Probabalistic Agglomerative Clustering for Redistricting

Matt Lawhon
Columbia University
New York, NY
matthew.lawhon@columbia.edu

Reetahan Mukhopadhyay
Columbia University
New York, NY
reetahan.m@columbia.edu

Pierre Tessier
Columbia University
New York, NY
pierre.tessier@columbia.edu

Ben Yang
Columbia University
New York, NY
by2297@columbia.edu

Abstract—As the impact of gerrymandering on the U.S. political system has come to light over the past 20 years, there has been an increasing focus on how we can quantitatively identify and describe gerrymandering [1]. The foremost method used in the courts and preferred by academics is known as *outlier analysis*, in which a proposed redistricting is compared to a representative sample of redistrictings on a variety of metrics, none of which are alone sufficient indicators of gerrymandering [2], [3]. A leading issue in using this method is generating a representative sample given the size of the sample space (in NC estimated to be on the order of 10^{37})[1]. Though the standard methods consist of variations of Markov Chain Monte Carlo sampling, this is computationally expensive [4]. In this paper we propose improved clustering algorithms to efficiently generate a representative sample of redistrictings for use in outlier analysis, and demonstrate their success in replicating underlying redistricting space distributions on key gerrymandering indicator metrics.

I. INTRODUCTION

A common format of representative democracy is to divide areas into districts such that each district elects a representative. A notable example is the U.S. House of Representatives, in which states are divided into districts of roughly equal population that all elect a single representative. In forming the U.S. House of Representatives, the districts must be periodically redrawn as populations change. The redrawing of districts is known as *redistricting* and it has long been known to be vulnerable to *gerrymandering*, a catch-all term referring to the act of redrawing district lines in such a way as to meet some other goal than political representation of the population. In particular, this has historically been used to undermine the political agency of minority racial and political groups in the U.S.

In *Vieth v. Jubelirer* (2004) was a U.S. Supreme Court Ruling in which the court upheld that the alleged political gerrymandering was not unconstitutional. Justice Kennedy noted, however, that if "new technologies" were developed to quantify and thus more precisely describe "the burdens gerrymanders impose," the court would find partisan gerrymandering claims justiciable [5]. This reignited an interest in quantifying and precisely describing gerrymandering so as to meet the burdens of proof required by the U.S. judicial system. As such, much of the work done in this field is recent, particularly the work requiring intense computation [6].

A. Redistricting Rules

Before examining districts for indications of gerrymandering, we must first understand what constraints are imposed on redistrictings by law. The majority of redistricting law is defined at the state level, though there are notable exceptions to this. One of the most important federal laws governing the redistricting process is preservation of minority voting blocks as per the Voting Rights Act of 1965 [7]. In effect, this may interpreted as stating that "if there is racially polarized voting and if a minority opportunity district can be drawn, then it must be drawn." A notable result of this is the legal requirement that Native American reservations and communities be grouped in the redistricting process.

The discussion state-level redistricting laws is complicated by the lack of consistency between states on these laws. The major requirements shared between states include considerations for population balance, contiguity, and in some case blindness to certain demographic features of the populations including registered political party.

B. Identifying Gerrymandering

A redistricting being legally valid certainly does not mean that it is fair, or free of gerrymandering. Further, it is still perfectly legal for states to politically gerrymander their congressional maps. Defining more precise notions
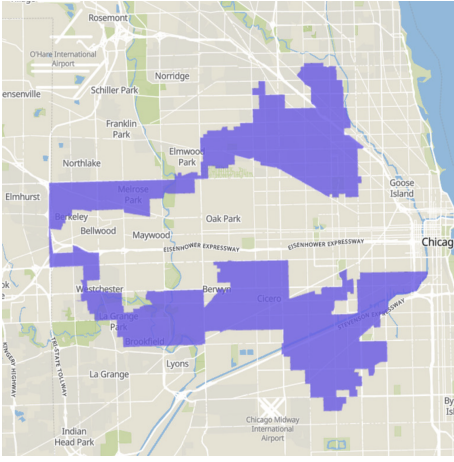
Figure 1. Illinois' district 4, that connects two large hispanic communities by a single stretch of highway, appears gerrymandered. The district's specific shape is largely a product of the VRA's requirement that minority voting bases not be split up via racial gerrymandering [8]
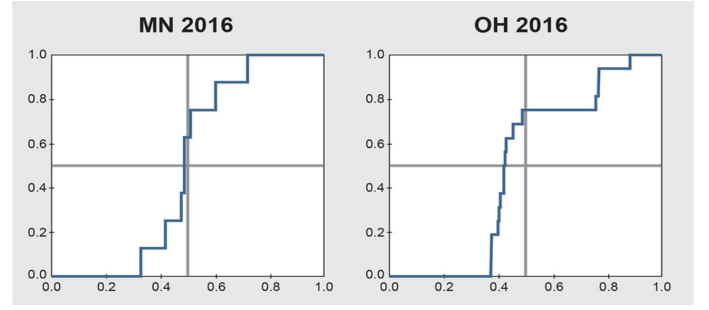

Figure 2. This indicates that Ohio's congressional map is gerrymandered, because across a wide variety of the possible proportion of seats won (30% to 70%), republicans have systematically higher representation in the legislature than democrats [1].

of fairness is one the central object of study related to gerrymandering, and the basis of the techniques proposed in this paper [2]. In this section we will explore a few important metrics that have been used to quantify gerrymandering.

1) Compactness: A simple measure of whether a map is gerrymandered can be derived from the intuition that often in order to gerrymander, one has to draw non-compact districts. Following this intuition, measuring compactness has been a central subject of focus in quantifying gerrymandering for decades. Among the standard compactness indexes are:

1) Skew measure = maximal length/perpendicular width.
2) Polsby-Popper = $4\pi \times$ Area / (Perimeter)^2
3) Reock = Area / (Area of smallest encompassing circle)

Compactness index values are known to be inconsistent particularly for states with coastlines, and non-compactness by design. The former point is illustrated by considering that coastline distances (which are not well defined to begin with), will artificially inflate compactness measures and accounting for this to preserve consistency of the metric is non-trivial. In the latter case, compactness must be violated often to preserve minority voting blocks (Fig 1).

2) Symmetry: By symmetry, we formalize the intuition that if party A gets $x\%$ of the vote and winds up with y seats, then if party B gets $x\%$ percent of the vote, it should also wind up with y seats. Thus we

can fully investigate symmetry in the two party system by, for party A, graphing vote percentage won vs seats won. See Fig 2 for symmetry graphs based on the 2016 redistrictings of Minnesota and Ohio, with proportion of R votes on the x axis, and proportion of R seats won on the y axis.

To calculate seats won for all possible values of vote percentage won, however, we must make assumptions about how the vote percentage increases in each district. Normally, this is done by assuming that representation of party A increases proportionally in all districts, in what is referred to as the *uniform partisan swing* assumption. This assumption is almost universally empirically false - voting population changes are much more complicated and linked to the particularities of state-level political geography and the associated population dynamics.

3) Proportionality: In proportionality, we measure the number of seats won by party A given the proportion of votes gathered by party A overall. Though perfect proportionality should not necessarily be the goal in redistricting, there is a reasonable range that is often sought out. The specifics of this are highly dependent on state political geography [9].

4) Efficiency gap: The *efficiency gap* metric aims to quantify the observation that one of the chief ways in which gerrymandering can occur is by means of forcing a losing parties' votes to be wasted, where a *wasted vote* is any vote cast for a losing candidate, or any vote beyond the required 50% majority required for a winning candidate. This is highly related to the notion of gerrymandering via *packing* and *cracking,* in which a controlling party minimizes the representation of the non-controlling party by creating districts that maximize the losing parties wasted votes. Namely, it maximizes the number of votes for a losing candidate, by *cracking*

non-controlling party populations into districts they have minority representation in, and maximizing the number of votes for a winning candidate beyond the required 50% by *packing* non-controlling party populations into districts.

The precise metric is as follows where LV denotes losing party wasted votes, WV denotes losing party wasted votes and TV denotes the total number of votes:

$$EG = \frac{LV - WV}{TV}$$

The efficiency gap is a new metric that has been featured prominently in many gerrymandering court cases within the past 5 years. The efficiency gap has been criticized for being unable to account for nuances within the behavior patterns of voters, including swing and third-party voters, being very dependent on non-gerrymandered features of local political geography, and being overly sensitive to minor redistricting changes [10].

C. Outlier Analysis

The current state of the art methodology for identifying Gerrymandering is *outlier analysis* [3]. Observing that all of the aforementioned metrics are insufficient indicators alone and heavily dependent on state-specific norms, in outlier analysis we present an ensemble approach to identifying maps as gerrymanders by capturing these state-specific norms. Outlier analysis consists of first drawing a representative sample from the space of redistrictings of adequate size, and then comparing a proposed redistricting to the sample across all the aforementioned metrics and more. Following this, at the discretion of human experts, an evaluation of the results of an outlier analysis can be used to make an informed decision about whether a proposed redistricting is gerrymandered.

II. SAMPLING FOR OUTLIER ANALYSIS

Of chief importance in the application of this method, then, is achieving a sufficiently representative sample from the redistricting space. This is not a statistically simple problem. Redistricting spaces can be extremely large. North Carolina, whose population is approximately 11'000'000 and thus ranks at the 9th most populous state, has an estimated redistricting state space size of 10^{37} [1]. In this paper, we review existing sampling methodologies and ultimately propose a novel approach using enhanced probabilistic clustering algorithms to sample from redistricting spaces.

A. Markov Chain Monte Carlo SoA

The State of the Art in methods of sampling from the redistricting space are MCMC (*Markov Chain Monte Carlo*) methods. Using notation provided by [11], we can view any state and the building blocks of any redistricting as a graph $G = (V, E)$ where each vertex corresponds to some area in the state and two vertices are adjacent in G iff their areas are adjacent on the map. Then a redistricting can be viewed as a k-partition of G: a collection of k disjoint subsets of V such that their disjoint union is V. Denote the set of all k-partitions as $\mathcal{P}_k(G)$, and the set of all valid k-partitions as $\mathcal{P}_k(G)*$.

To sample from $\mathcal{P}_k(G)$ using MCMC methods, we need to define a transition probability function over of the $\mathcal{P}_k(G)$ state space. Formally, this we need to define $\forall P, Q \in \mathcal{P}_k(G)$,

$$X_P(Q) : \mathcal{P}_k(G) \times \mathcal{P}_k(G) \rightarrow [0, 1]^{\mathcal{P}_k(G)} \qquad (1)$$

where $\sum_{Q \in \mathcal{P}_k(G)} X_P(Q) = 1$. In this application, we work specifically with *Ergodic Markov Chains*: Where s_t denotes the Markov chain state at time t, a Markov chain is Ergodic if $\exists T_0 : \forall i, j \in \mathcal{S}$, where $s_0 = i$, $\forall t > T_0, \Pr[s_t = j] > 0$

If the Markov chain associated with our redistricting space is Ergodic, then by the Fundamental theorem, there exists a unique stationary probability distribution of redistrictings. Further, we can effectively sample from this distribution by beginning from any state and iteratively applying the transition step some number of times, T, that will be depend on some parameter ϵ defining the required similarity between our sampling distribution and the true stationary distribution. This number T is known as the *mixing time* and is notoriously difficult to compute for arbitrary ϵ.

The SoA approaches in MCMC sampling of the redistricting space provide $\text{poly}(n, 1/\epsilon)$ guarantees of mixing time (where n is the number of building blocks for a given redistricting) and $O(n)$ guarantees of transition step run-time. The specifics of such of approaches are out of the scope of this paper. [11]

III. OUR APPROACH

Beyond the MCMC approaches, there have been limited attempts at applying clustering methods to generate redistrictings. The vast majority of these approaches have been studied outside of the context of outlier analysis, and provide no attempts to demonstrate practical or theoretical guarantees of the algorithms regarding their representativeness [12], [13], [14].

The papers that do explore the efficacy of clustering algorithms towards taking representative samples from the redistricting space for use in outlier analysis find that clustering methods are outperformed by MCMC methods for all metrics (where performance is closely approximating the distribution of a metric in a known redistricting space of small size) [15], [16].

A. Summary

The goal of our project is to generate a sample distribution of districting plans, where each districting plan has contiguous precincts and also has roughly the same population across all districts. After we have created a sample distribution of districting plans, we can then evaluate the entire sample distribution's metrics like efficiency gap and declination. Then, we can use outlier analysis to determine if any new districting plan is at low probability within our sample distribution of evaluated metrics. If that were to happen, then we have some evidence that the given districting plan may be gerrymandered.

B. Algorithm

As input, we take a file that contains the total amount of precincts p for a given region, each precinct i's voting age population, $VAP(i)$, a the target number of districts d, and a population deviance tolerance ρ, and a timeout time t. We then compute the ideal population per district T, which is simply the total population in the region divided by the target number of districts. Our final districts must have a population in $[T(1-\rho), T(1+\rho)]$. Next, we compute a graph G where each precinct is a node, and edges are drawn between nodes if the precincts share a border on the state map.

Initially, we set each precinct as its own district. Next, while our number of current districts is more than our target number of districts, we calculate the dissimilarity matrix upon our current set of districts. For each pair of districts i, j, we calculate a heuristic. $D(i, j) = T - (VAP(i) + VAP(j))$ if i, j are adjacent nodes, and $D(i, j) = 0$ if they are not. If $T - (VAP(i) + VAP(j)) < 0$ (the hypothetical merging of the districts would lead to a district population greater than the target), but not less than $-T\rho$, then we set $D(i, j) = \frac{-T}{T - (VAP(i) + VAP(j))}$. Finally, if $T - (VAP(i) + VAP(j)) < -T\rho$, then we set $D(i, j) = 0$. Then, we normalize the dissimilarity matrix, dividing each entry by the sum of all entries in the matrix: giving us a probability distribution. From these generated probabilities, we pull a random pair of edges i, j using the

effective probability distribution, which tells us to merge the districts (nodes) i and j. We update the graph by contracting the nodes representing the merged districts. We then update the dissimilarity matrix after removing the rows and columns corresponding to the two old districts and adding a row and column corresponding to the new district.

At the final merge, that gives us d districts, we check to see that each district is in $[T(1-\rho), T(1+\rho)]$. If this check is passed, we add this configuration as a sample districting plan to our distribution. If we fail this check, we have implemented the merging process recursively, so we backtrack recursively to allow us to randomly select a different pair of districts to merge, and try again. Every time all possible merge options are exhausted at a given level of current merge count, we go up one level in the recursion (to where one less merge had been performed), and recursively retry given a different merge selected. Thus is performed until we get a sequence of merges leading to a valid districting plan that we can write as a sample. Given our computational resources, if the entire process takes longer than timeout time t, we halt it, mark the attempt as failed, and restart fresh.

C. Theoretical Guarantees

Firstly, our algorithm is a valid method of sampling. Consider that any valid district plan must satisfy the contiguous district constraint. By only merging clusters who are adjacent in the graph, we guarantee our final districts will be contiguous. In addition, we perform a check each time the number of target districts is reached in the algorithm to ensure the population deviation constraint is being met.

However, we need to also consider that our sampling method must be representative: a representative sampling method must be capable of selecting every member of the space as a sample. In other words, our method must be capable of producing every representative districting plan as a sample plan in the space of representative districting plans - where a representative districting plan is defined as satisfying the population and contiguity constraints. In a final district plan, each district is composed of a combination of contiguous precincts. Thus, we can consider that any district in an arbitrary districting plan can be formed iteratively by combining a pair of contiguous clusters (merged smaller districts, which are initially just precincts) in the district in a given iteration. Therefore, if a representative sampling method needs to be capable of selecting every element of the districting space, our algorithm needs to be capable

of producing at least one of the possible sequences of merges that leads to each possible representative districting plan. Given that our algorithm assigns a non-zero probability of merging every pair of adjacent clusters at each iteration, we claim that our algorithm has some nonzero probability of producing every possible sequence of merges leading to representative districting plans. Therefore, we claim that our algorithm is capable of producing at least one sequence of merges leading to every possible representative districting plan with some nonzero probability. Thus, our algorithm is capable of choosing every plan in the redistricting space as a sample, making it representative.

D. Runtime

Next, let us consider the algorithm's runtime. Suppose we have target number of districts T and p amount of precincts. Merging two districts is $O(p^2)$, since it costs that amount of time to compute the 2-D dissimilarity matrix. Then, in the best case to generate a single sample, we do not need to recursively backtrack at all, meaning that we only have $p - T$ total merges, which is $O((p - T)p^2)$ total run time.

In the worst case, we backtrack all the way back up and compute all possible merges. The runtime of a single merge remains the same, but the number of merges goes up, to give us as overall runtime of $O((\frac{p!}{T!})^2)$. However, as stated, we set a time limit t on how long the algorithm can spend on a single districting plan, and if it fails to come up with one during this time - meaning it has been backtracking for too long, we considered it a failed sample. All other ones, which complete and output a sample, are successful samples. We can compute a success rate as $\frac{Successful\ Samples}{Total\ Samples}$. Given we used $t = 30$ seconds and $\rho = 20\%$, we achieved an average success rate of 75.7% for the Iowa dataset ($p = 99, d = 4$), and 88.1% for our subset of Florida dataset ($p = 25, d = 4$). This means on average, in practice, we do not recurse too far before building a valid districting plan, so the worse scenarios are usually not occurring.

As for the success rate, we can claim that if the tolerance was tighter, so ρ is smaller, then our success rate would be decrease. We would expect less samples to be able to satisfy this, so it is less likely for our algorithm to find a valid sample in the allotted time.

Note that our method can be seen as a form of hier-archical clustering, specifically agglomerative clustering - but with added non-determinism, where instead of selecting the pair of clusters that optimizes our metric to merge with certainty, we simply give that pair of clusters

the highest probability of being selected to merge, and each pair of clusters after that a probability proportional on how well they score on our metric - the heuristic.

E. Discussion

Lastly, we will make a remark about a step in the algorithm. Originally, our heuristic was $|T - (VAP(i) + VAP(j))|$. However, we quickly discovered this suffered from a case where a cluster that was too large, so more than $T(1 + \rho)$, would be the furthest from the target, by absolute value. Thus, we experienced a runaway effect where the largest cluster that was well over the desired target would keep getting selected to merge with a neighboring cluster, and creating an invalid plan. Thus, we saw a need to insert a safeguard to prevent clusters forming that would be over $T(1+\rho)$. Note while this prevents districts being over the tolerance from the target being formed, it does not stop districts that are under $T(1 - \rho)$ being formed, or being in scenarios where there are no possible merges that can be made because every possible remaining merge would create a cluster too large. However, we also realized we wanted to encourage clusters smaller than the target to be more likely to be selected compared to those over the target (that are still in a viable range), and the more over the target the merge would be, the less likely it should be picked - but all merges that would still fall in the valid range need a nonzero probability of occurring. So, our solution was what we have in the algorithm currently: if $-T\rho < T - (VAP(i) + VAP(j)) < 0$, then set $D(i,j) = \frac{-T}{T-(VAP(i)+VAP(j))}$. This generally worked, as can be seen with our results.

However, this is not ideal, because it will assign merges that would be slightly over the target with much higher probabilities than those slightly under the target. Thus, we propose a fix that currently remains untested but would be evaluated with further work on the topic. First, fill in the dissimilarity matrix as stated in the algorithm, but just leave the values $-T\rho < T - (VAP(i) + VAP(j)) < 0$ as their negative values $T-(VAP(i)+VAP(j))$, initially. Once the dissimilarity matrix is filled in, find the minimal positive value α. Now, for each $D(i,j) < 0$, recompute it as $D(i,j) = \frac{-\alpha}{D(i,j)}$. Since the population differences must be integers, thus every entry of D is an integer, we know that we cannot have $-1 < D(i,j) < 0$, ensuring that the largest possible probability a merge over the target will receive will be equal to the probability of the largest merge under the target. This choice also still maintains that the greater

over the target a merge would be, the less likely it would be picked.

IV. EXPERIMENTS

To evaluate our approach we used the data and tools provided by [16] in their *R* package dedicated to redistricting, called *redist*. We used two of their validation maps, Iowa, with 99 counties and 4 districts, and a subsample of Florida, with 25 precincts and 4 districts. For both we used a population deviation constraint of 20%; we couldn't accurately evaluate our method for lower deviation tolerance due to our limited computational resources.

We compared our method with three different MCMC variations, considered as the current standard, and with an estimation of the *ground truth*. Indeed, as proposed in [16], to evaluate our sampling algorithm we compare it with a complete enumeration of the possible redistricting plans, or at least with a large enough uniformly independently distributed set of samples. To get such a baseline, we used the *enumpart* algorithm provided in *redist*; but we were only able to use it for a deviation tolerance of at least 20% due to its high computational cost. We used 10.000 samples for each MCMC variation, and 1.500 with our method for each map. As ground truth, we used a complete enumeration of all possible redistricting plans of the Florida map, and 1.000.000 i.i.d. samples from Iowa which we then filtered with our population deviation constraint to have ∼25.000 samples.

Once all those samples have been collected, we computed the *Efficiency Gap* on each sample and compared the distribution of this metric for each sampling algorithm. By approximating the density function of those results we get the Figures 3 and 4. We also tried computing the KL-divergence of those density approximations, but the results are a bit unstable as we only have approximations of the density functions; the results are in Table I.

From those results, our method seems to return a much more representative sampling of all the possible redistricting plans. However, a deeper analysis would be needed to conclude that with certainty, as the only available validation maps are relatively small. It would also be interesting to run the same experiments with tighter population deviation constraints, but this would require large computational resources.

V. FUTURE WORK

One direction we propose for further research is exploring the use of such clustering methods in conjunction with MCMC methods. Proving rigorous bounds on
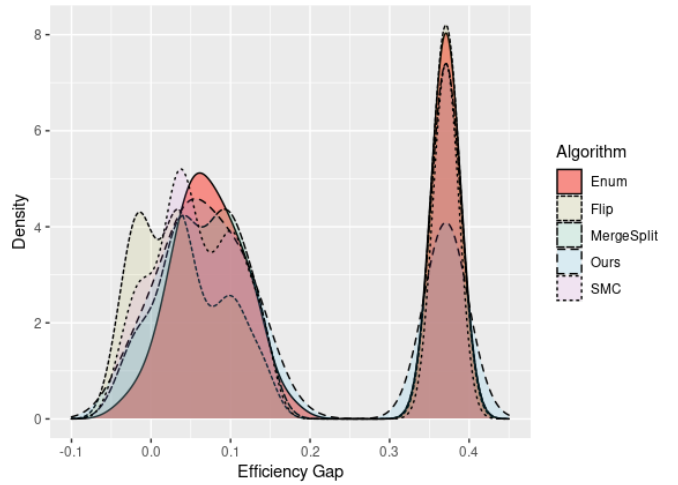

Figure 3. Efficiency gap distribution across all sampled redistricting plans for 25 precincts of Florida with 4 districts. The target distribution is *Enum*, which is a complete enumeration of all valid plans. Our method approximates better the left side of the distribution but underestimates the right side.
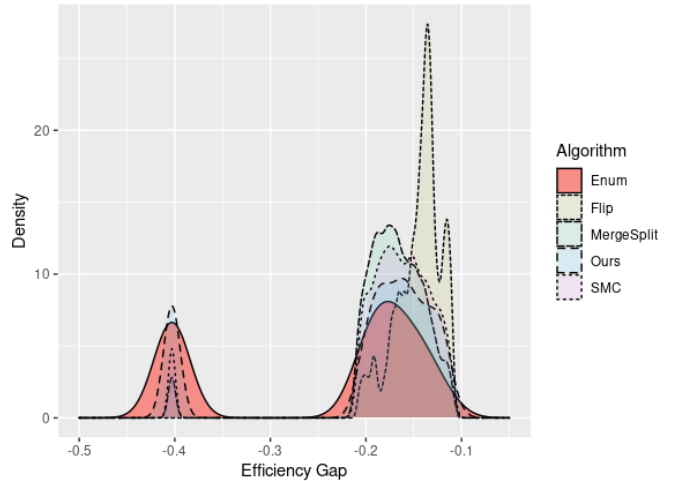

Figure 4. Efficiency gap distribution across all sampled redistricting plans for Iowa with 4 districts. The target distribution is *Enum*, which is ∼25k i.i.d. redistricting plans. Our method estimates the overall distribution much better, all the other algorithms underestimate the left side of the distribution.

mixing time is known to be difficult, so it would be very difficult to show that clustering-based redistricting plan seeding for MCMC methods reduces mixing time. Intuitively speaking, however, this is quite possible. An easy way of testing the efficacy of this approach could be in the following experiment. Suppose generating a clustering using one of our explored methods takes as long as performing t transition step iterations in a SoA MCMC method. One could investigate which method

	Florida Efficiency Gap	Iowa Efficiency Gap
MergeSplit	1.53×10^{-2}	1.85
Flip	2.38×10^{-2}	2.48
SMC	2.28×10^{-2}	-
Ours	$\mathbf{4.21 \times 10^{-3}}$	**0.933**

Table I

KL-DIVERGENCE ESTIMATION OF THE EFFICIENCY GAP DENSITY BETWEEN THE ENUMERATION (GROUND TRUTH) AND THE 4 TESTED SAMPLING ALGORITHMS.

achieves a better representation of the redistricting space within that time period, and according to which metrics.

A further direction for future work would be to address emerging perspectives on the fundamental shortcomings of how we have approached quantified gerrymandering research so far. In particular, Cox and Holden find that gerrymandering research is fundamentally flawed in its necessary inducing of voters into the binary political space [17]. They find that relaxing this assumption radically changes the way that one can optimally politically gerrymander. This necessarily casts doubt on the current state of the outlier analysis technique because many of the key metrics used to evaluate redistrictings are designed to capture the known techniques used to gerrymander in the binary political setting. This suggests the need for a reformulation of gerrymandering research in which we consider the voter in some higher dimensional, continuous space, as opposed to the single dimensional binary space, as was done in [18]. They further note the need to investigate how novel voting forms like ranked choice voting that are emerging in elections around the U.S. necessitate entirely different conditions in evaluating gerrymandering.

Lastly, on a broader note, it is worth considering that the merits of this sort of research going forward is uncertain. The US Supreme Court ruled in *Rucho v. Common Cause (2019)* that Gerrymandering was a political question, and beyond the reach of the federal courts [19]. Specifically, the court found that gerrymandering, at our current level of understanding, presents a political question because it lacks "judicially discoverable and manageable standards for resolving [it]."

Fundamentally, the court found that the thus far presented partisan gerrymandering cases have left unresolved the question of whether such claims are of legal right, resolvable according to legal principles, or political questions that must find their resolution elsewhere. As a result, it remains to be seen not only what the future holds for the federal judiciability of outlier analysis and

the associated techniques, but of all forms of quantified gerrymandering research. In the meantime, however, the methods discussed in this paper remain relevant at judiciating gerrymandering at state and more local levels of governance.

REFERENCES

[1] M. Duchin. Redistricting and representation. [Online]. Available: https://www.amacad.org/news/redistricting-and-representation

[2] B. Grofman and J. Cervas, "Tools for identifying partisan gerrymandering with an application to congressional districting in pennsylvania | elsevier enhanced reader," vol. 76. [Online]. Available: https://reader.elsevier.com/reader/sd/pii/S0962629818303342

[3] D. Gold and G. Ramachandran, "Using outlier analysis to detect partisan gerrymanders: A survey of current approaches and future directions," vol. 17, no. 4. [Online]. Available: https://www.liebertpub.com/doi/epdf/10.1089/elj.2018.0503

[4] D. DeFord, M. Duchin, and J. Solomon, "Recombination: A family of markov chains for redistricting." [Online]. Available: https://hdsr.mitpress.mit.edu/pub/1ds8ptxu/release/4

[5] "Vieth v. jubelirer, 541 US 267." [Online]. Available: https://supreme.justia.com/cases/federal/us/541/267/

[6] W. K. Tam Cho and Y. Y. Liu, "Toward a talismanic redistricting tool: A computational method for identifying extreme redistricting plans," vol. 15, no. 4, pp. 351–366. [Online]. Available: http://www.liebertpub.com/doi/10.1089/elj.2016.0384

[7] Redistricting criteria: The voting rights act - public mapping project. [Online]. Available: http://www.publicmapping.org/what-is-redistricting/redistricting-criteria-the-voting-rights-act

[8] L. Scales. Why chicago's 4th congressional district — those 'earmuffs' — are about fairness and not gerrymandering - chicago sun-times. [Online]. Available: https://chicago.suntimes.com/2020/3/11/21164850/illinois-4th-congressional-district-gerrymandering-voting-rights-act-census-liliana-scales

[9] R. J. Powell, J. T. Clark, and M. P. Dube, "Partisan gerrymandering, clustering, or both? a new approach to a persistent question," vol. 19, no. 1, pp. 79–100. [Online]. Available: https://www.liebertpub.com/doi/10.1089/elj.2019.0605

[10] P. Schuck. Efficiency gap. [Online]. Available: https://ballotpedia.org/Efficiency_gap

[11] E. A. Autry, D. Carter, G. Herschlag, Z. Hunter, and J. C. Mattingly, "Multi-scale merge-split markov chain monte carlo for redistricting." [Online]. Available: http://arxiv.org/abs/2008.08054

[12] O. Guest, F. J. Kanayet, and B. C. Love, "Gerrymandering and computational redistricting," vol. 2, no. 2, pp. 119–131. [Online]. Available: https://doi.org/10.1007/s42001-019-00053-9

[13] D. Joshi, L.-K. Soh, and A. Samal, "Redistricting using heuristic-based polygonal clustering," in *2009 Ninth IEEE International Conference on Data Mining*. IEEE, pp. 830–835. [Online]. Available: http://ieeexplore.ieee.org/document/5360319/

[14] Y.-C. Hung and L.-Y. Chen, "Using intelligent clustering to implement geometric computation for electoral districting," vol. 8, no. 9, p. 369, number: 9 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2220-9964/8/9/369

[15] B. Fifield, M. Higgins, K. Imai, and A. Tarr, "Automated redistricting simulation using markov chain monte carlo," vol. 29, no. 4, pp. 715–728. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/10618600.2020.1739532

[16] B. Fifield, K. Imai, J. Kawahara, and C. T. Kenny, "The essential role of empirical validation in legislative redistricting simulation," vol. 7, no. 1, pp. 52–68. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/2330443X.2020.1791773

[17] A. Cox and R. Holden. Reconsidering racial and partisan gerrymandering | the university of chicago law review. [Online]. Available: https://lawreview.uchicago.edu/publication/reconsidering-racial-and-partisan-gerrymandering

[18] P. Diaconis, S. Goel, and S. Holmes, "Horseshoes in multidimensional scaling and local kernel methods," vol. 2, no. 3. [Online]. Available: http://arxiv.org/abs/0811.1477

[19] "Rucho v. common cause." [Online]. Available: https://www.supremecourt.gov/opinions/18pdf/18-422$_9ol1.pdf$