Sampling and Graph Problems in Political Redistricting...
...or, what I did last summer

Austin Eide

September 18, 2018

Outline

1. Background
   - A brief intro to gerrymandering
   - Greatest Hits
   - The Takeaway
2. Graph Framework
   - The Dual Graph
3. Markov Stuff
   - (Reversible) Markov chains
   - Chikina, Frieze, Pegden $\sqrt{\varepsilon}$-test
   - Pennsylvania 2018
4. Sampling and Trees
   - Sampling Partitions
   - Tree Compactness

A brief intro to gerrymandering

A brief intro to gerrymandering

- U.S. Census taken every 10 years (with the next coming in 2020)

A brief intro to gerrymandering

- U.S. Census taken every 10 years (with the next coming in 2020)
- The 435 U.S. Reps are then (re)apportioned

A brief intro to gerrymandering

- U.S. Census taken every 10 years (with the next coming in 2020)
- The 435 U.S. Reps are then (re)apportioned
- States required to redistrict after each Census

A brief intro to gerrymandering

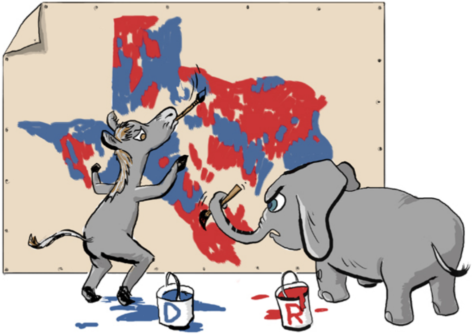
- U.S. Census taken every 10 years (with the next coming in 2020)
- The 435 U.S. Reps are then (re)apportioned
- States required to redistrict after each Census
- ...

A brief intro to gerrymandering

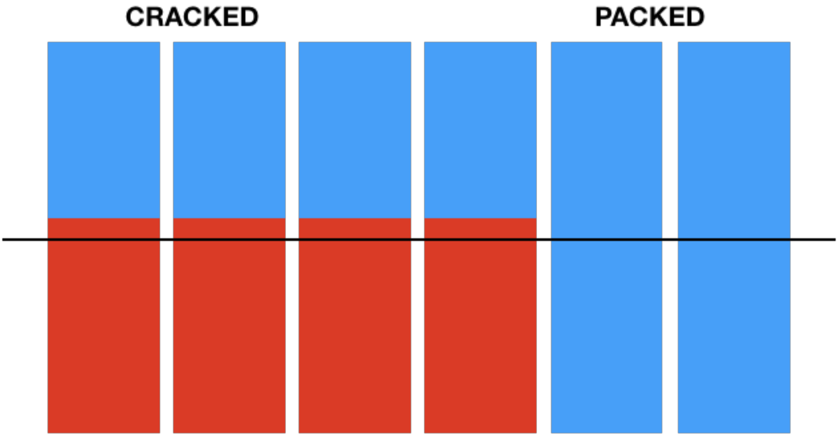
- U.S. Census taken every 10 years (with the next coming in 2020)
- The 435 U.S. Reps are then (re)apportioned
- States required to redistrict after each Census
- ...
- Lawsuits ensue

A brief intro to gerrymandering

- U.S. Census taken every 10 years (with the next coming in 2020)
- The 435 U.S. Reps are then (re)apportioned
- States required to redistrict after each Census
- ...
- Lawsuits ensue



How to gerrymander: cracking and packing



Red has ~ 1/3 of the vote share, but wins 2/3 of the seats

The VRA and racial gerrymandering

The VRA and racial gerrymandering

- Passed in 1965, the Voting Rights Act prohibits districting practices which "dilute" a racial minority group's voting power

The VRA and racial gerrymandering

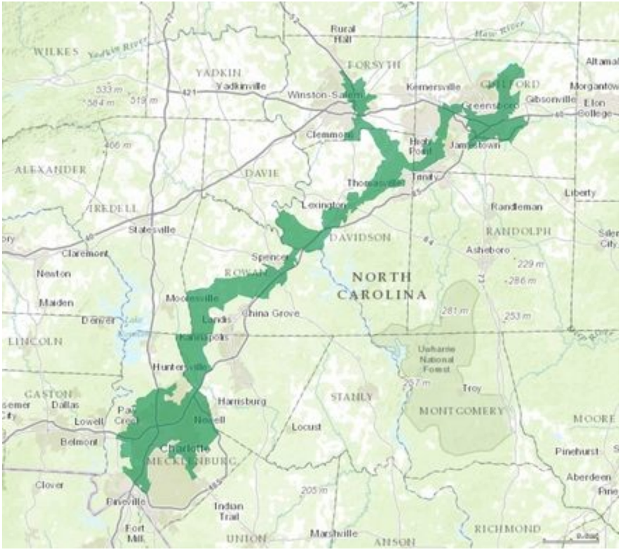
- Passed in 1965, the Voting Rights Act prohibits districting practices which "dilute" a racial minority group's voting power
- Thornburg v. Gingles (1986) established criteria for creating "majority-minority" districts

The VRA and racial gerrymandering

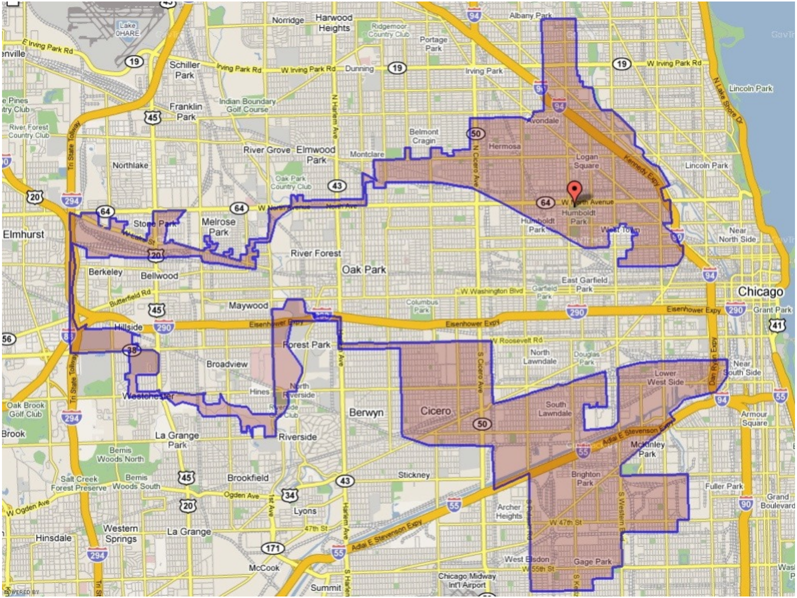
- Passed in 1965, the Voting Rights Act prohibits districting practices which "dilute" a racial minority group's voting power
- Thornburg v. Gingles (1986) established criteria for creating "majority-minority" districts
  1. Minority group must be "sufficiently numerous and compact" to make up a district
  2. Group members must vote similarly
  3. Majority group must vote as a bloc, generally in opposition to the minority group

Greatest Hits...

North Carolina 12



Illinois 4



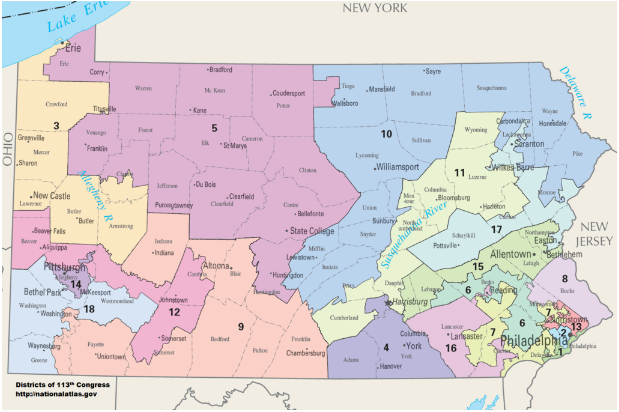
Norridge
Harwood Heights
Albany Park
Schiller Park
Portage Park
Irving Park
Lincoln Park
Franklin Park
Indian Boundary Golf Course
Belmont Cragin
Hermosa
Avondale
Logan Square
River Grove
Elmwood Park
Montclare
Northlake
Humboldt Park
West Town
Near North Side
Elmhurst
Stone Park
Melrose Park
River Forest
Oak Park
Berkeley
Bellwood
Maywood
East Garfield Park
Chicago
Hillside
Forest Park
Broadview
North Lawndale
Douglas Park
Near South Side
Lower West Side
Westchester
Hines
North Riverside
Berwyn
Cicero
South Lawndale
Armour Square
La Grange Park
Riverside
McKinley Park
Brighton Park
Bemis Woods North
Bemis Woods South
Salt Creek Forest Preserve
Brookfield
Stickney
Fuller Park
Grand Boulevard
Archer Heights
Lyons
Hinsdale
Western Springs
La Grange
West Elsdon
Gage Park
McCook
Summit
Chicago Midway Int'l Airport
Eisenhower Expy
Adlai E Stevenson Expy
Kennedy Expy
W Irving Park Rd
W North Avenue
W Roosevelt Rd
W Washington Blvd
W 47th St
Ogden Ave

Wisconsin State Assembly, 2012

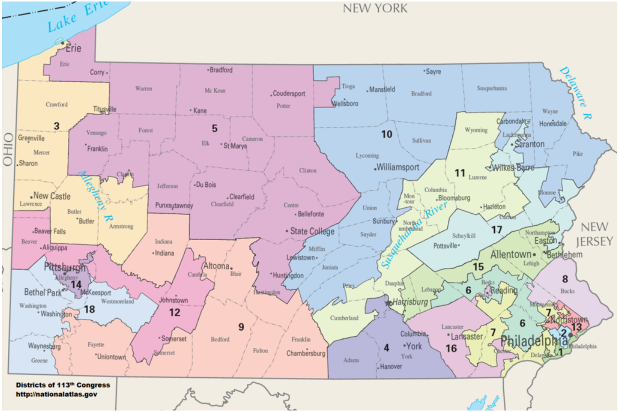
Wisconsin State Assembly, 2012

R's get 48% of the votes, 60/99 seats in 2012 under the above plan

Pennsylvania U.S. House, 2016

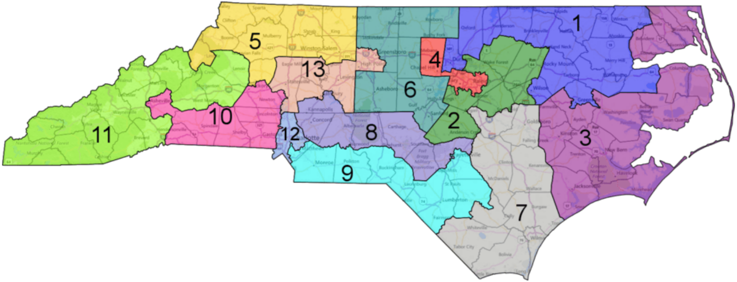


Pennsylvania U.S. House, 2016

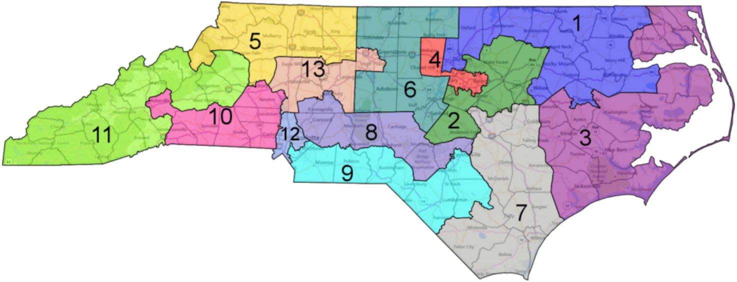


R's get 54% of the vote, 13/18 seats in 2016

North Carolina U.S. House, 2016



North Carolina U.S. House, 2016



R's get 53% of the vote, 10/13 seats in 2016

The Takeaway

The Takeaway

- Poorly shaped districts often indicate a gerrymander, but effective gerrymanders certainly take place without them

The Takeaway

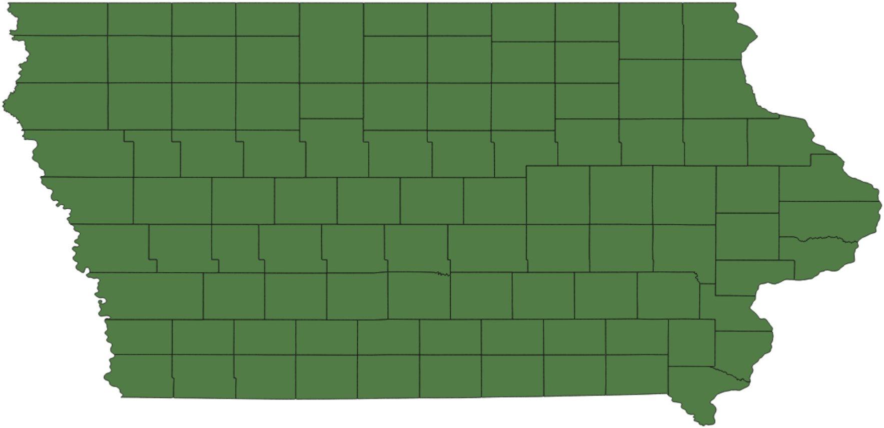
- Poorly shaped districts often indicate a gerrymander, but effective gerrymanders certainly take place without them
- In the absence of shape infractions, how can we determine with confidence that a given plan is a gerrymander?

The Takeaway

- Poorly shaped districts often indicate a gerrymander, but effective gerrymanders certainly take place without them
- In the absence of shape infractions, how can we determine with confidence that a given plan is a gerrymander?
  - Bad numbers aren't enough; must consider viable alternatives

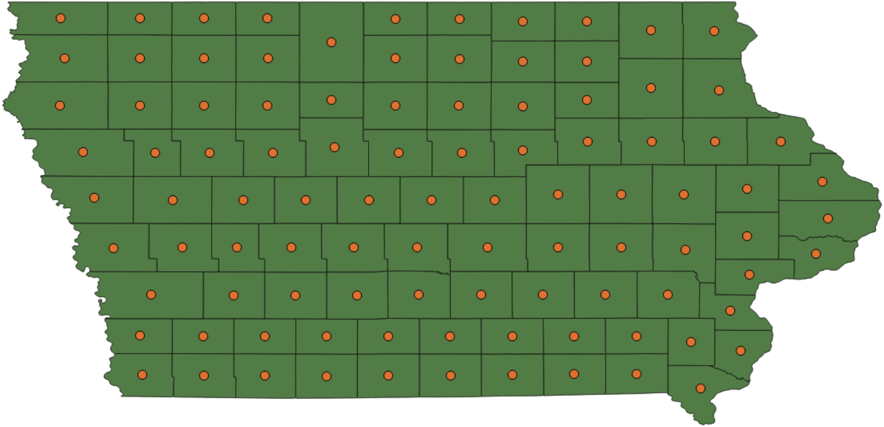
The "dual graph"

States are carved up into atomic units (often, Voting Tabulation Districts, or VTDs), which are agglomerated into congressional districts



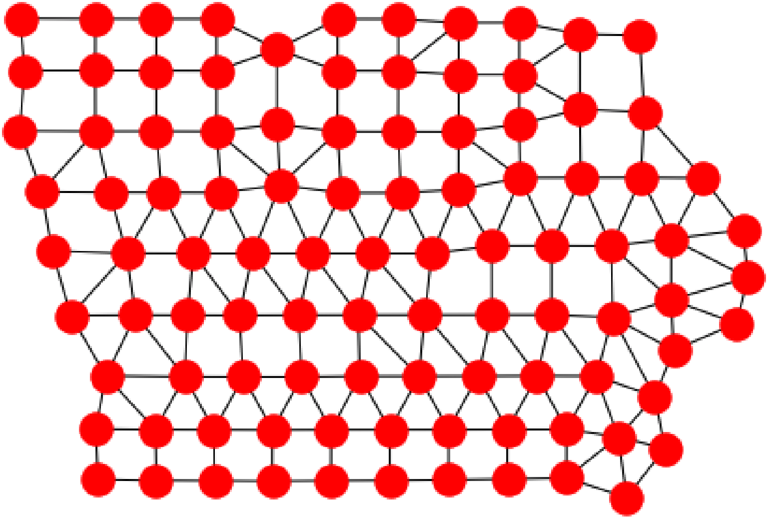
The "dual graph"

Treat VTDs as faces of a planar graph, and take the dual



The "dual graph"

Treat VTDs as faces of a planar graph, and take the dual



Benefits of the dual graph

Benefits of the dual graph

- Easy to work with in Python, etc.

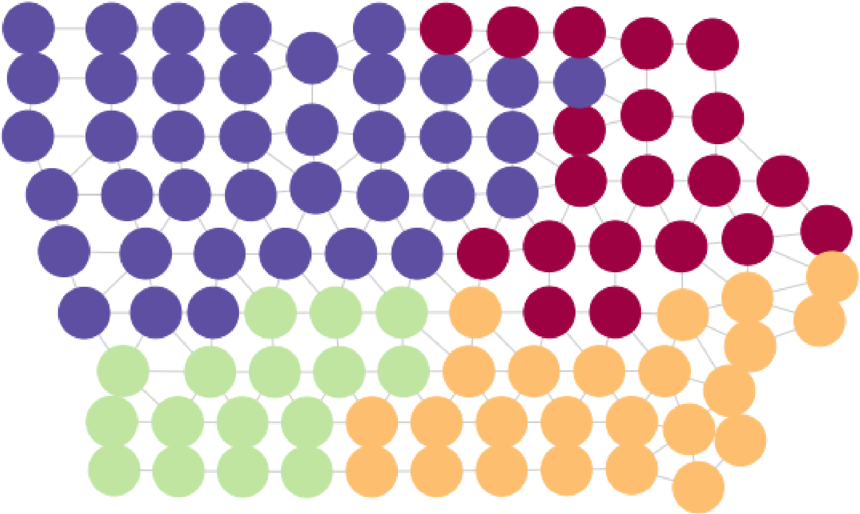
Benefits of the dual graph

- Easy to work with in Python, etc.
- Allows for measures of *discrete compactness*

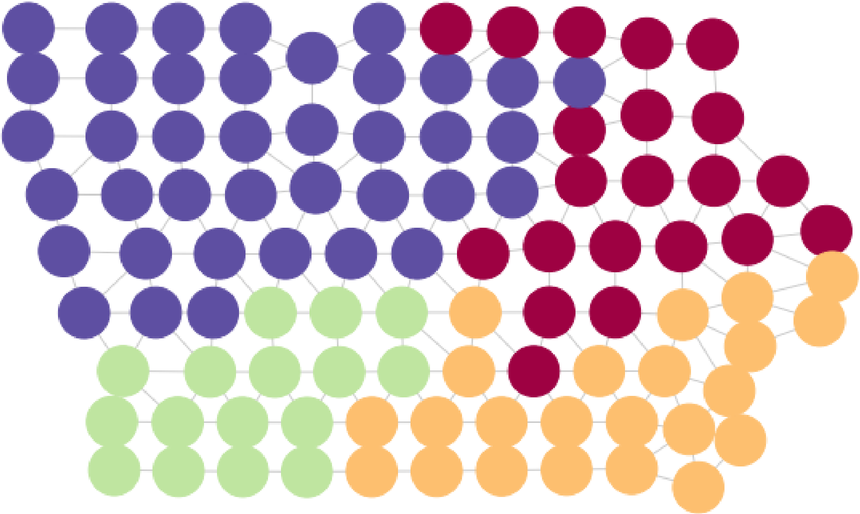
Benefits of the dual graph

- Easy to work with in Python, etc.
- Allows for measures of *discrete compactness*
- Districting plans are (connected) partitions of the dual graph

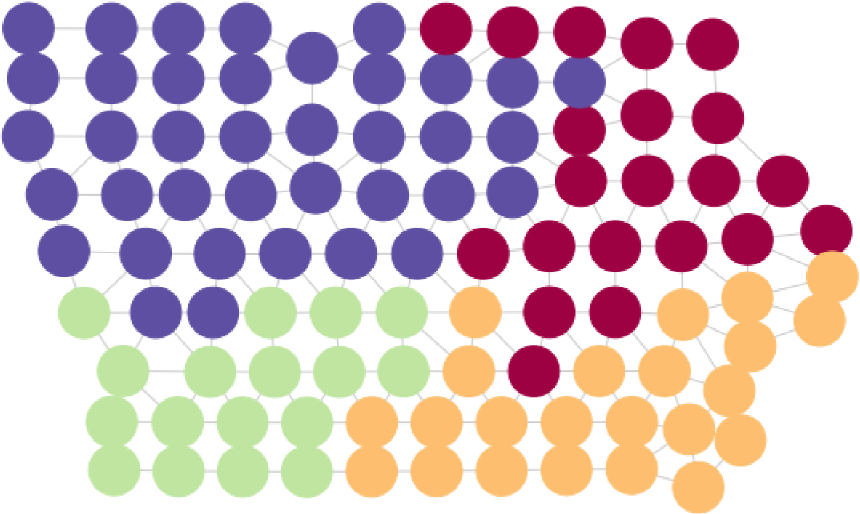
For example...



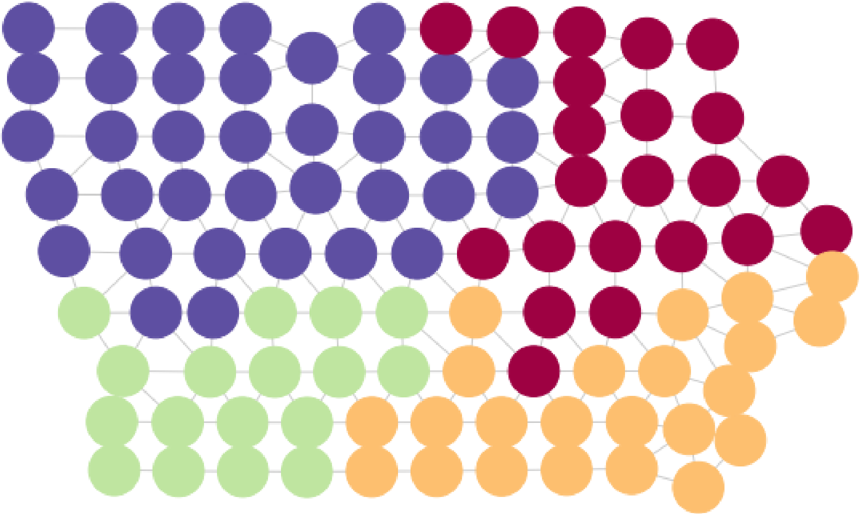
Exploring the set of districtings



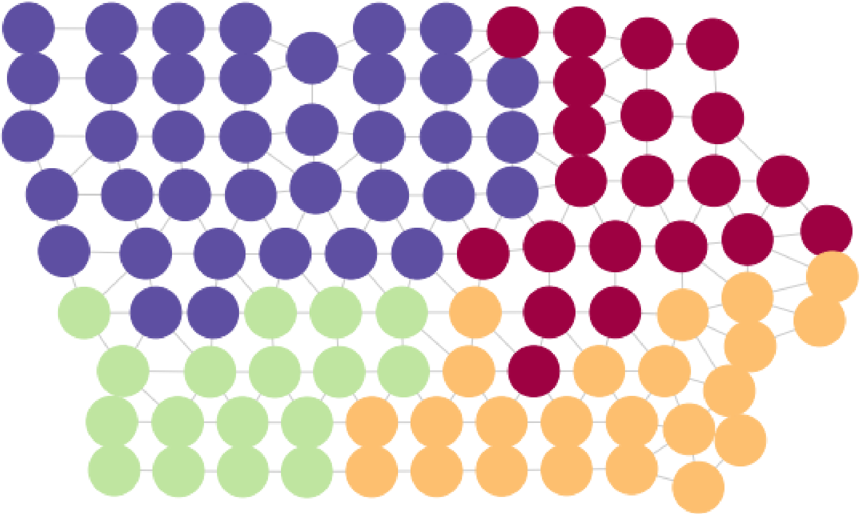
Exploring the set of districtings



Exploring the set of districtings



Exploring the set of districtings



Markov chains

Definition

A discrete time *Markov chain* $\mathcal{M}$ on state space Σ is a sequence of random variables $X_0, X_1, X_2, \dots$ taking values in Σ such that

$$\mathbb{P}(X_t = \sigma | X_0 = \sigma_0, X_1 = \sigma_1, \dots, X_{t-1} = \sigma') = \mathbb{P}(X_t = \sigma | X_{t-1} = \sigma').$$

This is the *transition probability* from σ' to σ, which we denote by $P(\sigma', \sigma)$

In our chain, the state space Σ will be the set of (reasonable) districting partitions of the state dual graph, and we transition between states by flipping border vertices

Finding outliers

At each step, compute some relevant measure of "gerrymandering"

Finding outliers

At each step, compute some relevant measure of "gerrymandering"

We define two measures of interest

Definition

Let σ be a districting of a state with d districts, and let $\delta_i \in [0,1]$ be the proportion of votes party A receives in district $i \in [d]$. The *variance score* of σ is

$$f_{\text{var}}(\sigma) := -\left(\frac{\sum_{i=1}^{d} \delta_i^2}{d} - \left(\frac{\sum_{i=1}^{d} \delta_i}{d}\right)^2\right)$$

The mean-median score of σ is

$$f_{\text{mm}}(\sigma) = \text{median}\{\delta_i\}_{i=1}^{d} - \text{mean}\{\delta_i\}_{i=1}^{d}.$$

Finding outliers

Walk along districting plans, computing f_{var} or f_{mm} at each step. Determine if any are suspiciously bad...

Finding outliers

Walk along districting plans, computing f_{var} or f_{mm} at each step. Determine if any are suspiciously bad...

But what if our walk is very slow? How to determine significance?

The $\sqrt{2\varepsilon}$ test

Theorem (Chikina, Frieze, Pegden '17)

Let $\mathcal{M} = X_0, X_1, \ldots$ be a reversible Markov chain on state space Σ with stationary distribution π, and let $f : \Sigma \to \mathbb{R}$. If $X_0 \sim \pi$, then for any fixed k the probability that $f(X_0)$ is in the lowest ε percentile of $\{f(X_i)\}_{i=0}^k$ is at most $\sqrt{2\varepsilon}$.

Stationary distributions and reversibility

Definition

A probability distribution π on Σ is stationary for a Markov chain $\mathcal{M}$ if $X_0 \sim \pi$ implies $X_i \sim \pi$ for all $i \geq 1$.

Stationary distributions and reversibility

Definition

A probability distribution π on Σ is stationary for a Markov chain $\mathcal{M}$ if $X_0 \sim \pi$ implies $X_i \sim \pi$ for all $i \geq 1$.

Definition

A Markov chain $\mathcal{M}$ is reversible if

$$\mathbb{P}(X_0 = \sigma_0, X_1 = \sigma_1, \ldots, X_k = \sigma_k) = \mathbb{P}(X_0 = \sigma_k, X_1 = \sigma_{k-1}, \ldots, X_k = \sigma_0)$$

for all k.

Theorem, rephrased

Make π uniform on our state space, so that sampling from π is akin to sampling a "typical" districting plan.

Theorem, rephrased

Make π uniform on our state space, so that sampling from π is akin to sampling a "typical" districting plan.

Theorem (Rephrased)

If the random walk on districting plans is reversible, then under the null hypothesis that it starts at a "typical" plan, the probability of the starting plan being an extreme outlier among the resulting sample is small.

League of Women Voters of PA v. Commonwealth of PA

- Lawsuit filed in 2017 to challenge 2011 congressional districts in Pennsylvania

League of Women Voters of PA v. Commonwealth of PA

- Lawsuit filed in 2017 to challenge 2011 congressional districts in Pennsylvania
- In Jan, 2018, court orders Republican state legislature and Democratic governor to compromise

League of Women Voters of PA v. Commonwealth of PA

- Lawsuit filed in 2017 to challenge 2011 congressional districts in Pennsylvania
- In Jan, 2018, court orders Republican state legislature and Democratic governor to compromise
- After this order, maps from 8 different groups (including PA House Dems, PA House Repubs, and the governor) are submitted to the court

League of Women Voters of PA v. Commonwealth of PA

- Lawsuit filed in 2017 to challenge 2011 congressional districts in Pennsylvania
- In Jan, 2018, court orders Republican state legislature and Democratic governor to compromise
- After this order, maps from 8 different groups (including PA House Dems, PA House Repubs, and the governor) are submitted to the court
  - Amazingly, partisan concerns were not considered in the creation of any of these maps!
    - (...not really.)

Duchin's PA Analysis

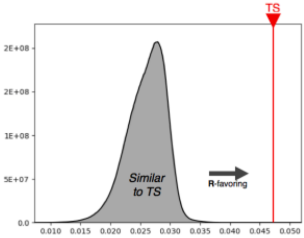
- In February, PA Governor Tom Wolf (D) hires Moon Duchin to do an analysis of the submitted maps

Home > News > Governor Wolf to Enlist Non-Partisan Mathematician to Evaluate Fairness of Redistricting Maps

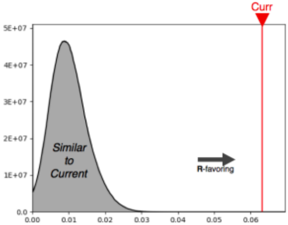
Governor Wolf to Enlist Non-Partisan Mathematician to Evaluate Fairness of Redistricting Maps

January 26, 2018

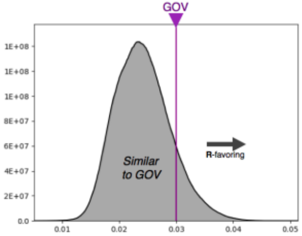
Some convincing figures...



Mean-median
TS plan
SenW



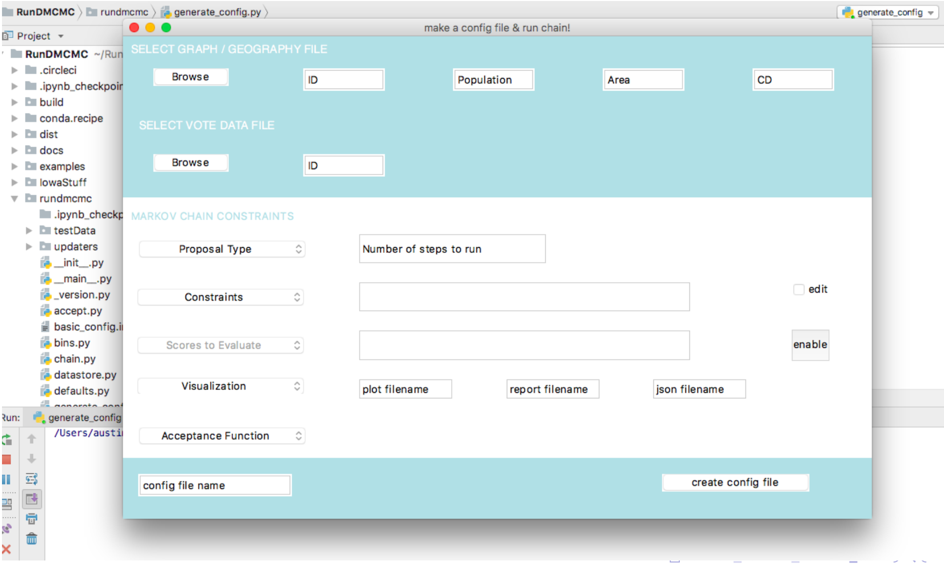
Mean-median
Current plan
SenW



Mean-median
GOV plan
SenW

Try it at home!

RunDMCMC:



Issues

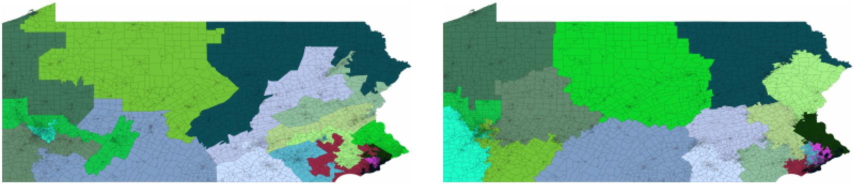


FIG. 2: **Left:** The current districting of Pennsylvania. **Right:** A districting produced by the Markov chain after 2^{40} steps.

(Figure taken from “Assessing Significance in a Markov Chain without mixing” by Chikina, Frieze, and Pegden.)

A graph partition sampling problem

- Want to better understand the random walk on graph partitions.

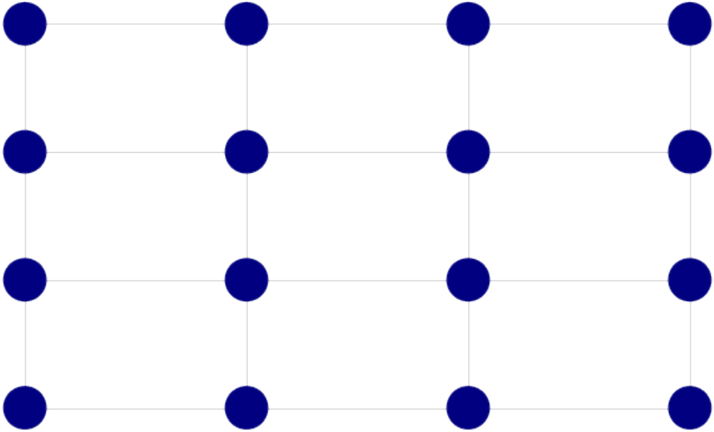
A graph partition sampling problem

- Want to better understand the random walk on graph partitions.
  - Walk shown before is quite slow.

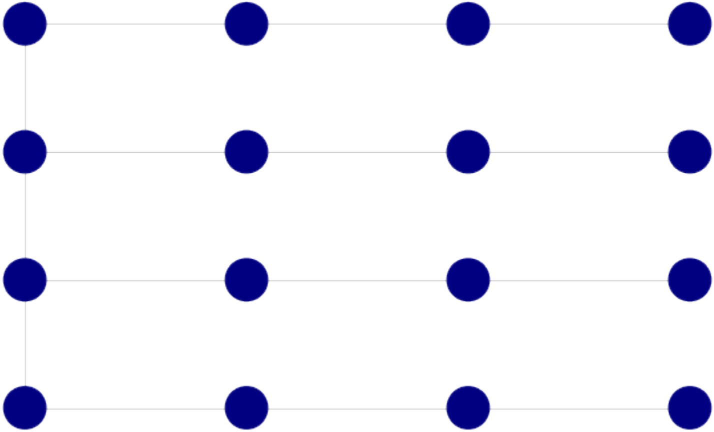
A graph partition sampling problem

- Want to better understand the random walk on graph partitions.
  - Walk shown before is quite slow.
- An "easy" way to get a connected partition of a graph G into k parts: take a spanning tree T of G and cut $k-1$ edges:

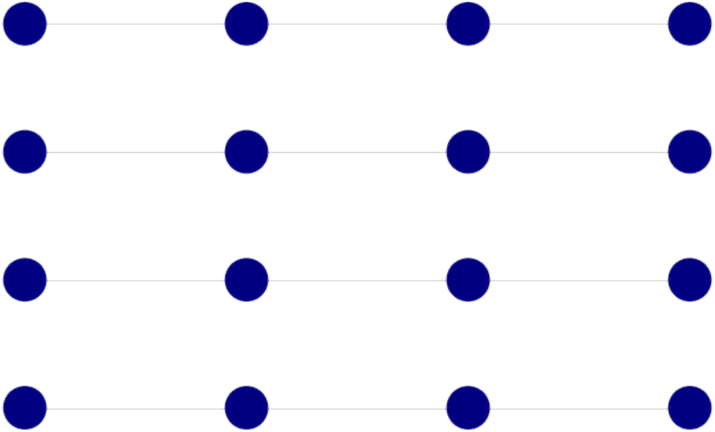
Example: the 4×4 grid, aka Wyoming



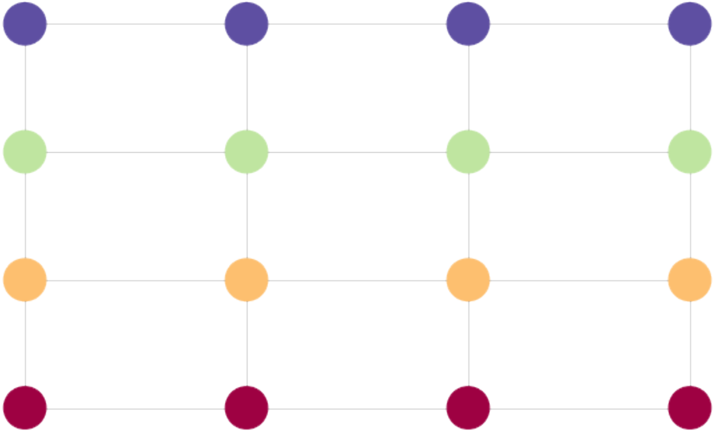
Example: the 4×4 grid, aka Wyoming



Example: the 4×4 grid, aka Wyoming



Example: the 4×4 grid, aka Wyoming



A random walk on spanning trees, and a "shadow walk" on partitions

Algorithm (Spanning Tree Walk).

Let T_i be a spanning tree of G

1. Pick $e \in E(T_i)$ and $e' \in E(G)$ u.a.r.
2. If $T_i - e + e'$ is a tree, set $T_{i+1} = T_i - e + e'$
3. Else, $T_{i+1} = T_i$

The above describes a Markov chain on the set of spanning trees of G.

A random walk on spanning trees, and a "shadow walk" on partitions

Algorithm (Spanning Tree Walk).

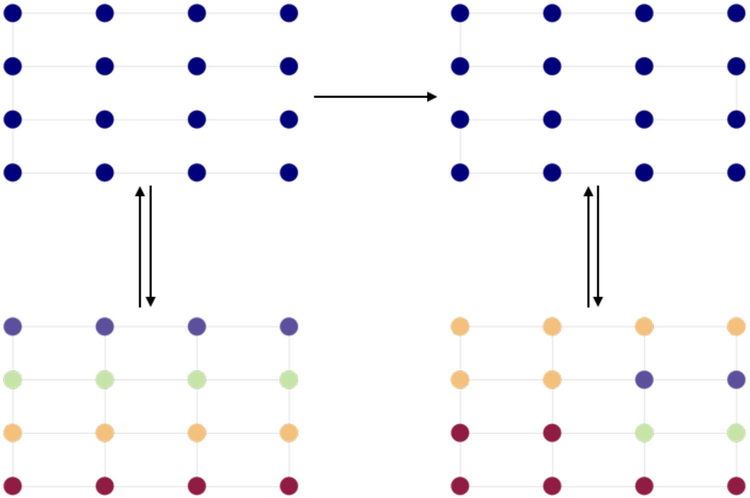
Let T_i be a spanning tree of G

1. Pick $e \in E(T_i)$ and $e' \in E(G)$ u.a.r.
2. If $T_i - e + e'$ is a tree, set $T_{i+1} = T_i - e + e'$
3. Else, $T_{i+1} = T_i$

The above describes a Markov chain on the set of spanning trees of G.

Choose $k - 1$ edges to cut from each T_i.

A random walk on spanning trees, and a "shadow walk" on partitions



Partition Spectra

Call the set of partitions of $V(G)$ obtainable by cutting $k-1$ edges from a spanning tree T the *k-partition spectrum* of T.

Partition Spectra

Call the set of partitions of $V(G)$ obtainable by cutting $k - 1$ edges from a spanning tree T the *k-partition spectrum* of T.

Question: can we bound the overlap of the k-partition spectra of spanning trees T and T' in terms of $d(T, T')$?

Some evidence:

#SpanningTrees as a measure of compactness

The sampling method described above is far from uniform partition. In fact, we can compute the bias relatively easily.

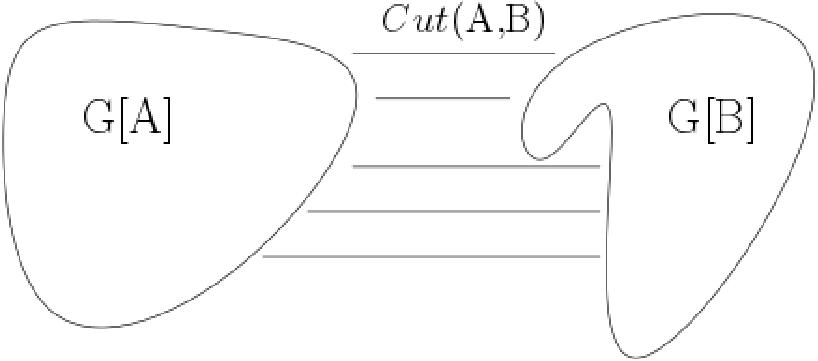
#SpanningTrees as a measure of compactness

The sampling method described above is far from uniform partition. In fact, we can compute the bias relatively easily. For instance, if we cut one edge from a uniform tree T, then:

$$\mathbb{P}(\text{drawing partition } P = (A, B)) = \frac{\tau(G[A])\tau(G[B]))\text{Cut}(A, B)}{\tau(G)(n-1)}$$

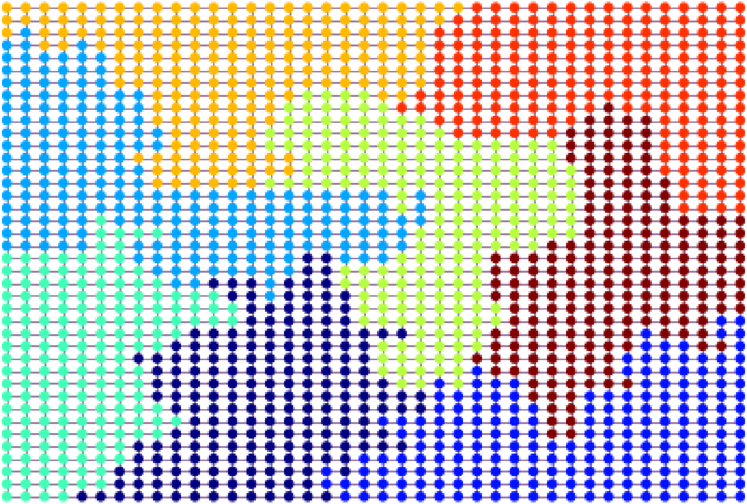
where $\tau(H) =$ the number of spanning trees of a graph H.

#SpanningTrees as a measure of compactness

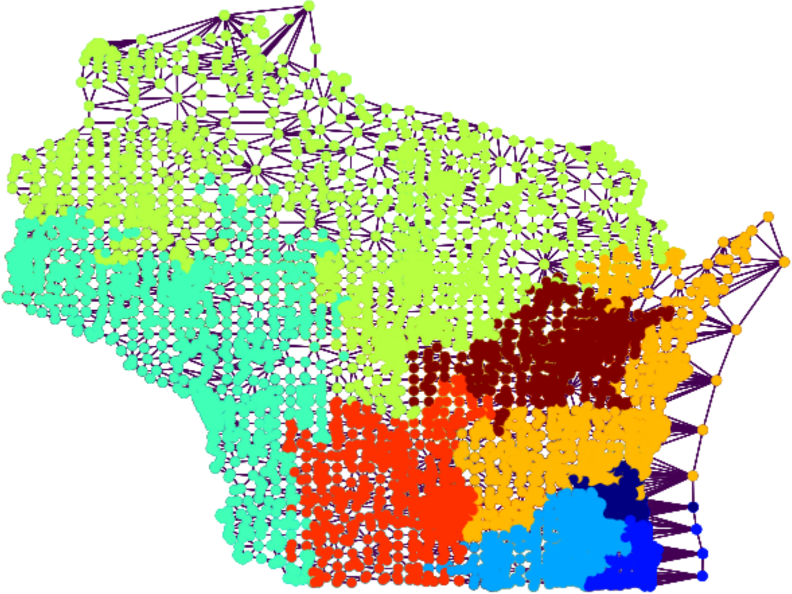


$$\mathbb{P}(\text{drawing partition } P = (A, B)) = \frac{\tau(G[A])\tau(G[B]))\text{Cut}(A, B)}{\tau(G)(n-1)}$$

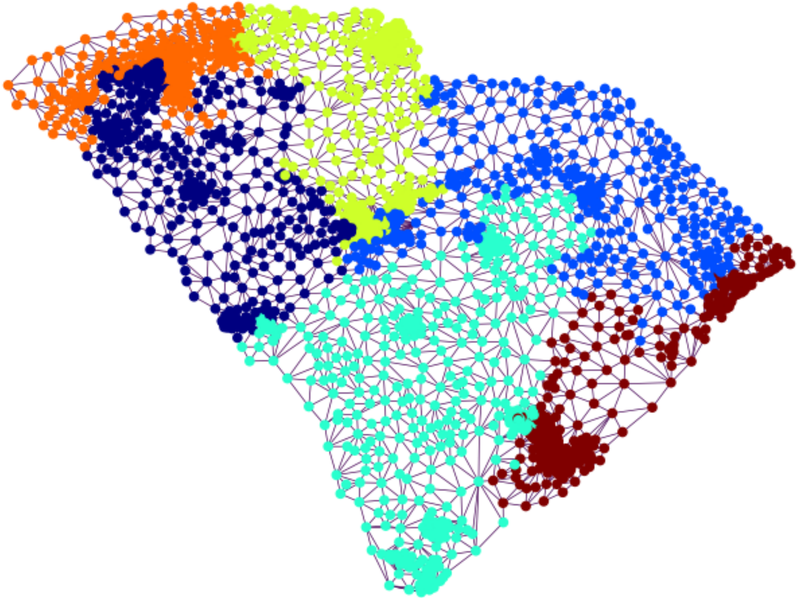
#SpanningTrees as a measure of compactness



#SpanningTrees as a measure of compactness



#SpanningTrees as a measure of compactness

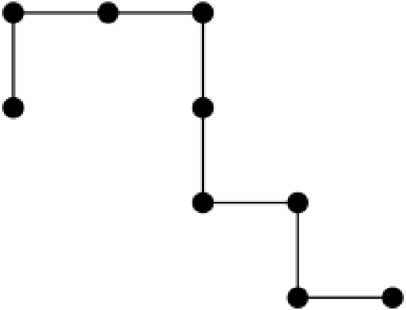


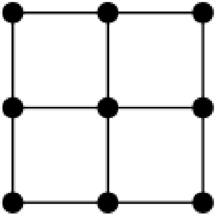
Maximizing #SpanningTrees

Question: which n-vertex connected (induced) subgraph of the infinite grid has the most spanning trees?

Maximizing #SpanningTrees

Question: which *n*-vertex connected (induced) subgraph of the infinite grid has the most spanning trees?





Kirchoff's (Matrix-Tree) Theorem

Theorem (Kirchoff's Theorem)

Let G be a connected graph on n vertices, L be the Laplacian of G, and $\widehat{L}_j$ be L with the jth row and column removed. Then

$$\tau(G) = \det(\widehat{L}_j)$$

for any $j \in [n]$.

Kirchoff's (Matrix-Tree) Theorem

Theorem (Kirchoff's Theorem)

Let G be a connected graph on n vertices, L be the Laplacian of G, and $\widehat{L}_j$ be L with the jth row and column removed. Then

$$\tau(G) = \det(\widehat{L}_j)$$

for any $j \in [n]$.

A result of Fan Chung

Theorem (F. Chung, 1999)

Let H be a connected induced subgraph of the infinite grid. Then

$$ce^{c_1|H|-c_2|\partial H|} \leq \tau(H) \leq c' e^{c_1|H|+c_3\frac{|H|}{|\partial H|}}$$

where c, c' are constants which depend only on the infinite grid, and c_1, c_2, c_3 are constants which depend on $|H|$.

A result of Fan Chung

Theorem (F. Chung, 1999)

Let H be a connected induced subgraph of the infinite grid. Then

$$ce^{c_1|H|-c_2|\partial H|} \leq \tau(H) \leq c' e^{c_1|H|+c_3\frac{|H|}{|\partial H|}}$$

where c, c' are constants which depend only on the infinite grid, and c_1, c_2, c_3 are constants which depend on $|H|$.

But, the bounds are not particularly tight.

An optimization approach

Theorem (Dissanayake, Huang, Sukhatme, Khosoussi '18)

Let $V = V(K_n)$, $E = E(K_n))$ and $E_{init} \subseteq E$ such that $G = (V, E_{init})$ is connected. The function $\log TG : \mathcal{P}(E) \to \mathbb{R}$ given by

$$\log TG(S) = \log\det[\widehat{L}(S \cup E_{init})] - \log\det[\widehat{L}(E_{init})]$$

is monotone and submodular.

An optimization approach

Theorem (Dissanayake, Huang, Sukhatme, Khosoussi '18)

Let $V = V(K_n)$, $E = E(K_n))$ and $E_{init} \subseteq E$ such that $G = (V, E_{init})$ is connected. The function $\log TG : \mathcal{P}(E) \to \mathbb{R}$ given by

$$\log TG(S) = \log\det[\widehat{L}(S \cup E_{init})] - \log\det[\widehat{L}(E_{init})]$$

is monotone and submodular.