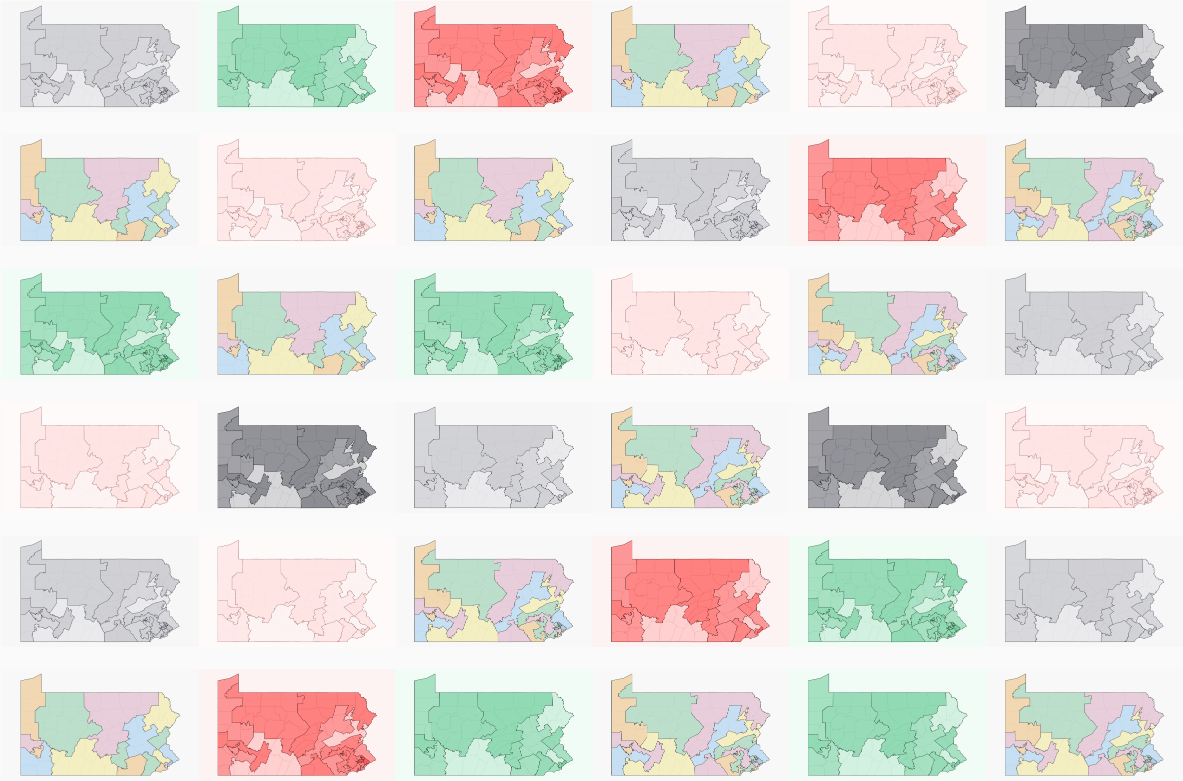
Sequential Monte Carlo for Sampling Balanced and Compact Redistricting Plans

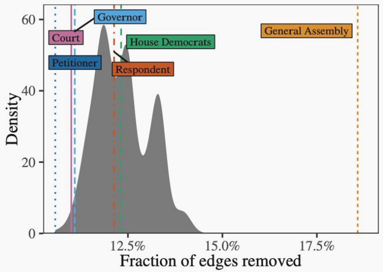
Cory McCartan and Kosuke Imai

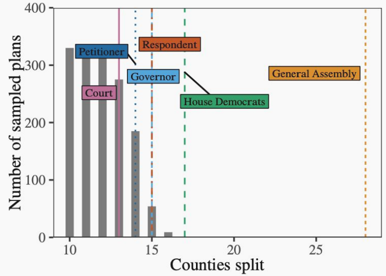
September 13, 2021

Redistricting Evaluation using Ensembles



Summary statistics





Overview of Existing Redistricting Algorithms

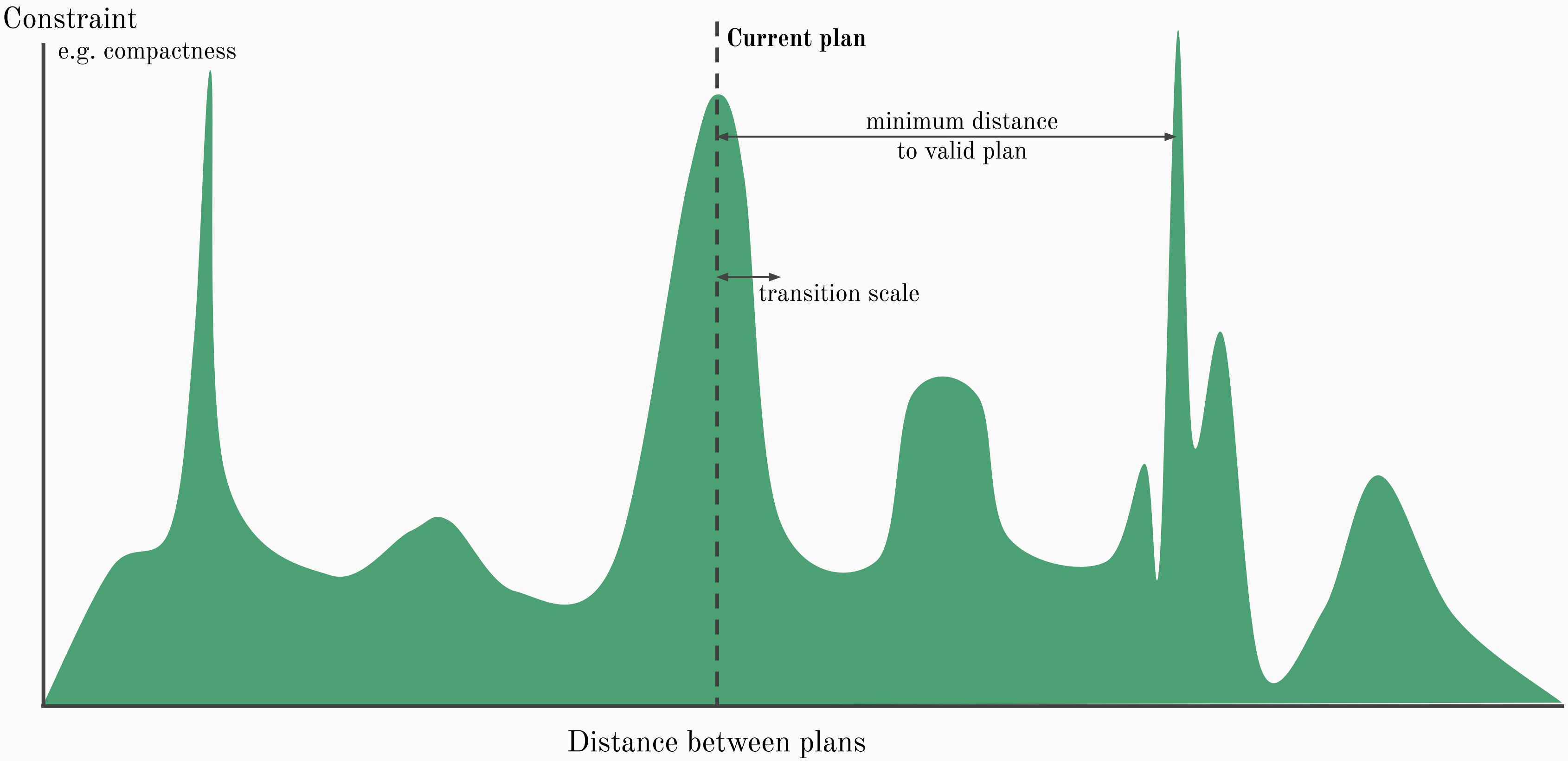
- **Optimization-based methods**
    - Goal: generate maps that have certain characteristics
    - Scalable and flexible
        - Liu et al. 2016
- **Constructive Monte Carlo**
    - Seed-and-grow algorithms
    - Similar to optimization methods
        - Chen & Rodden 2013
- **Enumeration**
    - Goal: general all possible maps
    - Works for small problems
        - Fifield et al. 2020b

- **Markov chain Monte Carlo (MCMC)**
    - Goal: generate *representative* maps under constraints
    - Explicit target distribution
    - Start with an existing map and change it bit by bit
    - Flip algorithms
        - change boundaries
        - Fifield et al. 2014/2020a; Mattingly & Vaughn 2014; Chikina et al. 2017
    - Merge-and-split algorithms
        - much improved mixing
        - Deford et al. 2019/2021; Carter et al. 2019

Challenges of Generating Ensembles

- More plans than atoms
    - Can’t enumerate except for small problems
- Specific target distribution
    - So that we can understand what plans are being generated
- Flexible and realistic set of constraints
- Scale to large problems
    - So that the algorithms are applicable to real-world problems

Difficulties with MCMC



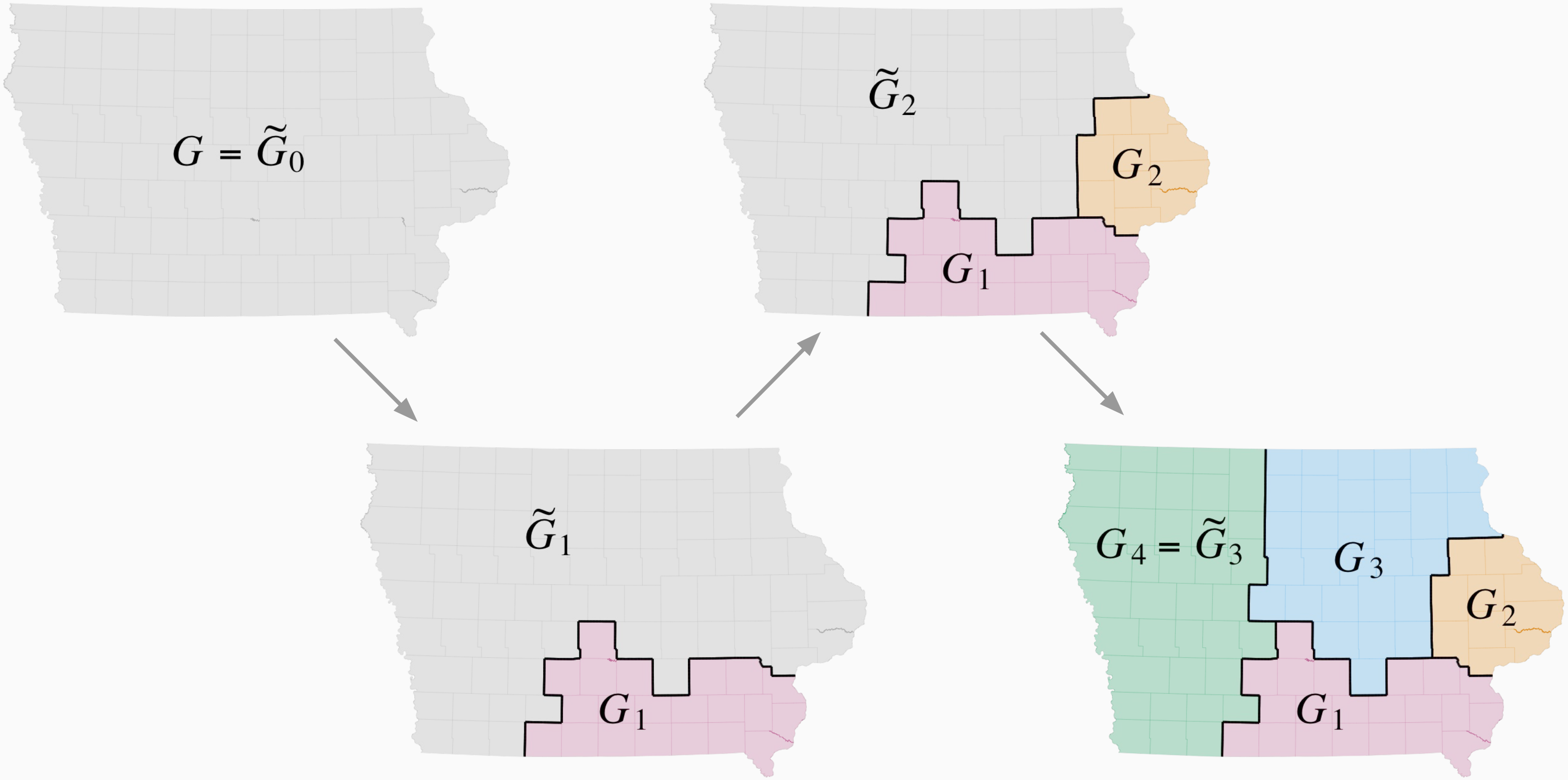
A New Algorithm: Sequential Monte Carlo (SMC)

- Generate (nearly) independent maps
- Can incorporate basic set of constraints *by design*
    - Population constraint
    - Contiguity
    - Compactness
    - County-split constraint
- Other constraints can also be incorporated indirectly as usual via “plan score”
- Efficient and Scalable - applicable to any state in the US

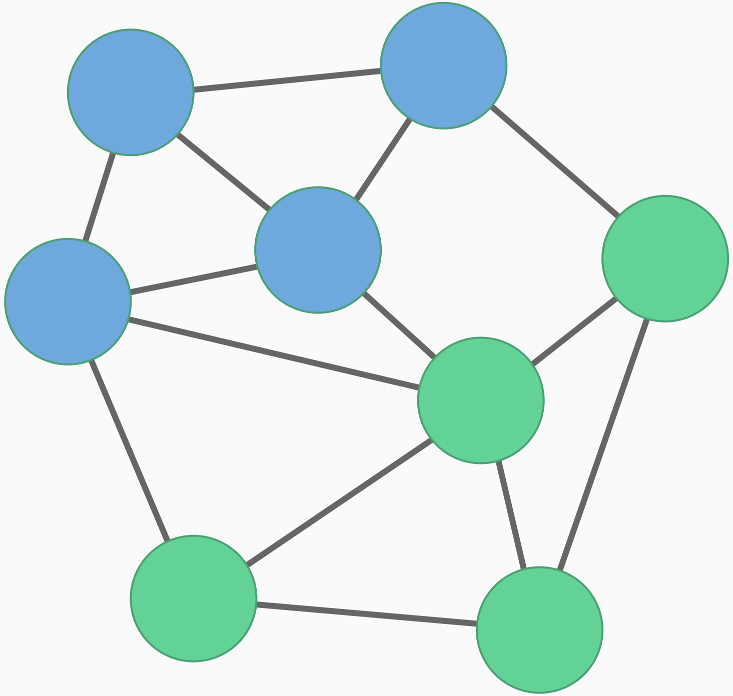
- Limitation: strict constraints lead to inefficient sampling
- Can be combined with the merge-split MCMC

The Algorithm

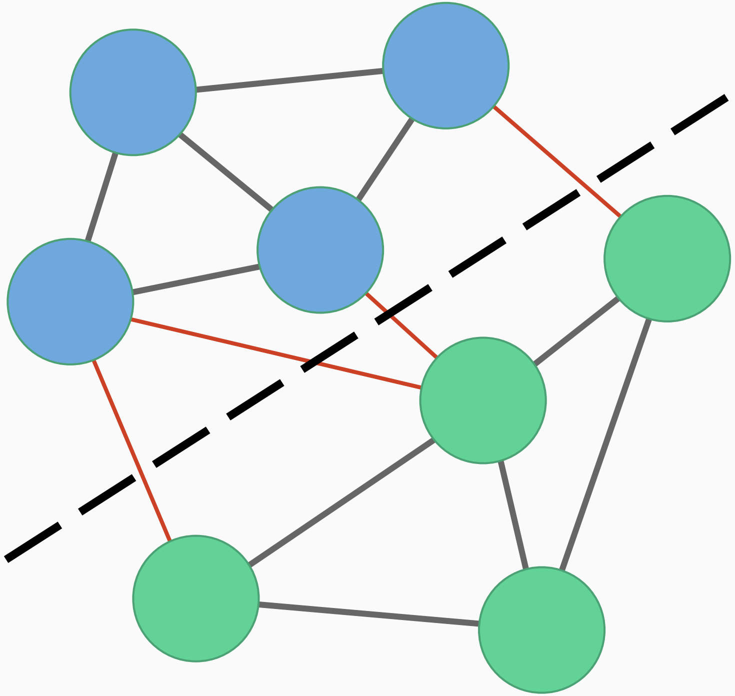
Sequential Splitting



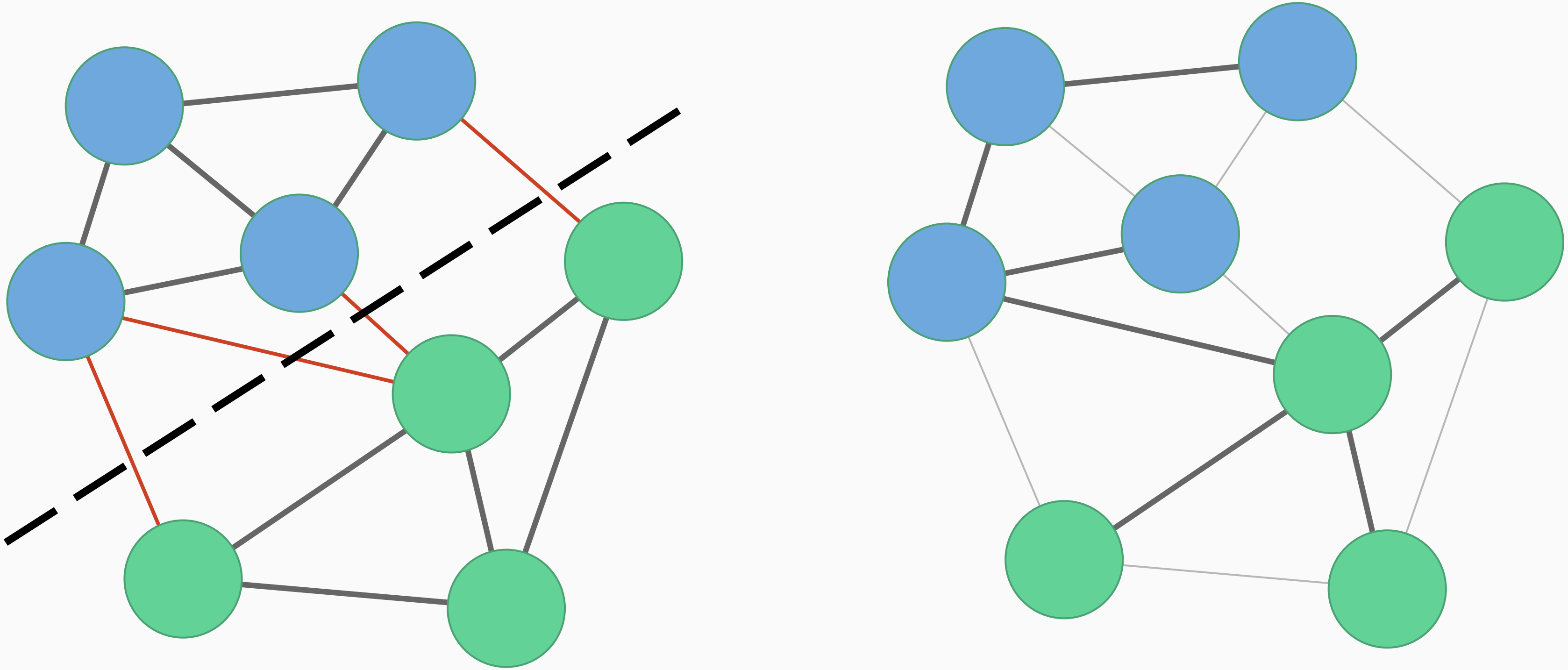
Partitioning a Graph



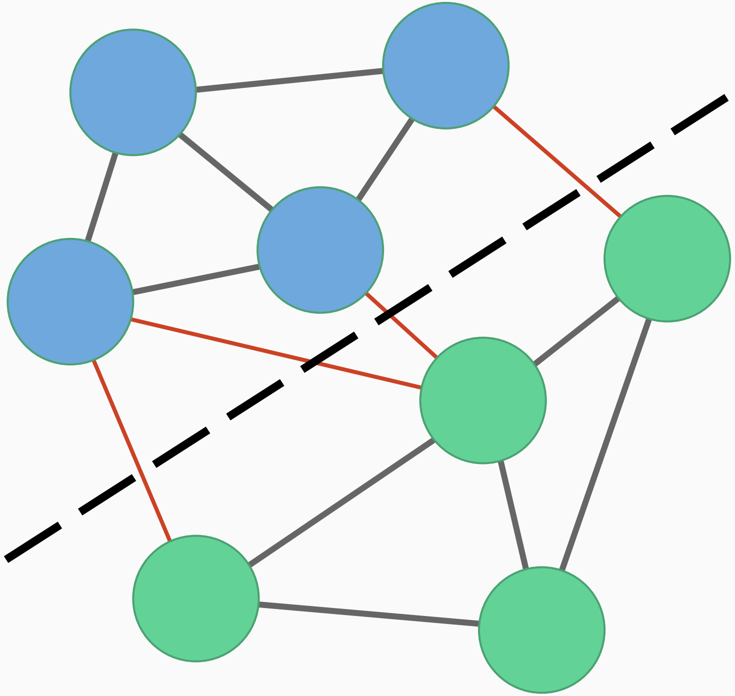
Partitioning a Graph

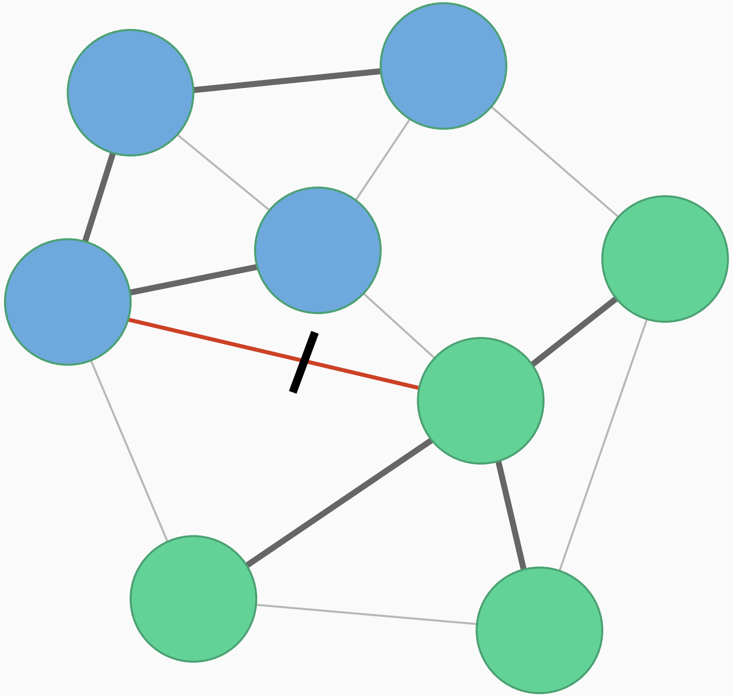


Spanning Trees



Spanning Trees



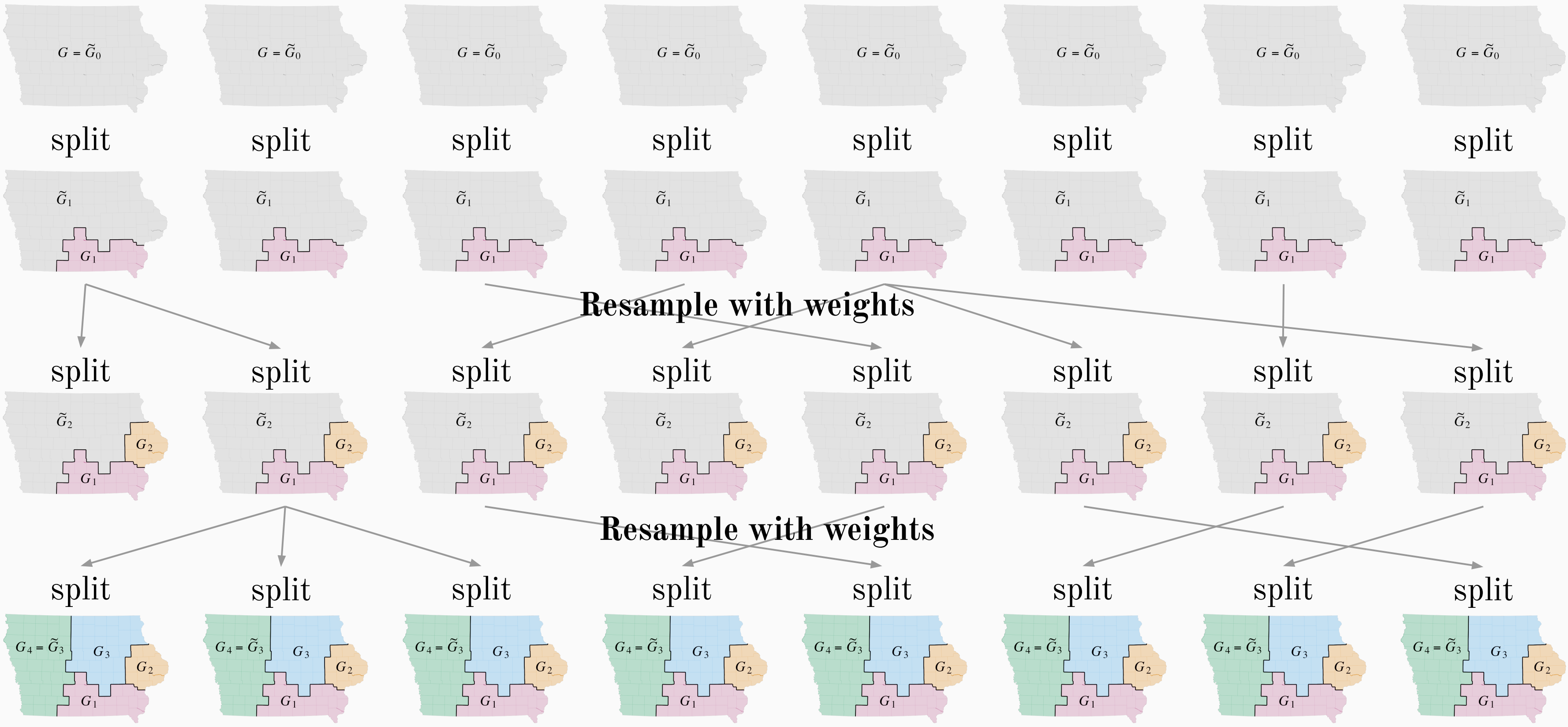


The Splitting Procedure

Repeat n–1 times
to generate n districts

1. Generate a uniform spanning tree
2. Sort edges by population deviation
3. Sample one edge from top k and remove it
4. Check population bounds

Sequential Monte Carlo



split split split split split split split split

Resample with weights

split split split split split split split split

Resample with weights

split split split split split split split split

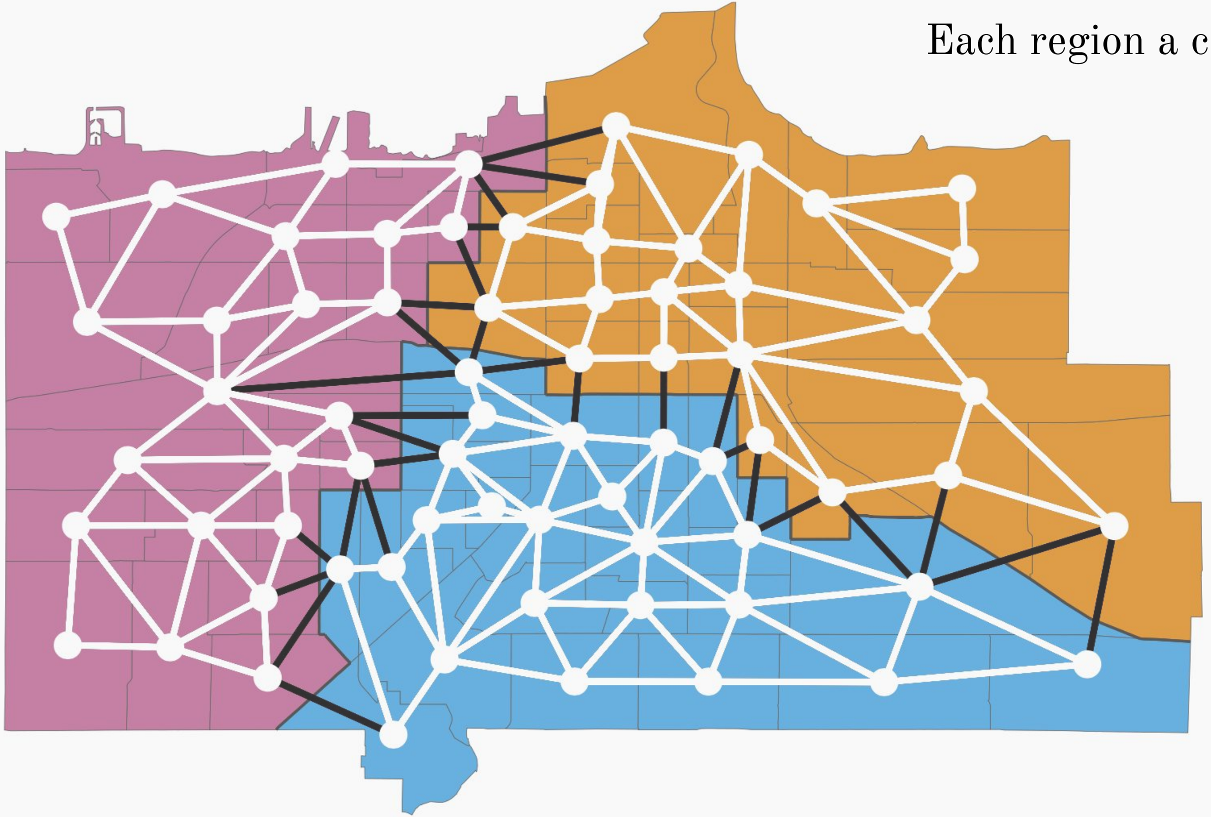
Resample with weights

The SMC Algorithm

To generate a properly weighted sample of S (nearly) independent redistricting plans

```
1.  Generate $S$ initial copies of map
    Set all weights to 1
2.  For $i \in \{1, 2, ..., n - 1\}$:
        a. Until there are $S$ successes:
              i. Sample a map according to the weights
             ii. Use the Splitting Procedure to split off a new district from each of the existing maps
            iii. Reject if population outside bounds
        b. Calculate new weights based on splitting probability
3.  Calculate final weights
4.  Output complete plans and weights
```

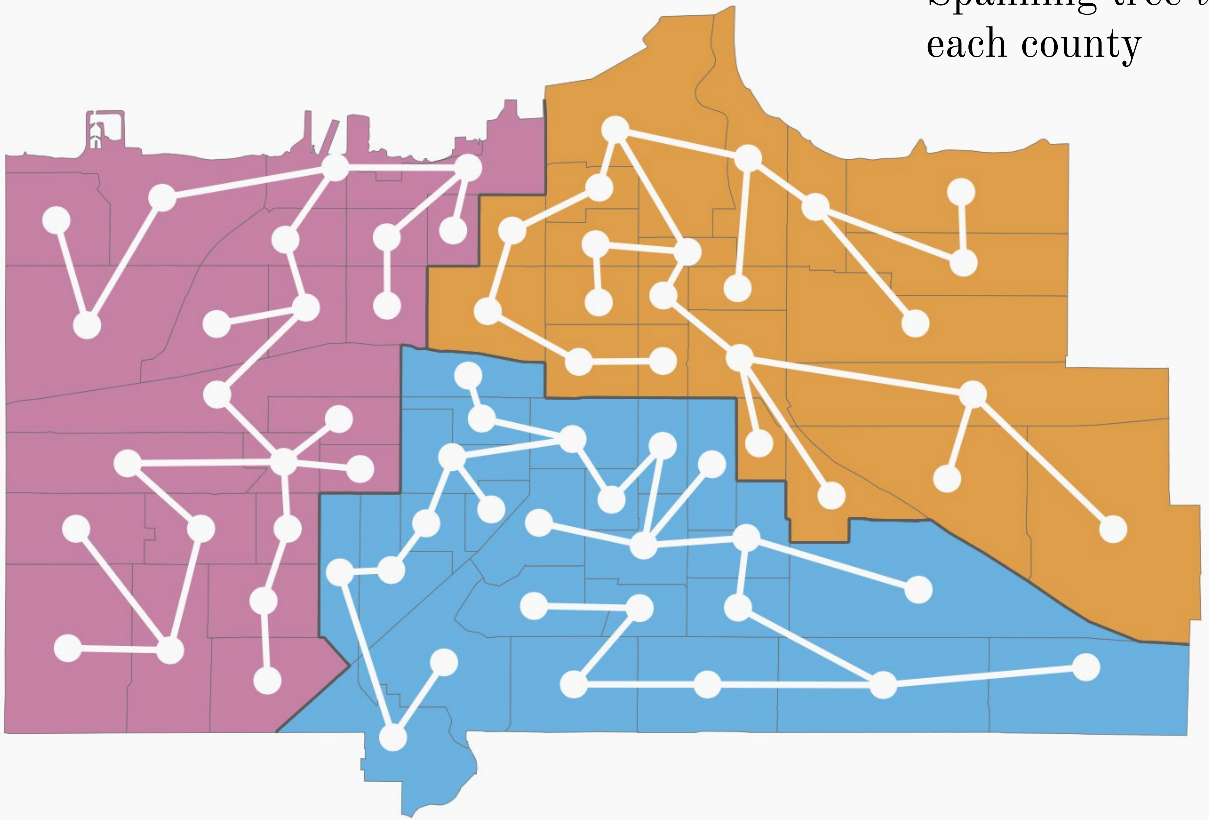
Hierarchical Sampling



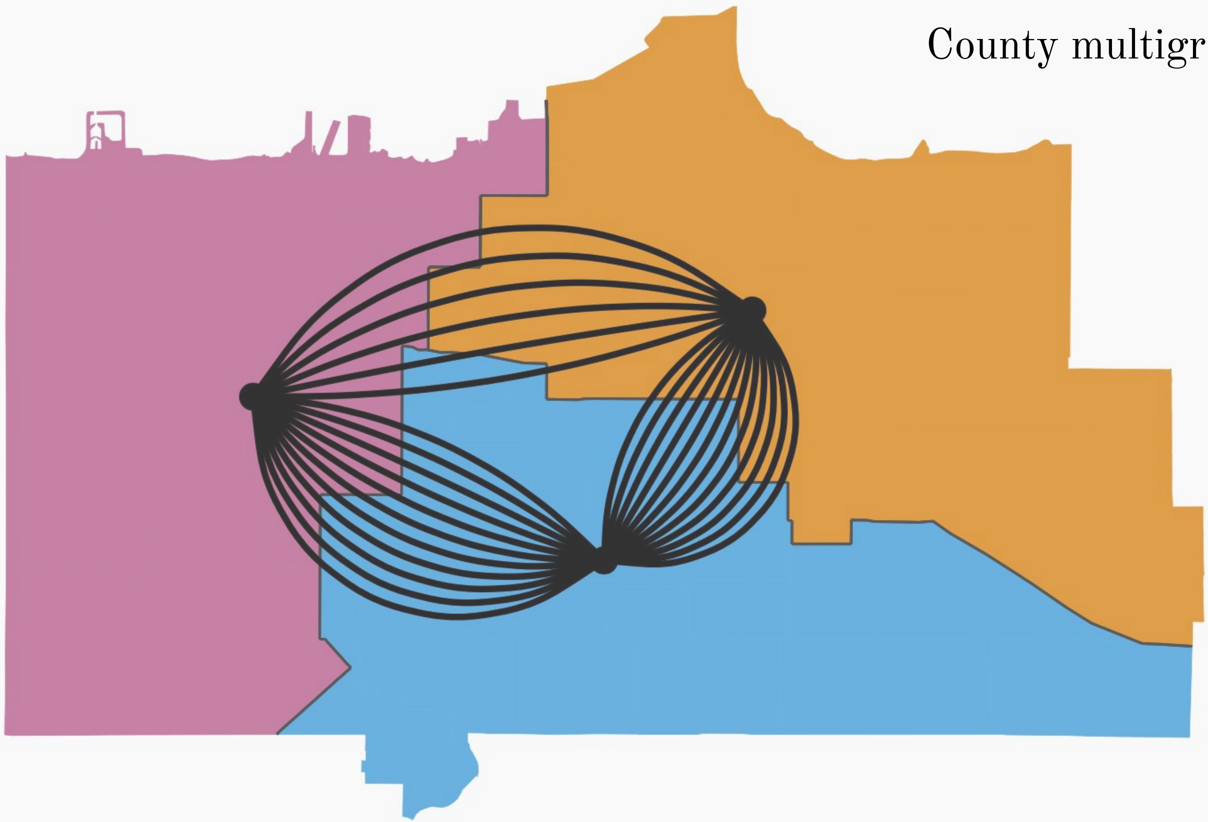
Each region a county

Hierarchical Sampling

Spanning tree *within* each county



Hierarchical Sampling

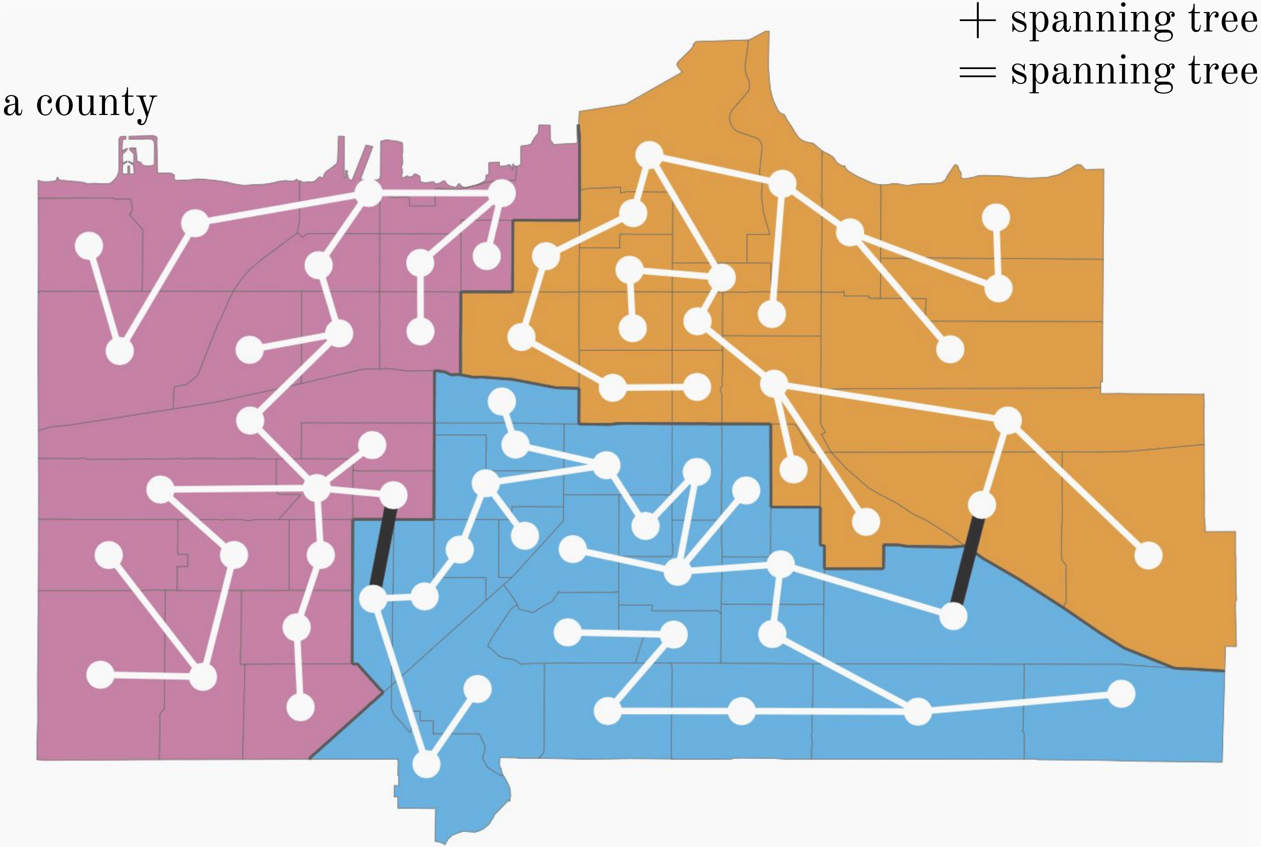


County multigraph

Avoiding County Splits

Spanning tree on
+ spanning trees on counties
= spanning tree on

Each region a county



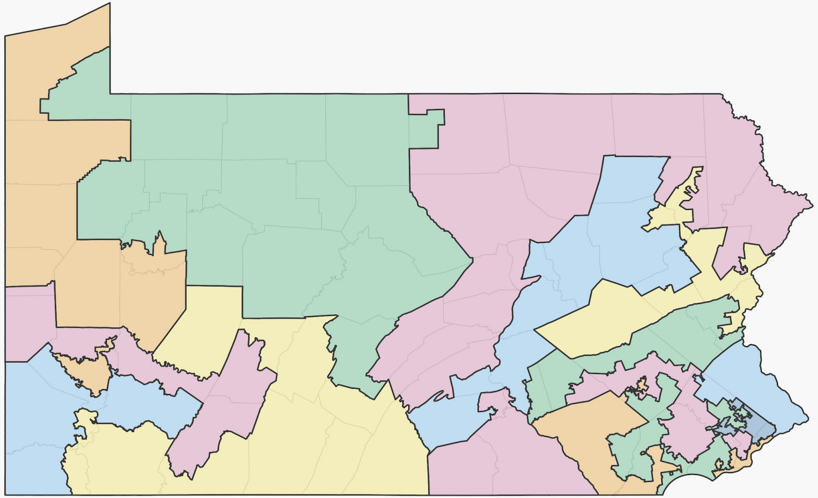
see also Autry, E. A., Carter, D., Herschlag, G., Hunter, Z., & Mattingly, J. C. (2020). Multi-scale merge-split Markov chain Monte Carlo for redistricting.

The 2011 Pennsylvania Redistricting

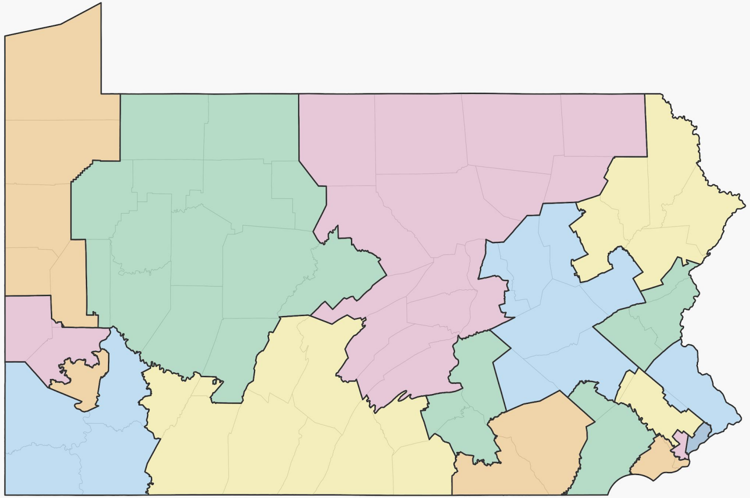
> “First, the Court finds as a matter of law that the Congressional Redistricting Act of 2011 **clearly, plainly and palpably violates the Constitution** of the Commonwealth of Pennsylvania, and, on that sole basis, we hereby strike it as unconstitutional.”

JANUARY 22, 2018, PENNSYLVANIA SUPREME COURT
League of Women Voters of Pennsylvania *v.* Commonwealth of Pennsylvania

General assembly plan



Court's remedial plan



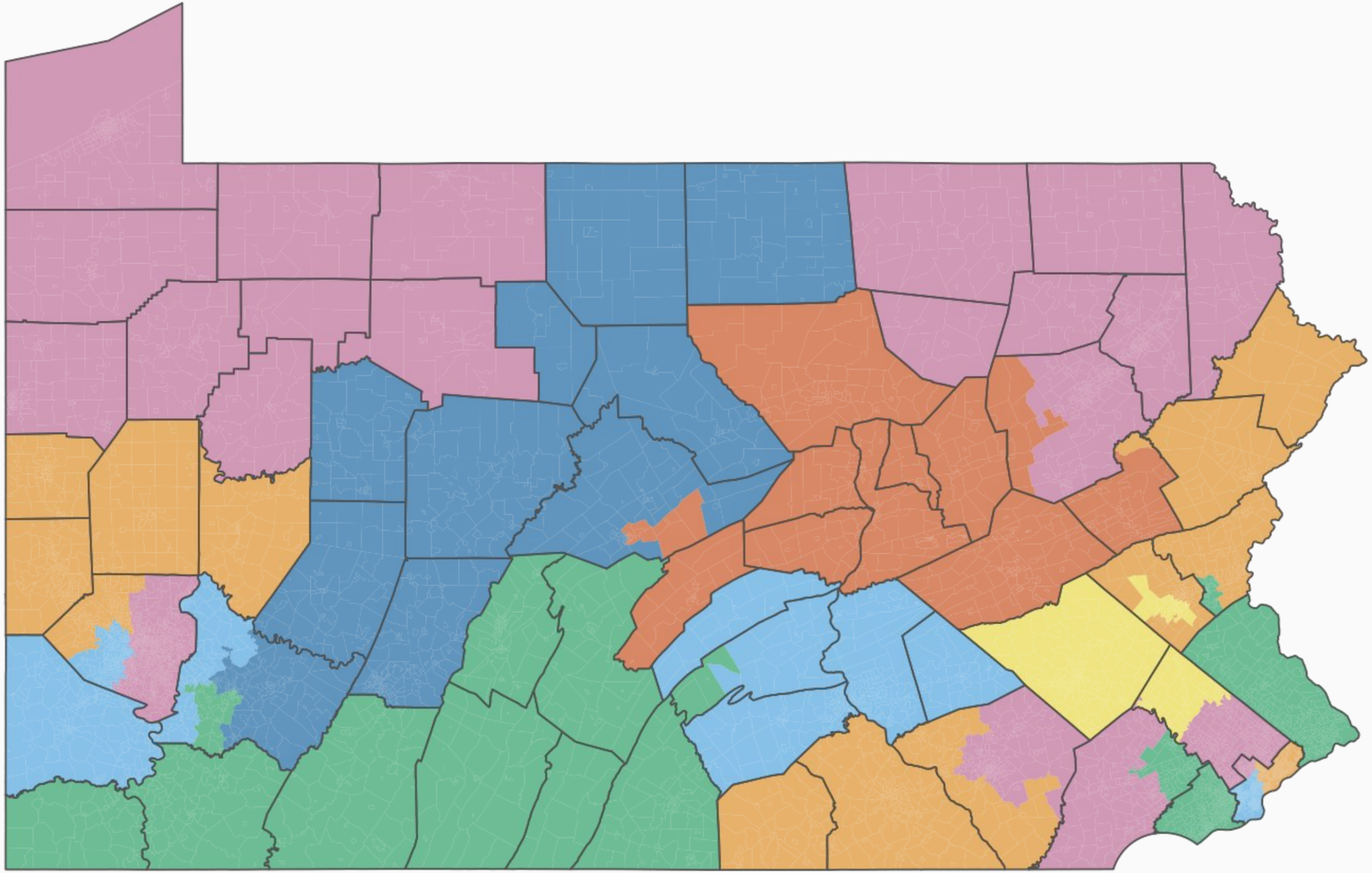
Six plans

1. (Original) General Assembly plan (Republican)
2. Court’s remedial plan
3. Governor’s plan (Democratic)
4. House Democrats’ plan
5. Petitioner’s plan
6. Respondent’s plan (Republican)

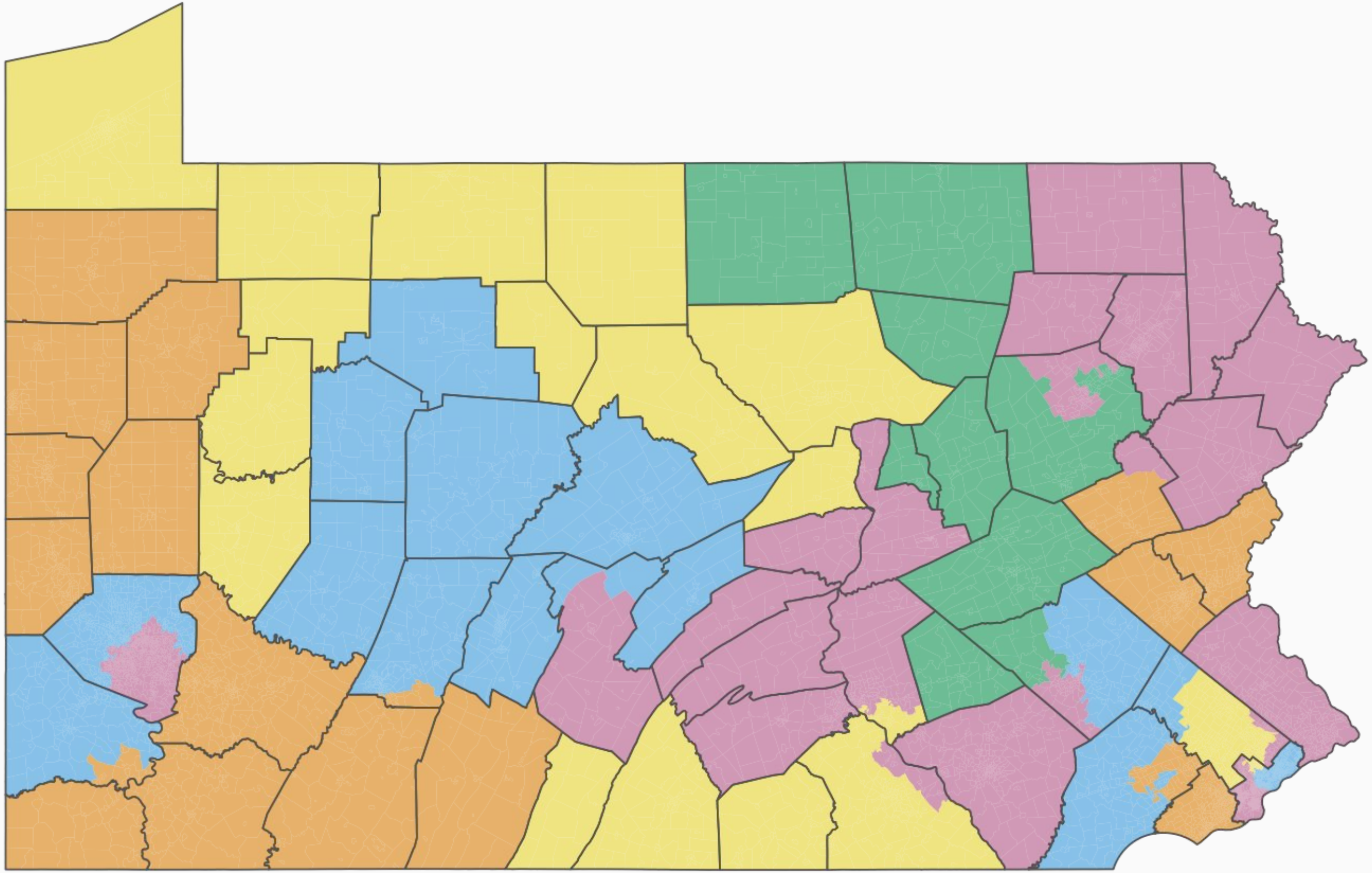
Sampling Details

- 1,500 samples
- 9,256 precincts, 18 congressional districts
- Maximum 0.1% population deviation (± 700 people)
- Compact districts ($\varrho = 1$)
- Maximum 17 county splits

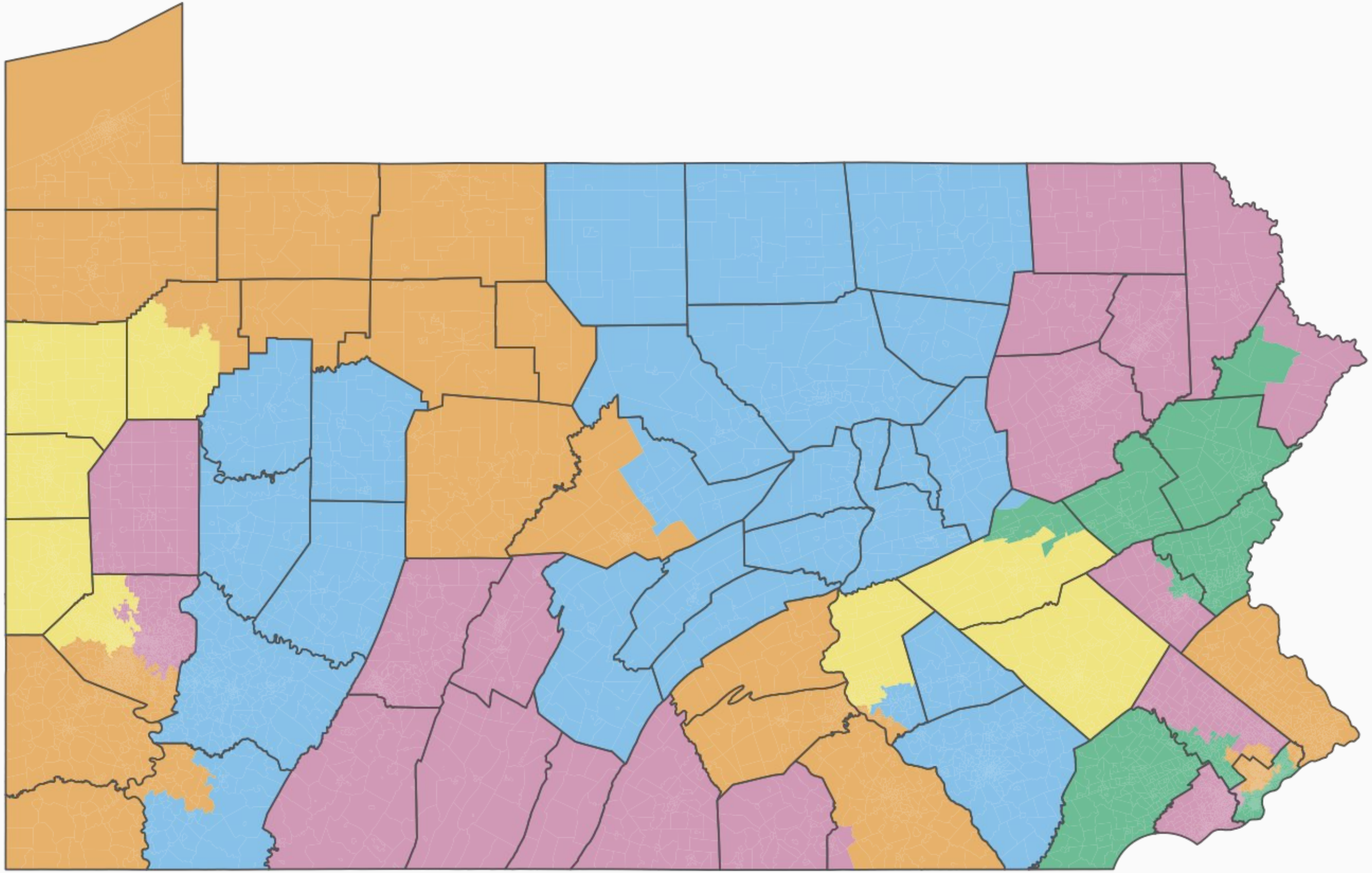
Some Samples



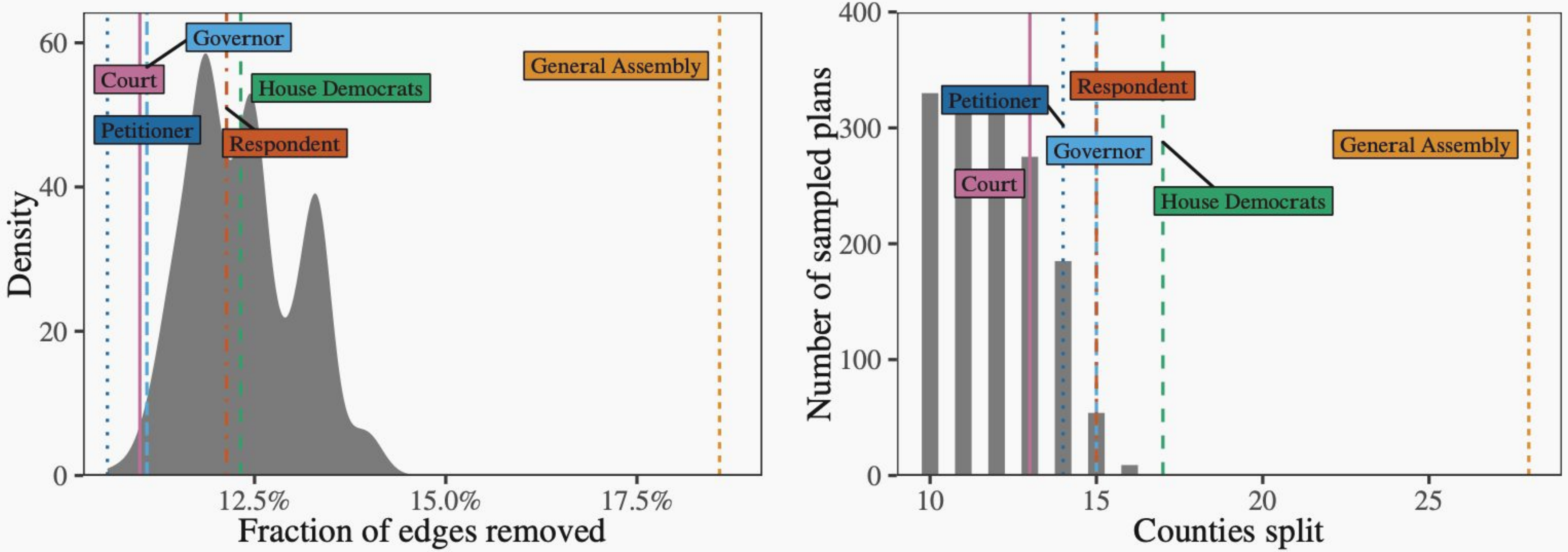
Some Samples



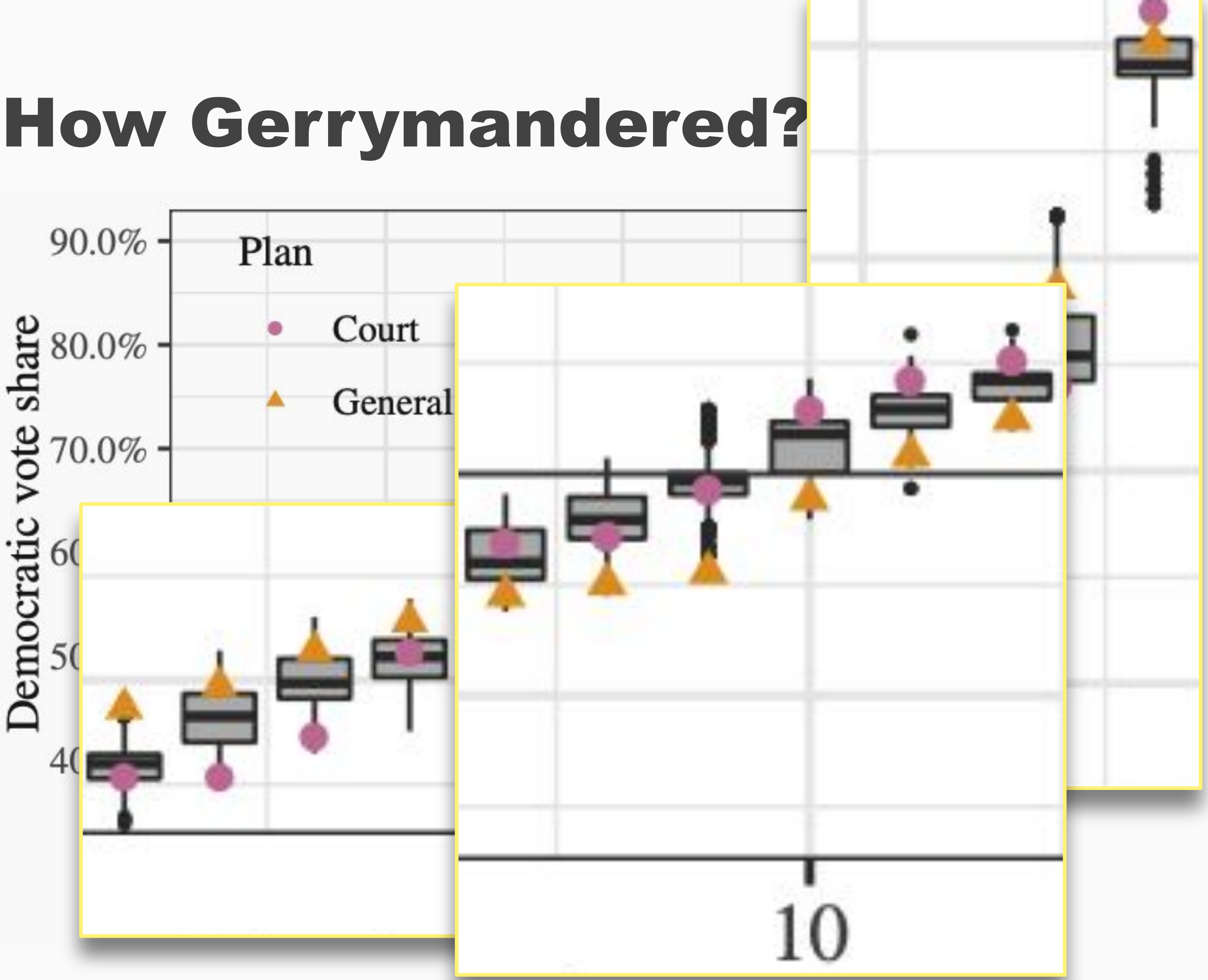
Some Samples



Compactness and County Splits



How Gerrymandered?



Efficiency Comparison

	Gerrymandering index		
	SMC	ReCom	Merge-split
Nominal samples	1,500	1,500	1,500
Effective samples	**580.2**	76.0	27.1
Efficiency	**38.7%**	5.1%	1.8%

Open-source
R package `redist`

Algorithm-assisted redistricting analysis by citizen data scientists

- Implemented algorithms:
  - SMC
  - Merge-split
  - Flip MCMC
  - Enumeration
  - Short-burst
- Various metrics:
  - Population deviation
  - Compactness
  - Competitiveness
  - Partisan fairness
- Tools:
  - Painless data preparation
  - Easy to summarize and analyze redistricting plans
  - Can be used on one's laptop
- Visualization:
  - Easy automatic visualization for quantities of interest
  - Plot redistricting plans
  - Interactive visualization
- Website: https://alarm-redist.github.io/