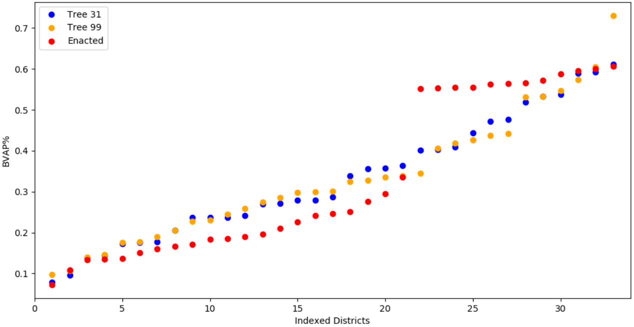
# AR Ensembles



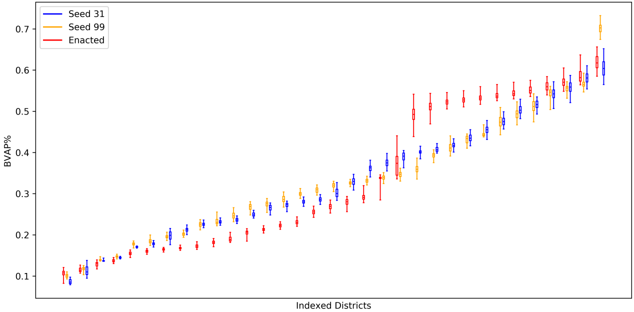
# PA Recombination Steps

# Initial Seeds



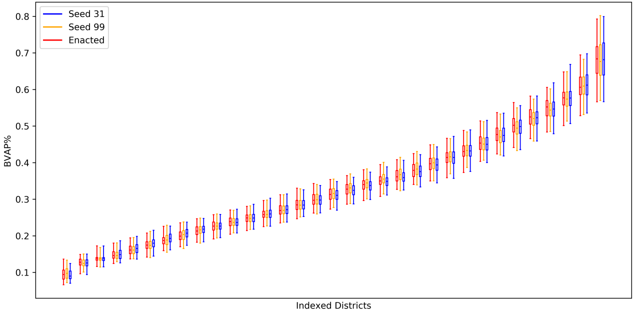

Initial

# Boundary Flip Mixing – Seeds



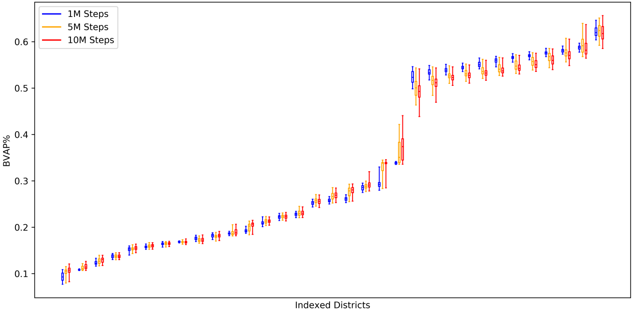

10,000,000 Flip Steps

# Recombination Mixing – Seeds

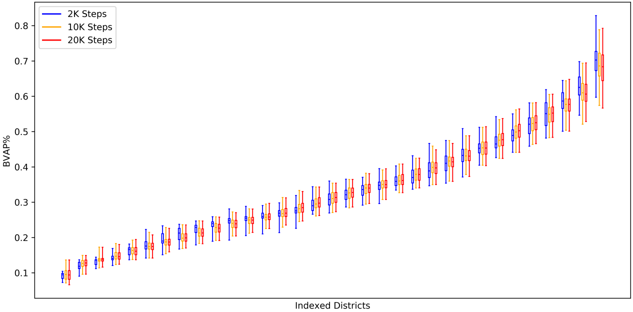



20,000 Recombination Steps

# Boundary Flip Mixing – Length



# Recombination Mixing – Length




20,000 Recombination Steps

# Applications

**Nos. 18-422, 18-726**

IN THE
**Supreme Court of the United States**

ROBERT A. RUCHO, ET AL.,
*Appellants,*
v.
COMMON CAUSE, ET AL.,
*Appellees.*

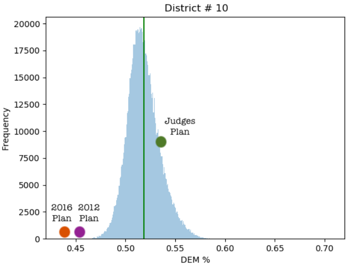
*On Appeal from the United States District Court for the Middle District of North Carolina*

LINDA H. LAMONE, ET AL.,
*Appellants,*
v.
O. JOHN BENISEK, ET AL.,
*Appellees.*

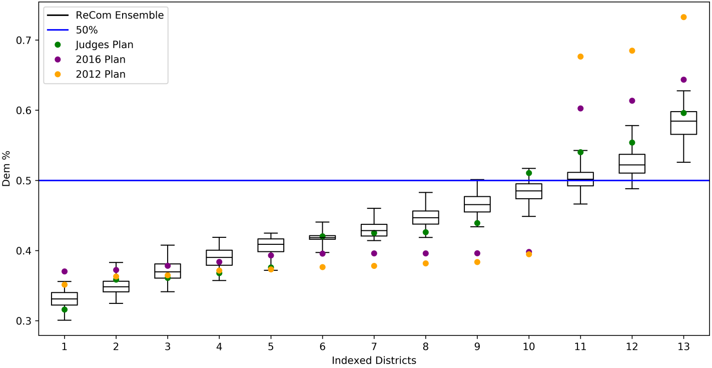
*On Appeal from the United States District Court for the District of Maryland*

**AMICUS BRIEF OF MATHEMATICIANS, LAW PROFESSORS, AND STUDENTS IN SUPPORT OF APPELLEES AND AFFIRMANCE**

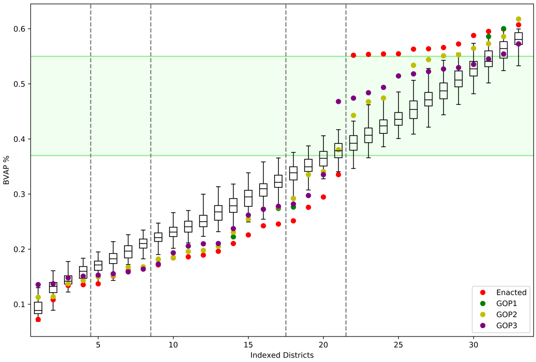
# Outlier Example: NC



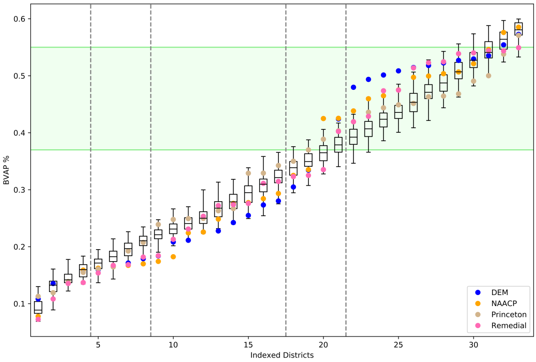
# Outlier Example: NC



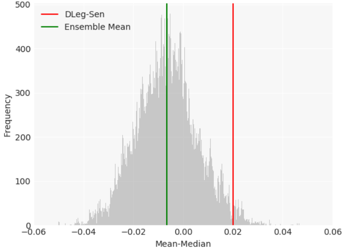
# Outlier Example: VA



# Outlier Example: VA



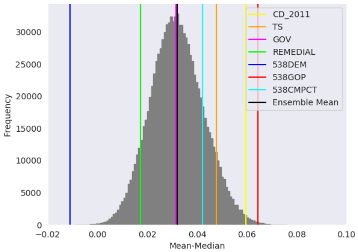
# Baseline Example: VA



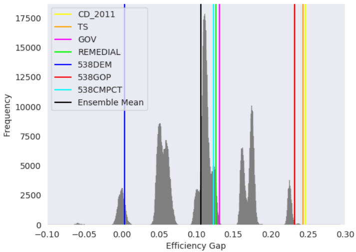

Mean–Median

Efficiency Gap

# Baseline Example: PA

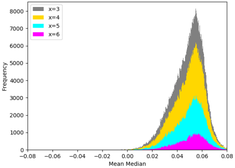



Mean–Median

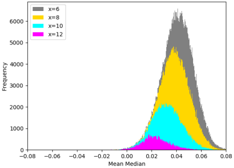



Efficiency Gap

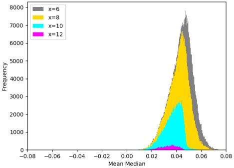
# Reform Example: Competitiveness



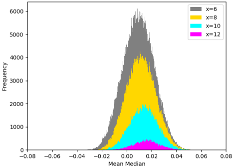

UT



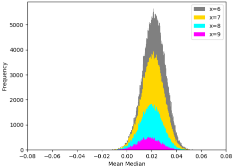

GA




WI




VA




MA

# MORAL:

Computational Redistricting is NOT a solved problem!

# Try it at home!

- Draw your own districts with **Districtr**
  - `https://districtr.org`
  - Easy to generate complete districting plans in browser or on a tablet
  - Measures district demographics and expected partisan performance
  - Identifies communities of interest
- Generate your own ensembles with **GerryChain**[1]
  - `https://github.com/mggg/gerrychain`
  - Flexible, modular software for sampling graph partitions
  - Handles the geodata processing as well as the MCMC sampling
  - Current support for a
  - Successfully applied in VA, NC, PA, etc.
- Data is available for your favorite state!
  - Census dual graphs with demographic information:
  - `https://people.csail.mit.edu/ddeford/dual_graphs`
  - Precincts with electoral results
  - `https://github.com/mggg-states`

[1] Originally RunDMCMC

# The End

Thanks!

# General Tree Proposals

1. Form the induced subgraph on the complement of the cut edges
2. Add some subset of the cut edges
3. Uniformly select a maximal spanning forest
4. Apply a Markov chain on trees
5. Partition the spanning forest into $k$ population balanced pieces

# Special Cases

- Uniform Trees: Add all cut edges
- $k$–edges: Uniformly add $k$ cut edges
- Recombination: Add all cut edges between one pair of districts.
- Super-Recombination: Take a maximal matching on the dual graph to the districts and add all cut edges between matched districts.
- Bounce Walk: Add a single cut edge between enough pairs of districts to make a tree in the dual graph of districts.

# Special Cases

- Uniform Trees: Add all cut edges
- $k$–edges: Uniformly add $k$ cut edges
- Recombination: Add all cut edges between one pair of districts.
- Super-Recombination: Take a maximal matching on the dual graph to the districts and add all cut edges between matched districts.
- Bounce Walk: Add a single cut edge between enough pairs of districts to make a tree in the dual graph of districts.

**Question**

*What are the steady state distributions (and mixing times) of these walks?*

# Tree Partitioning Questions

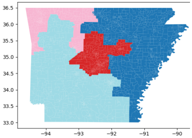
- Characterizing the distribution on partitions defined by cutting trees!
- How bad is the best cut?
- Criteria for determining when a tree is $\varepsilon$ cuttable?
- Criteria for determining when all spanning trees of a graph are $\varepsilon$ cuttable?
- How hard is it to find the mininum $\varepsilon$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

# General Merge Proposals

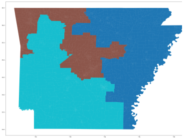
1. At each step, select two adjacent **districts**
2. Merge the subunits of those two districts
3. Bipartition the new super–district
4. Repeat
5. (Optional) Mix with single edge flips

# General Merge Proposals

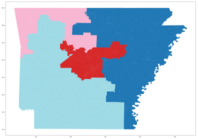
1. At each step, select two adjacent **districts**
2. Merge the subunits of those two districts
3. Bipartition the new super–district
4. Repeat
5. (Optional) Mix with single edge flips




Before



During



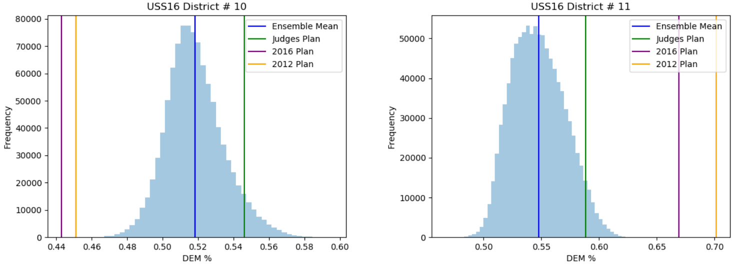
After

# Bipartitioning Methods

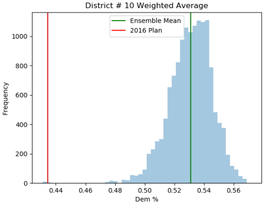
- Trees!
- Flood Fills
- Path Fills
- Agglomerative/Hierarchical
- Spectral
- Min Cut

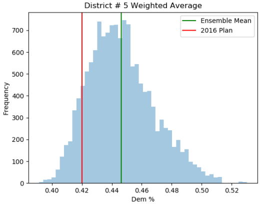
More details (and colorful figures) at:
`https://www.overleaf.com/read/zpmyzqmpvmnx`

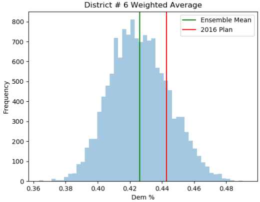
# Ensemble Example: NC

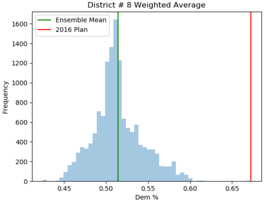


# Ensemble Example: NC

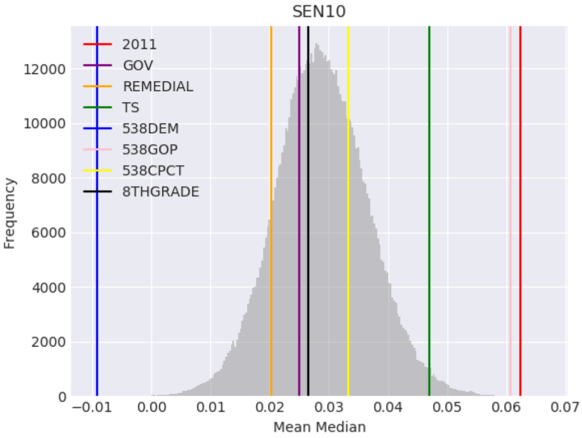


# Ensemble Example: PA



# Ensemble Example: PA