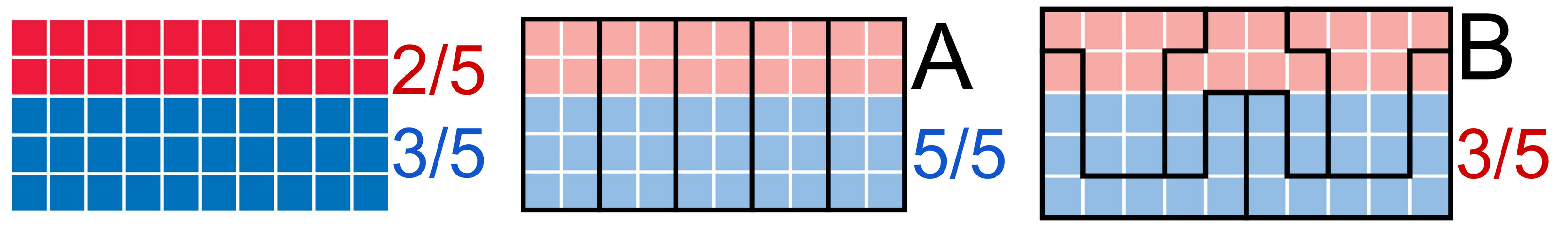
Fairmandering: A Column Generation Heuristic for Fairness Optimized Political Redistricting

Wes Gurnee [1, 2] David B. Shmoys [1]

[1]Cornell University [2]MIT

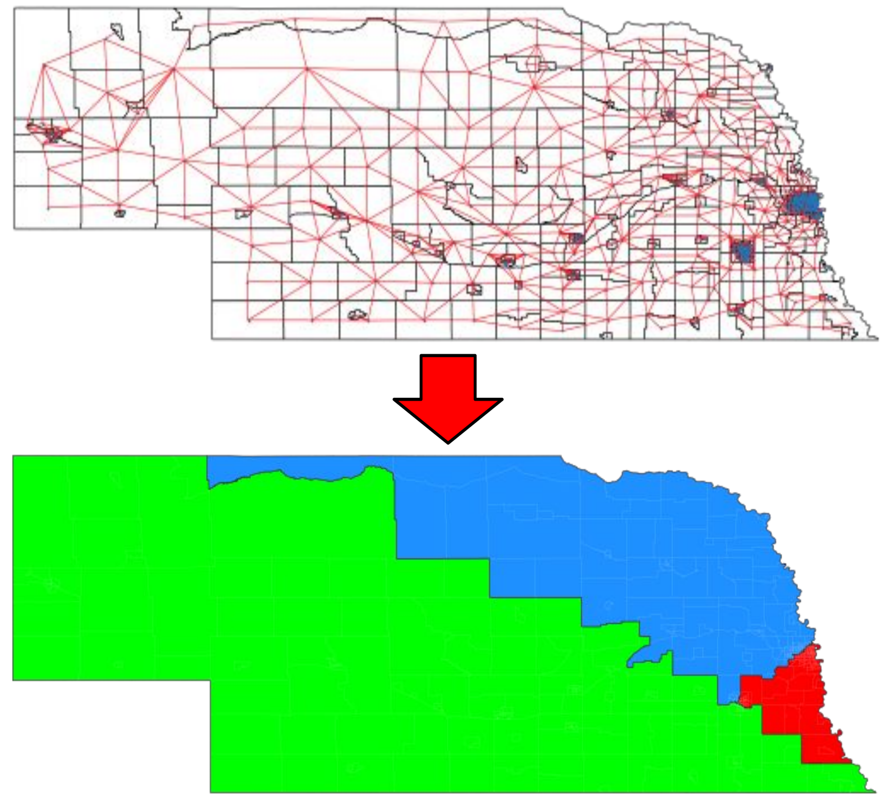
Redistricting and Gerrymandering

Every 10 years in the United States, 428 congressional, 1938 state senate, and 4826 state house districts are redrawn, cementing the partisan power balance for the following decade in a process known as redistricting. In most states, politicians get to draw these lines, enabling partisans to secure a partisan advantage, suppress the vote of minority groups, and protect incumbents from competition. Such practices, broadly known as gerrymandering, are accomplished by "cracking" and "packing" opposition voters to strategically dilute their power. The goal of our work is to use these mechanisms to instead generate more representative districts with a scalable algorithm.



Political Districting Problem (PDP)

- Input
  1. Atomic Geographic Blocks (B) - census units or precincts
  2. Adjacency Graph $G = (B, E)$ where p_i denotes the population for each $i \in B$ and d_{ij} denotes the distance of each edge $i, j \in E$
  3. Parameters - k is the number of districts in a plan and ϵ_p is the population tolerance.
- Constraints
  1. Population balanced (deviation $< \epsilon_p$).
  2. Contiguous - it is possible to reach any block from any other block without leaving the district.
  3. Geographically compact - the districts form geometrically simple shapes.
- Objective
  1. Compactness (e.g., moment of inertia, Roeck, Polsby Popper, total perimeter, edge cuts).
  2. Fairness (e.g., proportionality, partisan symmetry).
- Output - An assignment of blocks to districts that forms a connected k-partition.



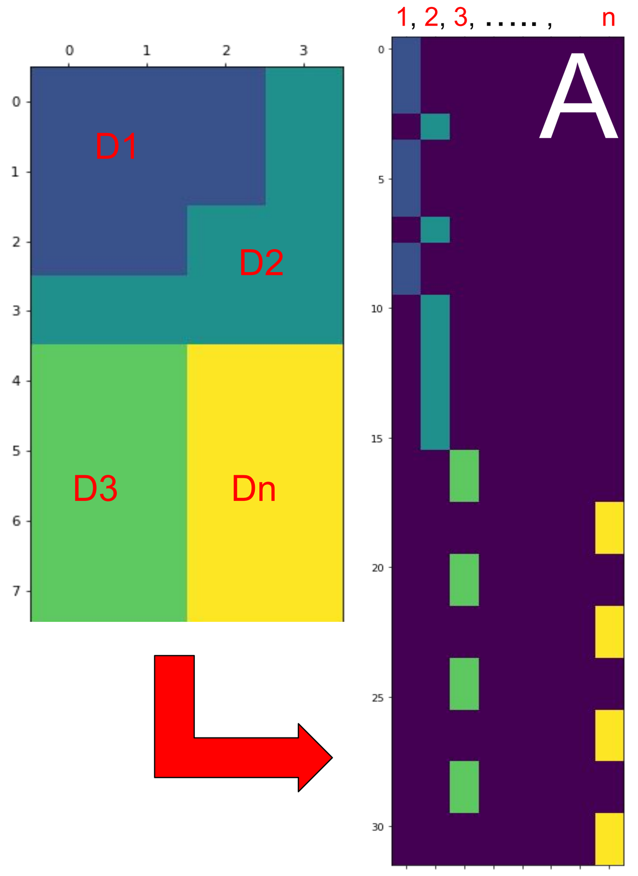
Decoupled Formulation

Instead of optimizing a plan in one shot, we use a decoupled formulation and first generate a large set of legal districts D (traditional on-the-fly column generation does not work because of the degeneracy of the master problem). These districts are collected into a binary block-district matrix A which encodes the assignment of blocks to districts. Assuming all districts are contiguous and population balanced, the set of all feasible plans is

$$F = \{x : Ax = 1, \quad ||x||_1 = k, \quad x \in \mathbb{Z}_2^n\}$$

and the optimal plan $\hat{x} \in F$ maximizes some linear objective cx.

Importantly, the complexity of solving $max\ cx$ s.t. $x \in F$ scales with n, the number of districts, not $|F|$, the size of the feasible set. Therefore, we want to generate a representative set of districts, that are also *efficient*, meaning that $\frac{|F|}{n}$ is large.



Stochastic Hierarchical Partitioning (SHP)

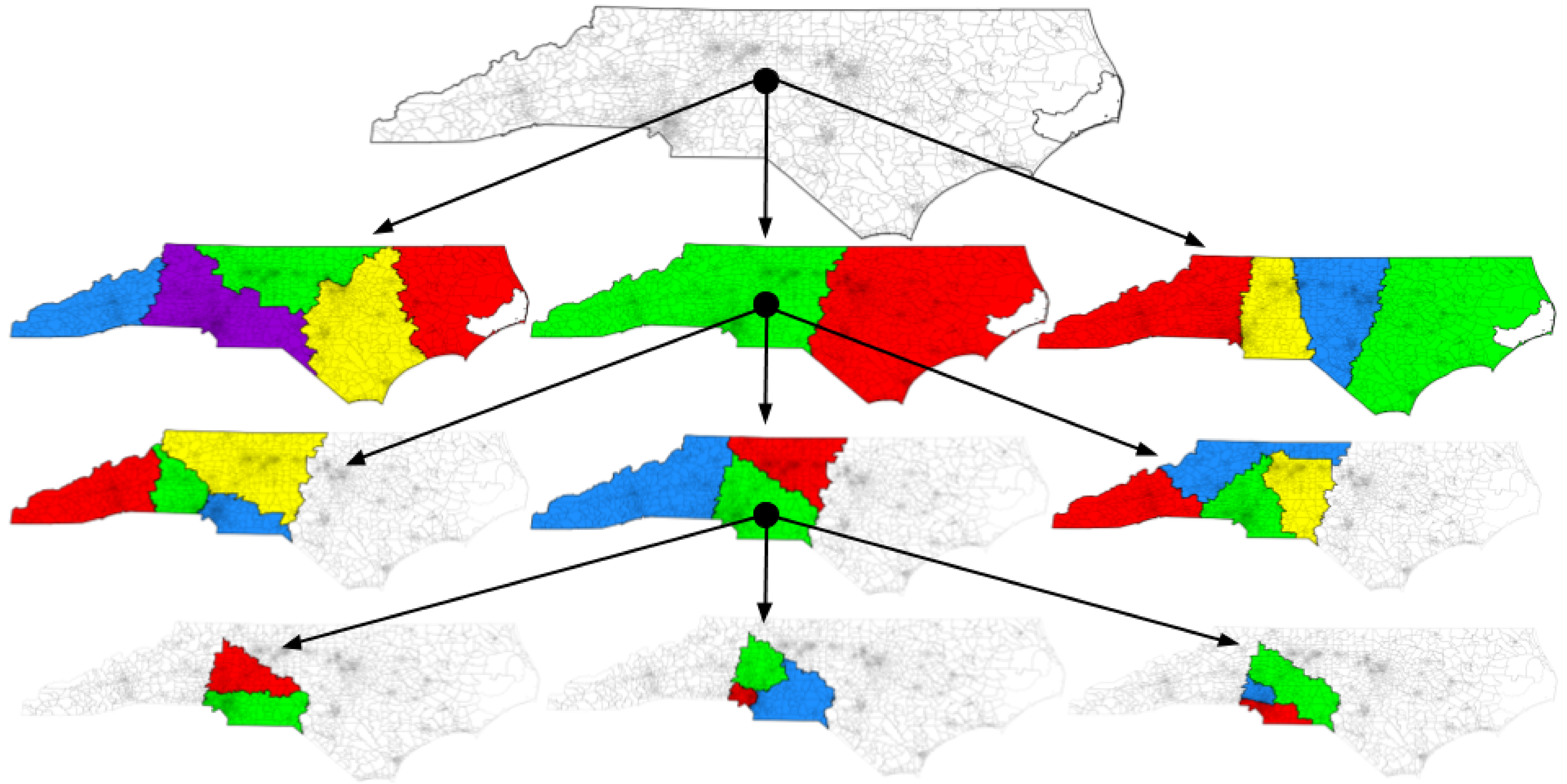
We generate efficient columns by recursively sampling region partitions and organizing these into a *sample tree*. For each partition step, we sample a split size z, sample a capacity of each center s_i, sample the position of the centers c_i, and finally solve a partitioning integer program to perform the final assignment of blocks to regions. The partition integer program (right), is essentially a transportation problem with a contiguity constraint (5). The set

$$S_{ij} = \{k : (k, j) \in E_R,\ d_{ik}^R < d_{ij}^R\}$$

is the set of neighbors of j that are closer to i.

$$\begin{aligned}
\text{minimize} \quad & \sum_{i \in C} \sum_{j \in R} (d_{ij})^\alpha p_j x_{ij} && (1)\\
\text{s.t.} \quad & \sum_{i \in C} x_{ij} = 1, && \forall j \in R \quad (2)\\
& \sum_{j \in R} p_j x_{ij} \le \hat{p}(s_i + \epsilon_p) && \forall i \in C \quad (3)\\
& \sum_{j \in R} p_j x_{ij} \ge \hat{p}(s_i - \epsilon_p) && \forall i \in C \quad (4)\\
& \sum_{k \in S_{ij}} x_{ik} \ge x_{ij}, && \forall i \in C, \quad \forall j \in R \quad (5)\\
& x_{ij} \in \{0, 1\}, && \forall i \in C, \quad \forall j \in R \quad (6)
\end{aligned}$$

This enforces that a district is a subtree of the shortest path tree rooted at block i (ensuring contiguity) [2]. Center blocks are sampled iteratively with probability proportional to the product of the distances to the already sampled centers. We continue sampling region partitions until $s = 1$, which yields a legal district.



A sample tree of North Carolina with sample width $w = 3$ and split sizes $z \in [2, 5]$.

Theorem 1 *Consider a set of blocks B to be partitioned into k districts. For a sample tree with root node (B, k) and with nodes corresponding to distinct partitions, with constant sample width w, and arbitrary split sizes $z' \in [2, z]$, the tree admits $P(B, k)$ total distinct partitions where*

$$w^{\frac{k-1}{z-1}} \le P(B, k) \le w^{k-1}.$$

Master Selection Problem (MSP)

The second step is to optimize over the generated districts to select the k districts of the optimal solution. We solve a MSP once for every root partition because columns without shared parents are unlikely to be compatible. This makes the whole pipeline arbitrarily parallelizable because root partitions can be generated and optimized independently. We can also add arbitrary linear constraints to encode additional legal requirements for the Voting Rights Act or state specific rules.

$$\begin{aligned}
\text{minimize} \quad & \left| \sum_{j \in D} c_j x_j \right| && (1)\\
\text{s.t.} \quad & \sum_{j \in D} a_{ij} x_j = 1, \quad \forall i \in B; && (2)\\
& \sum_{j \in D} x_j = k; && (3)\\
& x_j \in \{0, 1\}, \quad \forall j \in D. && (4)
\end{aligned}$$

Objective Function

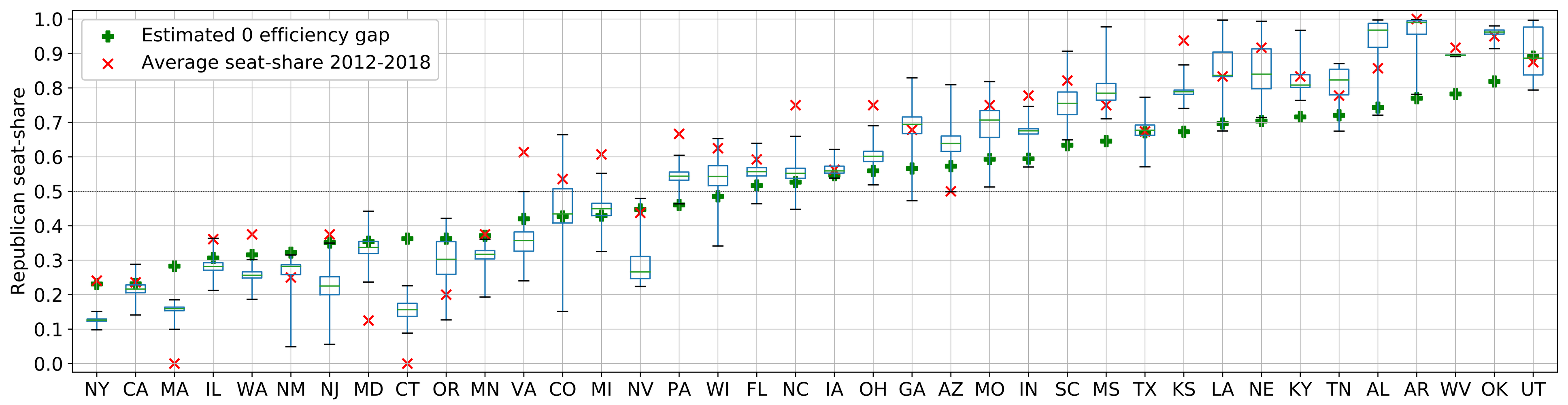
Using historical precinct returns for statewide elections, we can estimate the probability that an arbitrary district will have a greater proportion of Democrats or Republicans: $\nu_i \sim \mathcal{N}(\mu_i, \sigma_i^2),\ \psi_i \sim \mathcal{B}(P(\nu_i > .5))$. By linearity of expectation, the expected difference between the statewide seat-share and statewide vote-share is the sum of the differences between the expected district-level seat-share and statewide vote-share:

$$E\left[\frac{1}{k}\sum_{i=1}^{k} \hat{\nu} - \psi_i\right] = \frac{1}{k}\sum_{i=1}^{k} \hat{\nu} - \left(1 - \Phi\left(\frac{\mu_i - .5}{\sigma_i}\right)\right).$$

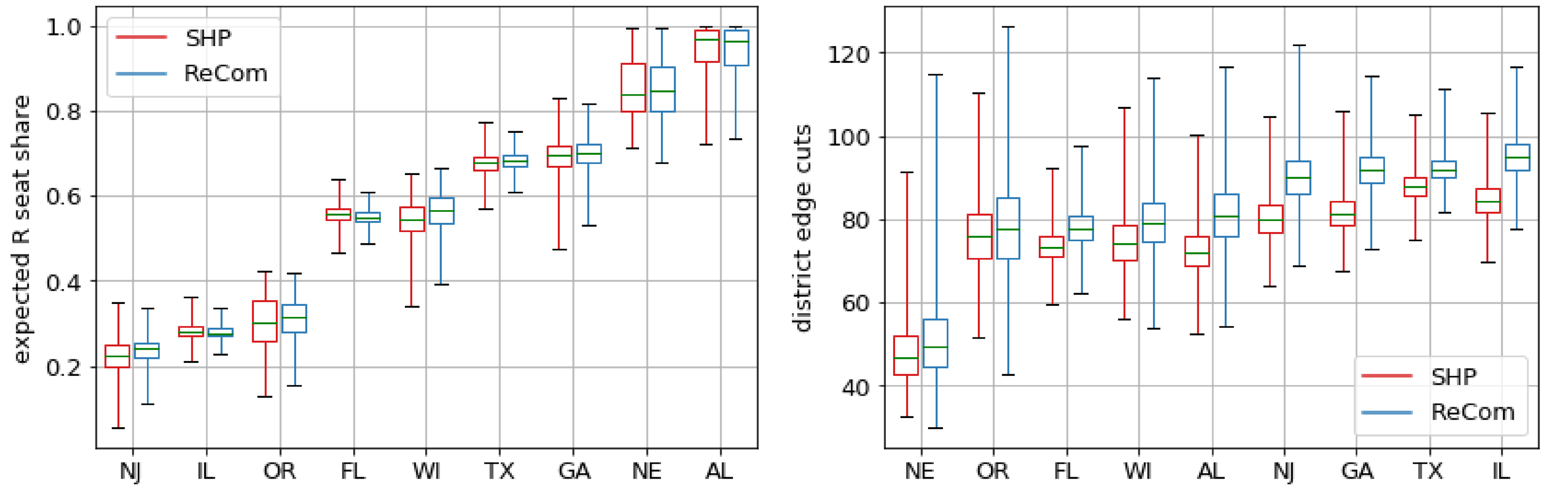
Importantly, we can also minimize $E(h(v_i) - \psi_i)$, enabling arbitrary ideal mappings of seats to votes. The efficiency gap (EG) [3], a popular fairness metric, measures the difference between wasted (surplus or losing) votes for the two parties. Assuming uniform turnout, the efficiency gap assumes an ideal mapping $h(v) = 2v - 0.5$.

Results

We run our algorithm on all 43 multi-district states and compare the distribution of partisan outcomes with the average partisan composition of the past decade and the point that would minimize the expected efficiency gap.



Compared to the standard quantitative tool in redistricting, recombination Markov chains [1], our method generates plans with a wider partisan range while better maintaining district compactness, especially in larger states.



Our results show that with just using *natural* districts, those that are of a reasonable shape and neutrally generated, we can change the partisan composition of the House of Representatives by about 20%. Furthermore, we demonstrate the efficacy of our decoupled design and scalability of our hierarchical generation method.

References

[1] Daryl DeFord, Moon Duchin, and Justin Solomon. Recombination: A family of Markov chains for redistricting. *arXiv preprint arXiv:1911.05725*, 2019.

[2] Anuj Mehrotra, Ellis L Johnson, and George L Nemhauser. An optimization based heuristic for political districting. *Management Science*, 44(8):1100–1114, 1998.

[3] Nicholas O Stephanopoulos and Eric M McGhee. Partisan gerrymandering and the efficiency gap. *U. Chi. L. Rev.*, 82:831, 2015.