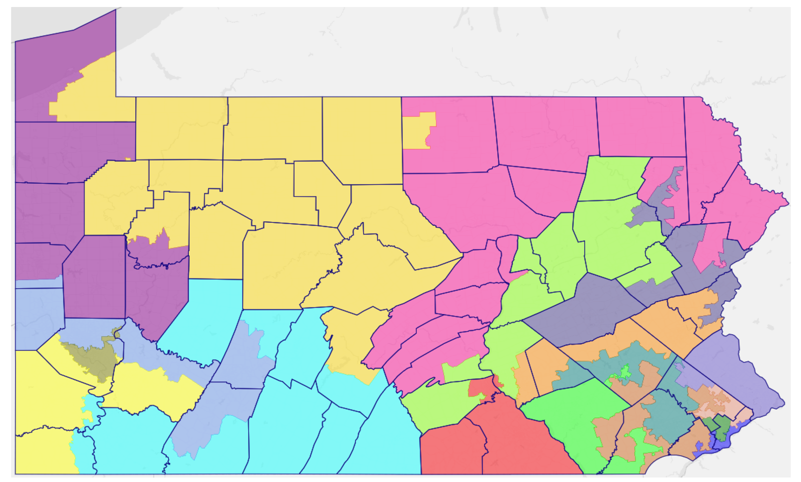
Political Districting to Minimize County Splits

Maral Shahmizad   Austin Buchanan

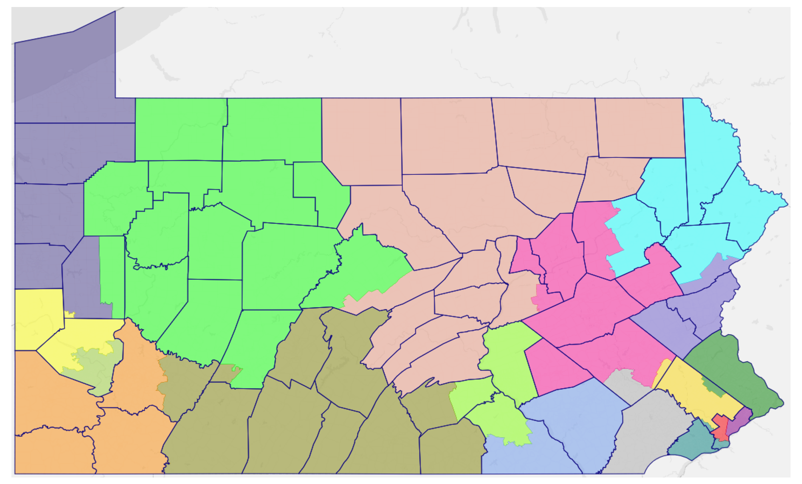
Abstract

When partitioning a state into political districts, a common criterion is that political subdivisions like counties should not be split across multiple districts. This criterion is encoded into most state constitutions and is sometimes enforced quite strictly by the courts. However, map drawers, courts, and the public typically do not know what amount of splitting is truly necessary. In this paper, we provide answers for all congressional, state senate, and state house districts in the USA using 2020 census data. Our approach is based on integer programming. The associated codes and experimental results are publicly available on GitHub.

League of Women Voters v. Pennsylvania (2018)



(a) Pennsylvania's overturned districts
28 split counties, 37 county splits



(b) Pennsylvania's court-manded districts
13 split counties, 17 county splits

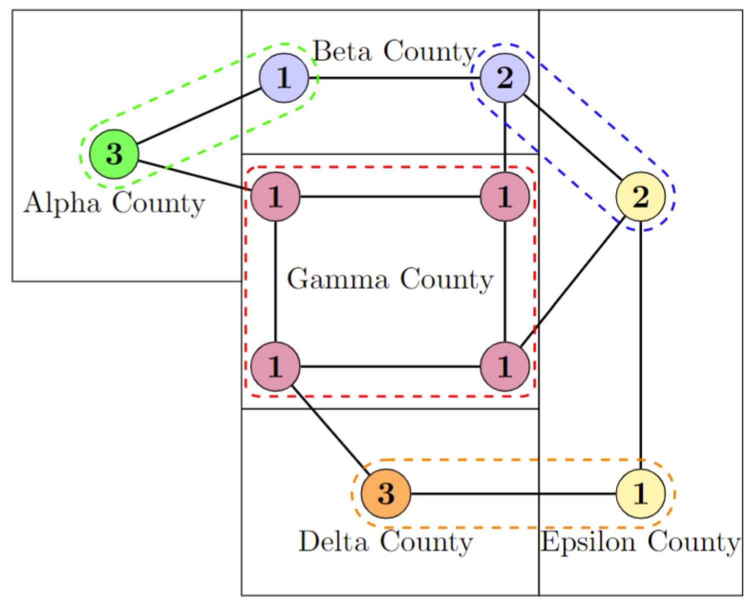
Definition (County Clustering)

A county clustering is a partition $(C_1, C_2, \ldots, C_c)$ of the counties along with associated cluster sizes $(k_1, k_2, \ldots, k_c)$ such that:

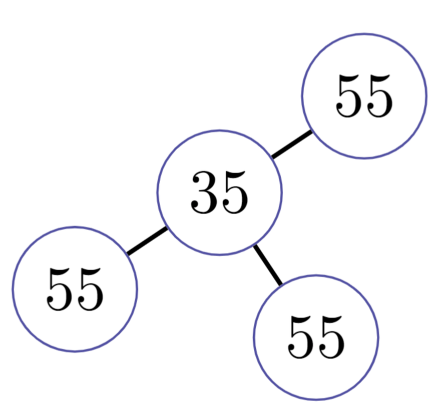
1. the cluster sizes $(k_1, k_2, \ldots, k_c)$ are positive integers that sum to k;
2. each cluster C_j is contiguous, i.e., induces a connected subgraph of G_C;
3. each cluster C_j satisfies population balance, i.e., $Lk_j \leq p(C_j) \leq Uk_j$.

Theorem (Carter et al., 2020)

We have `(min # splits)`$= k -$`(max # clusters)`, "except in rare circumstances".



(a) Example: Splitigan has `(min # splits)`$= 2$, $k = 4$ districts, and `(max # clusters)`$= 2$



(b) Counterexample(!): this claw instance requires more than $k-1$ county splits

Definition (Split Duality)

A districting instance exhibits **weak split duality** if `(min # splits)`$\geq k -$`(max # clusters)`. It exhibits **strong split duality** if `(min # splits)`$= k -$`(max # clusters)`.

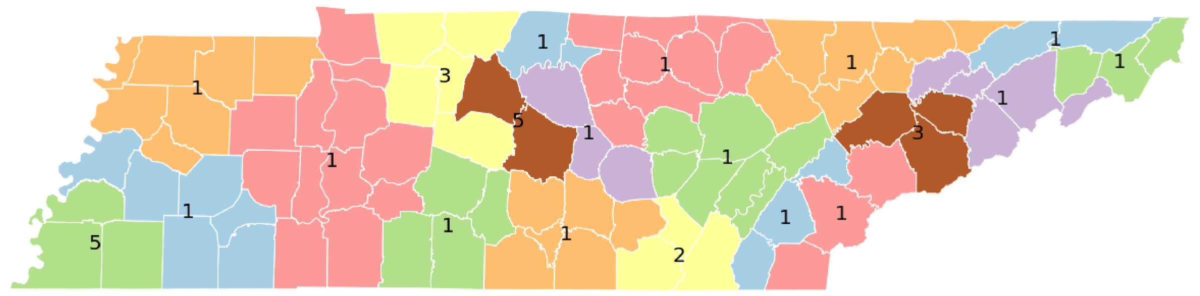
Methodology

Theorem (Weak Split Duality)

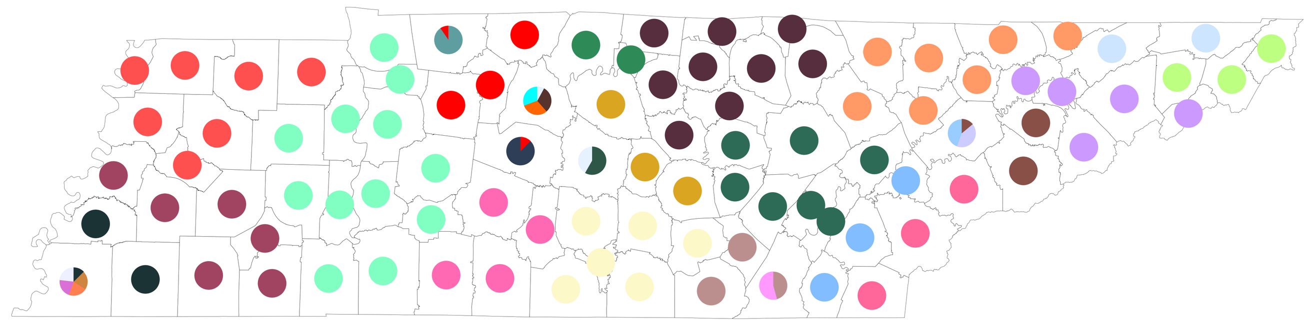
Weak split duality always holds. Strong split duality does not always hold.

Our approach exploits weak split duality and has three steps (each solved with IP techniques).

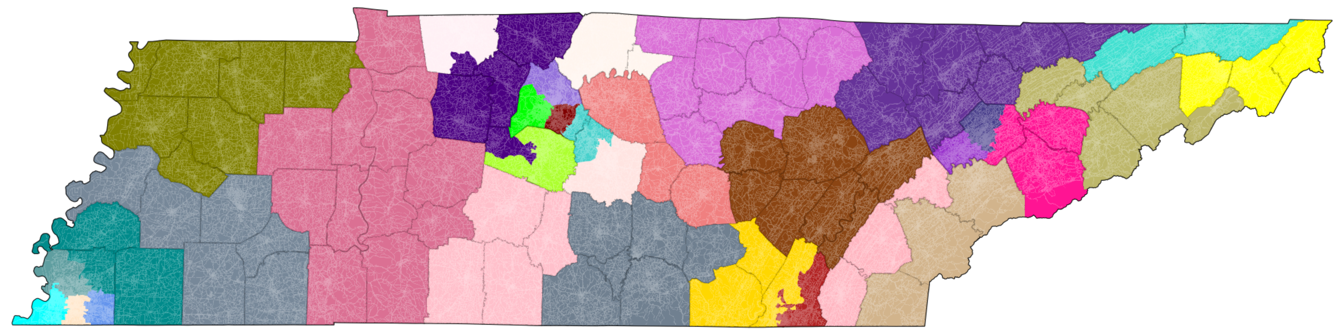
1. *Cluster*. Partition the counties into a maximum number of county clusters $(C_1, C_2, \ldots, C_c)$ with associated cluster sizes $(k_1, k_2, \ldots, k_c)$.



2. *Sketch*. For each cluster C_j, sketch a districting plan for it that has k_j districts and $k_j - 1$ county splits.



3. *Detail*. For each cluster C_j, find a detailed districting plan that abides by the sketch's *support*.



If step 1 is successful, then $k - c$ splits is a lower bound. If steps 2 and 3 are successful, then $(k_1 - 1) + (k_2 - 1) + \cdots + (k_c - 1) = k - c$ splits is an upper bound.

Step 1: MIP for *Cluster* (see paper for Steps 2 and 3)

$$
\begin{aligned}
\max \quad & \sum_{j \in C} x_{jj} & \\
\text{s.t.} \quad & \sum_{j \in C} x_{ij} = 1 & \forall i \in C \\
& \sum_{j \in C} y_j = k & \\
& C_j = \{i \in C \mid x_{ij} = 1\} \text{ is connected} & \forall j \in C \\
& L y_j \leq \sum_{i \in C} p_i x_{ij} \leq U y_j & \forall j \in C \\
& x_{ij} \leq x_{jj} & \forall i, j \in C \\
& x_{ij} \in \{0, 1\} & \forall i, j \in C \\
& y_j \in \mathbb{Z}_+ & \forall j \in C.
\end{aligned}
$$

Theorem (Rounding Inequalities)

Let t be a positive integer, and let j be a county. The following rounding inequality is valid.

$$\sum_{i \in C} \left\lfloor \frac{t p_i}{U + 1} \right\rfloor x_{ij} \leq t y_j - x_{jj}.$$

Results: All instances solved to optimality!

Table 1. For each state and district type (congressional, state senate, state house), what is the maximum number of county clusters (c), minimum number of county splits (s), and the enacted number of county splits (s_e)? We impose a $\pm 0.5\%$ deviation for congressional instances and a $\pm 5\%$ deviation for legislative instances.

		Congressional				State Senate				State House			
state	$\lvert C \rvert$	k	c	s	s_e	k	c	s	s_e	k	c	s	s_e
AL	67	7	7	0	6	35	19	16	35	105	29	76	115
AZ	15	9	2	7	15	30	6	24	44	30	6	24	44
CA	58	52	11	41	72	40	14	26	56	80	20	60	95
CO	64	8	6	2	20	35	13	22	42	65	18	47	73
FL	67	28	9	19	31	40	16	24	32	120	26	94	112
GA	159	14	12	2	21	56	31	25	60	180	57	123	209
IL	102	17	8	9	53	59	20	39	135	118	31	87	220
IN	92	9	8	1	8	50	28	22	48	100	39	61	129
LA	64	6	6	0	15	39	18	21	77	105	29	76	116
MA	14	9	2	7	22	40	6	34	59	160	10	150	182
MD	24	8	4	4	9	47	10	37	45	47	10	37	67
MI	83	13	9	4	21	38	18	20	64	110	32	78	154
MN	87	8	6	2	12	67	26	41	100	134	34	100	176
MO	115	8	7	1	10	34	20	14	16	163	47	116	137
NC	100	14	11	3	13	50	28	22	24	120	40	80	80
NJ	21	12	3	9	20	40	10	30	56	40	10	30	56
NY	62	26	8	18	26	63	20	43	66	150	26	124	179
OH	88	15	11	4	14	33	20	13	20	99	35	64	77
PA	67	17	10	7	17	50	23	27	47	203	39	164	186
SC	46	7	6	1	10	46	16	30	68	124	24	100	145
TN	95	9	7	2	11	33	20	13	15	99	36	63	74
TX	254	38	19	19	59	31	19	12	41	150	50	100	101
VA	133	11	9	2	11	40	24	16	34	100	38	62	98
WA	39	10	6	4	11	49	13	36	59	49	13	36	59
WI	72	8	7	1	13	33	20	13	73	99	30	69	159

Conclusion

- Preserving political subdivisions is a key traditional redistricting principle; it is encoded into many states' laws; it may inhibit gerrymandering
- In most cases, the minimum number of county splits was not previously known, but we can find it with integer programming techniques
- We apply to congressional, state senate, and state house districting
- Strong split duality *does* hold in practice
- Many states' districting plans divide counties much more than necessary
- Disclaimer: We make *no* claims that the generated maps are "good" or legal

References

[1] Daniel Carter, Zach Hunter, Dan Teague, Gregory Herschlag, and Jonathan Mattingly.
Optimal legislative county clustering in North Carolina.
Statistics and Public Policy, 7(1):19–29, 2020.

[2] Daryl DeFord, Moon Duchin, and Justin Solomon.
Recombination: A family of Markov chains for redistricting.
Harvard Data Science Review, 3(1), 2021.

[3] Cory McCartan and Kosuke Imai.
Sequential Monte Carlo for sampling balanced and compact redistricting plans.
arXiv preprint arXiv:2008.06131, 2020.

[4] John Nagle.
Euler's formula determines the minimum number of splits in maps of election districts.
Available at SSRN 4115039, 2022.