Quantifying Gerrymandering: Sampling the Space of Possible NC Congressional Districting Plans

Sam Eure, Jacob Shulman, Tiffany Mei, Luke Farrell
Duke University Department of Mathematics

***Abstract*— The number of possible precinct to district assignments for North Carolina's US House of Representatives Congressional Map is intractable. However, understanding the composition of this space has become essential in efforts to quantify the severity of partisan gerrymandering among implemented maps. Towards this goal we implement (1) a parallel tempering sampling procedure and (2) umbrella sampling procedure to sample the state the space of all North Carolina redistricting maps. We compare these methods to previously explored sampling procedures in this domain, provide specific implementation details, and preliminary results.**

I. INTRODUCTION

Much recent work [1][2][3] seeking to rigorously quantify partisan gerrymandering has focused on building ensembles of redistricting plans. Ensembles are formed by (1) sampling the space of all possible maps, (2) simulating the ensemble's distribution of electoral outcomes, and (3) then comparing these electoral outcomes with those of implemented redistricting plans. If an implemented plan's electoral outcome falls far outside the distribution outcomes generated by the ensemble of maps than it is considered gerrymandered. By building an ensemble, the natural geo-political landscape of the state is captured [2] and many of the other pitfalls suffered by popular methods like the efficiency gap [7] or proportionality are avoided. As a result, the ensemble method has found success as evidence in litigation challenging partisan gerrymandering including in the Rucho v. Common Cause (2018) decision.

One of the major challenges of the ensemble method lies in the first step: sampling the space of all possible maps. As the number of possible precinct to district assignments for North Carolina's US House of Representatives Congressional Map is intractably large, ensuring that the space is sampled accurately is a key challenge. The accuracy of the sampling method affects the distribution of electoral outcomes, the judgment of the magnitude of the partisan gerrymander, and subsequently the legal viability of the ensemble methodology.

The success of a sampling method thus depends on how well it can traverse the space across variable energy ranges to find a representative sample of compliant maps.

II. REVIEW OF SAMPLING PROCEDURES

A. Sampling the State Space of Maps

The sampling of NC congressional redistricting plans can be accomplished by modeling NC as a graph with precincts for nodes and where edges between nodes exist for precincts which share a border. From this graph, each partition into 13 sets of precinct nodes can be considered a redistricting plan. To make this more precise, allow $V = \{\text{set of all precincts of NC}\}$, $D = \{1, 2, \ldots, 13\}$ represent each of the 13 districts of NC. Let a redistricting plan be a function $\xi : V \to D$ which maps precincts onto district numbers. Given this notation, the i^{th} district of a redistricting plan ξ may be represented as $D_i(\xi) = \{v \in V \mid \xi(v) = i\}$.

In order to sample redistricting plans which satisfy specific criteria of population deviations between districts, compactness of district boundaries, and minimization of county splitting, a score function, S, is derived in order compare the probability of observing any two particular redistricting plans. For a redistricting plan ξ, the score of this plan is designed to be proportional to $S(\xi) \propto e^{-J(\xi)\beta}$, for a scaling constant β, where, if Φ is the space of all redistricting plans, $J : \Phi \to \mathbb{R}^+$ is a function which maps a redistricting plan onto a positive "energy" value, where J is defined as

$$J(\xi) = w_{pop}J_{pop}(\xi) + w_{iso}J_{iso}(\xi) + w_{county}J_{county}(\xi)$$

For penalty, scalar parameters w_{pop}, w_{iso}, & w_{county} (also known as *weights*) which allow for the tuning of the score function S, and J_{pop}, J_{iso}, & J_{county} are separately defined functions which calculate a population deviation energy, a compactness energy, and a county splitting energy, respectively, for a given ξ. Each of these sub-functions are defined below. In order to define J_{pop}, let $\text{pop}[D_i(\xi)]$ represent the population of $D_i(\xi)$ given the plan ξ. Using this notation,

$$J_{\text{pop}}(\xi) = \sum_{i=1}^{13} \left(\frac{pop[D_i(\xi)]}{pop[NC]/13} - 1 \right)^2$$

where the summation from 1 to 13 comes from the fact that there are 13 districts in North Carolina and $\frac{pop[NC]}{13}$ represents the ideal population of each of the 13 districts. This component of J penalizes districting plans which have large differences in populations between different districts.

Next, for any $D_i(\xi)$, let $pD_i(\xi)$ be defined as the length of the boundary $D_i(\xi)$ share with other district boundaries (i.e. the geographical boundary of $D_i(\xi)$ excluding the boundary that $D_i(\xi)$ shares with other states, such as TN, SC, & VA). Similarly, letting $area[D_i(\xi)]$ be defined as the geographical area of $D_i(\xi)$,

$$J_{iso}(\xi) = \left[\sum_{i=1}^{13} \frac{[pD_i(\xi)]^2}{\text{area}[D_i(\xi)]} \right]$$

calculates the energy attributed to a districting plan given the geographical compactness of its districts.

In order to properly consider county splitting, let C represent the set of all counties of North Carolina. We define the function $A : (\xi, C) \to \mathbb{Z}$ to be the function from $c \in C$ to the number of times c is split given plan ξ.

$$J_{county}(\xi) = \sum_{c \in C} \sum_{a=1}^{A(\xi,c)} \left| \phi_i(\xi, c)a \right| + T(a)$$

where the function $\phi_i : (\xi, C) \to \mathbb{R}$ smooths the transition in energy from no county splits to heavy county splits and which is defined as

$$\phi_i(\xi, c) = \frac{pop[\rho_{c,i}(\xi)]}{pop[c]}$$

where $\rho_{c,i}(\xi)$ represents the partition of c with the i^{th} least population given ξ, and where $T : \mathbb{N} \to \mathbb{R}$ is a function which is zero for all values less than three, and which is otherwise large enough to prevent the occurrence of a redistricting plan with more than two splits in any given county. This definition of J_{county} penalizes multiple splits of a single county more than single splits of multiple counties and smooths the transition from no county splits to county splits, preventing large fluctuations of energy when they are not desired.

Using this score function, changes to an initial redistricting plan ξ_0 can be obtained by using the Markov Chain Monte Carlo (MCMC) algorithm described in *Fifield et. al* [1] given a scaling parameter β_0. Although the method at which changes are proposed and accepted to a redistricting plan remain universal, there exists a variety of sampling schemes which can be used to generate redistricting plans given an initial starting point, a few of which are discussed in detail below.

B. Constant Temperature Sampling

here Constant temperatue sampling is a base Metropolis-Hastings sampler. The constant temperature sampling scheme samples while keeping the value of β_0 constant. This sampling scheme can be appealing for computational purposes; however, there are several disadvantages to this method. For example, if the state space one is sampling is highly multi-modal, then the sampler may have a difficult time exploring areas of the state space that are separated by large energy barriers.

C. Simulated Annealing

This sampling scheme is similar to that of constant temperature sample, however, unlike constant temperature sampling, the value of β changes gradually over time (typically from $(\beta_{start} = 0) \to (\beta_{end} = 1)$), which begins the sampling scheme under very "hot" conditions (i.e. a very low value of β) and slowly "cools down" the map. To explain this further, recall that $S(\xi) \propto e^{-J(\xi)\beta}$, and thus, for very low values of beta, changes in $J(\xi)$ have little effect on the overall score of ξ, thus allowing changes to ξ to occur more easily given the Metropolis-Hastings algorithm. Starting the redistricting simulation at such a temperature allows the initial districting plan is move far away from its starting point (i.e. the initial plan) and explore the state space in far away areas – perhaps even crossing several energy barriers that a sampler running at constant temperature would have a difficult time surmounting. Furthermore, as the temperature cools down, the value of $J(\xi)$ begins to wield more influence in the decisions of which steps are being accepted and rejected during redistricting; this forces the plan of ξ to "move" towards a plan more compliant to the criteria we have specified given our energy function J, until the plan nears a local minimum in terms of its score. These plans at local minimums are the plans which intuitively are compliant.

Although simulated annealing is a reliable sampling method for sampling, it has a few disadvantages. One such disadvantage is that it is inefficient. For example, once a local minimum has been discovered at the end of a sampling run, the plan may still be rejected due to a lack of compliance, at which point, and entire iteration of simulated annealing will have been carried out essentially for nothing. This is troubling given the possibility that slight modifications of the ξ_{end} redistricting plan may lead to the discovery of a new compliant plan, or, the redistricting plan fell into a sub-optimal local minimum when a lower energy valley may have been near by. Under these conditions the sampling scheme will reset it's β_{start} value to zero and begin exploring the state space once again, leaving this potentially promising area of the state space. Under this belief, it may be useful to have a sampling scheme which allows for the surmounting of energy barriers as well as jump from one local minimum to another. One such sampling scheme that allows this is parallel tempering, which is discussed in the following section.

III. Parallel Tempering

A. Overview of Procedure

As alluded to above, parallel tempering is a separate sampling scheme which allows for the exploration of multiple local minimums in the state space while allowing jumps over large energy barriers in a reasonable amount of time. The sampling scheme begins by modifying the initial districting plan ξ_0 using the constant temperature sampling scheme at a predetermined $n \in \mathbb{N}$ number of processors, each of which holds unique β values which range from 0 to 1.

What makes parallel tempering different from n processes of constant temperature sampling is the "swapping" of these β values across processors every period of m or so steps of the sampler (for example, after every 100 steps of the sampler, ask two processors to exchange the β values with which they are currently constant temperature sampling. Moreover, samples are only sampled from the processor currently holding $\beta_n = 1$ since this is the *coldest* processor at the time. The general sampling scheme is as follows:

1) Choose an increasing sequence of n β values which range from $\beta_0 = 0$ (or close to zero), to $\beta_n = 1$.

2) On separate processors, begin the exploration of redistricting plans using constant temperature sampling at each β value, however only try to sample plans from the processor holding $\beta_n = 1$
3) Using a predetermined period of m, propose a swapping of β values using the following algorithm:

Algorithm 1 Swapping of Beta Values

$\beta_i \sim$ discrete uniform$\{\beta_1, \beta_2, \ldots, \beta_n\}$

sample $u \sim$ uniform$(0, 1)$

set $r = \left(\frac{S(\xi_{i-1}, \beta_i)}{S(\xi_i, \beta_i)} \frac{S(\xi_i, \beta_{i-1})}{S(\xi_{i-1}, \beta_{i-1})} \right)$

if $u \leq \min\{1, r\}$ **then**

$\beta_{i-1} \rightleftharpoons \beta_i$ swap the locations of β_i & β_{i-1}

end if

Under this system, if 4 different β vales are being used to parallel temper (i.e. there is a world size of 4), then β values should be proposed to swap with each other such that if the first swap is proposed at time (s), using the notation of $P_i(\beta_j)^{(s)}$ representing the state of processor i hosting β_j at time s, if β_3 is proposed to swap with β_2 and the swap is accepted, the following change in the configuration of processors and beta values from time s to time $s+1$ will occur:

$$\begin{array}{cccc} P_0(\beta_0)^{(s)} & P_1(\beta_1)^{(s)} & P_2(\beta_2)^{(s)} & P_3(\beta_3)^{(s)} \\ \hline P_0(\beta_0)^{(s+1)} & P_1(\beta_1)^{(s+1)} & P_2(\beta_3)^{(s+1)} & P_3(\beta_2)^{(s+1)} \end{array}$$

This swapping of beta values can continue over multiple periods (i.e. a swap is proposed ever s steps of the sampler) which results in a history of β value swaps which can be visualized as the following:

$$\begin{array}{cccc} P_0(\beta_0)^{(s)} & P_1(\beta_1)^{(s)} & P_2(\beta_2)^{(s)} & P_3(\beta_3)^{(s)} \\ P_0(\beta_0)^{(s+1)} & P_1(\beta_1)^{(s+1)} & P_2(\beta_3)^{(s+1)} & P_3(\beta_2)^{(s+1)} \\ \vdots & \vdots & \vdots & \vdots \\ P_0(\beta_0)^{(2s)} & P_1(\beta_1)^{(2s)} & P_2(\beta_3)^{(2s)} & P_3(\beta_2)^{(2s)} \\ P_0(\beta_0)^{(2s+1)} & P_1(\beta_2)^{(2s+1)} & P_2(\beta_3)^{(2s+1)} & P_3(\beta_1)^{(2s+1)} \\ \vdots & \vdots & \vdots & \vdots \\ P_0(\beta_0)^{(3s)} & P_1(\beta_3)^{(3s)} & P_2(\beta_1)^{(3s)} & P_3(\beta_2)^{(3s)} \\ P_0(\beta_0)^{(3s)} & P_1(\beta_3)^{(3s)} & P_2(\beta_1)^{(3s)} & P_3(\beta_2)^{(3s)} \end{array}$$

Where β_3 is first proposed to swap with β_2 and the swap is accepted, then β_2 is proposed to swap with β_1 and the swap is accepted, and then finally, β_1 is proposed to swap with β_0 and the swap is rejected. Using this method, β values will theoretically move from processor to processor such that a redistricting plan which was initially being modified using constant temperature sampling with a β value of $\beta_0 = 0$ will recurrently be sampled at $\beta_n = 1$ (similar to the method used in simulated annealing).

Perhaps clear to the reader by now, there are many decisions which must be made in order to perform parallel tempering, such as

1) How many distinct β values should be used (i.e. what is the world size) ?
2) How often should the swapping of β values be proposed?
3) How should β values be spaced from one another?

Answers to these questions come more from art than science, as there are no known procedures which will lead to lead to optimal results prior to trial and error [1].

B. Beta Value Distribution

As proposed in the section above, the distribution of beta values must be taken into consideration prior to parallel tempering, of few such distributions are described here.

a) Constant Spacing: One perhaps trivial choice for the spacing of β values is to have $\beta_i = \beta_{i+1} - \frac{1}{n-1}$ for $i < n$ where n is the world size begin used. This spacing is simple to implement and results in values of beta having a relationship to $\frac{1}{n-1}$ via $\beta_i = \frac{i}{n-1}$ for $\beta_0, \ldots, \beta_{n-1}$. Constant spacing for beta values is shown as the black line in the Figure 1 below.

b) Log Transformed Spacing (concave): In certain situation, it may be appealing to have beta values such that $e^{\beta_0}, \ldots, e^{\beta_{n-1}}$ is constantly spaced. In other situations, it may be wise to have beta values which "bunch up" near values near 1 and "spread out" near values around zero. (i.e. the distribution of beta values is concave). Both of these conditions can be satisfied by taking β values which were originally constantly spaced and applying a log transformation to the beta values. This distribution is plotted in the red line within Figure 1.

c) Exponentially Transformed Spacing (convex): Just as having beta values which are concave may be beneficial in certain situation, having beta values which are convex may be beneficial in others. These situations may arise when beta values between $(0, .2)$ are having a difficult time swapping with each other (i.e. the vast majority of proposed swaps is rejected). In this situation, an exponential transformation of the constantly spaced beta values as well. In addition to this transformation, the slope of the exponentially transformed beta distributions can be tuned by choosing a parameter $m \in \mathbb{N}$ such that

$$\beta_i = \frac{e^{im/(n-1)}}{e^m}$$

Two distributions of this nature using values of $m = 3$ and $m = 5$ can be seen in Figure 1 as the blue and gray lines, respectively.

C. Multi-parameter Parallel Tempering

Parallel tempering on β values can be generalized to parallel tempering over a set of hyperparameters. For example,

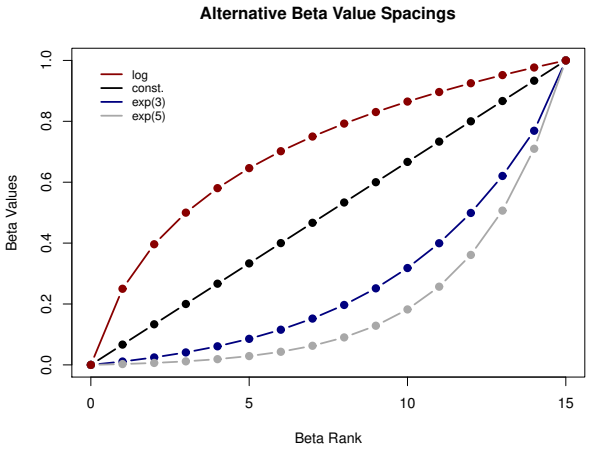

Fig. 1. Various Beta Value Spacings

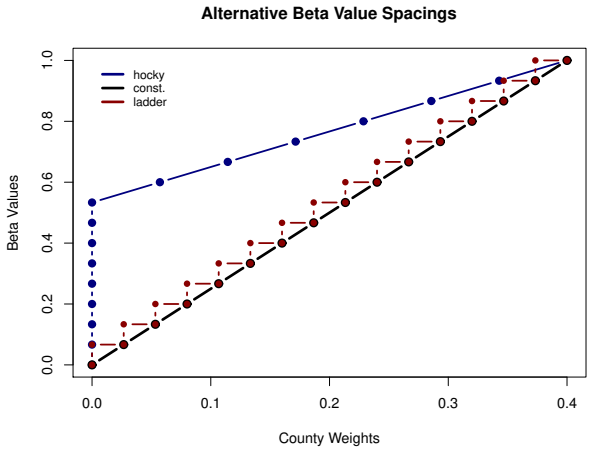

Fig. 2. Parallel Tempering of Beta Values and County Split Weights

since $S(\xi) \propto e^{-\beta J(\xi)}$, the score of a districting plan can be "cooled down" by either increasing β or changing the way $J(\xi)$ is calculated. For example, one may trade on the weights w_{pop}, w_{iso}, or w_{county} as well as the values of β. If this is desired, then the simple correction of

$$r = \left(\frac{S(\xi_{i-1}, w_{county,i-1}, \beta_i)}{S(\xi_i, w_{county,i}, \beta_i)} \frac{S(\xi_i, w_{county,i}, \beta_{i-1})}{S(\xi_{i-1}, w_{county,i-1}, \beta_{i-1})} \right)$$

for the calculation of r is all that is needed for the calculation of the acceptance or rejection of swaps. In this situation, increasing the any of these weights would be equivalent to increasing the value of β which a given processor is using. As a more concrete example, consider a parallel tempering scheme where tuple values of $(\beta_i, w_{county,i})$ are being swapped across processors. In this situation, there is now a choice of how to space β values from one another as well as w_{county} values.

One such spacing would be constant spacing across beta values as well as county splitting weights, where each swap of tuples leads to an increase or decrease in the beta value a sampler is running as as well as a corresponding increase or decrease in the county weight. This may be problematic, as two values are changing at once, which may cause large changes in the scores of any two particular distracting plans.

Alternatively, one may hold the value of county weights constant for some predetermined number of processors and then gradually increase county weight values along with beta values, allowing a perhaps smoother transition; this type of spacing would represent that of a "hockey stick" when β is plotted as a function of w_{county}.

Lastly, one may wish to only change either a β value or a w_{county} values during any given swapping of tuples. This may be accomplished following a ladder type approach of beta value spacings and county weights, as shown in Figure 2 below.

D. Visualizing Parallel Tempering

As discussed previously, "good" parallel tempering should occur when beta values are swapping with each other at a rate which allows beta values which started at a specific processor to have a uniform stationary distribution across all processors over time. In order to evaluate such "mixing" of beta values, one may implement parallel tempering and track the location of beta values as a function of the number of proposed swaps up to that point. As a results, plots such as those shown in Figure 3, which was obtained from sampling on the Union Anson county cluster using a world size of 8 and constant beta spacing with proposed swaps ever 500 steps of the sampler.

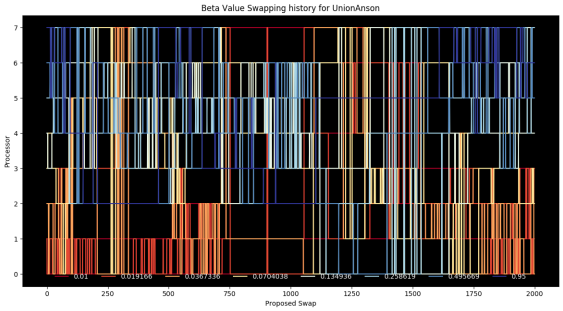

Fig. 3. Parallel Tempering on the Union Anson County Cluster

In the above graph, the color of each line references a separate β value and each integer along the y-axis refers to the processor at which the corresponding beta value is located. As shown, the colors of the beta values quickly mix with each other and change location. These plots are commonly referred to as "mixing plots" given their use.

E. Discussion of Parallel Tempering on North Carolina

The parallel tempering sampling scheme has been implemented to sample redistricting plans for North Carolina Congressional Districts. Much of the work from the authors of this report has entailed the exploration of different beta spacing as well as the effects of county weight. As a result of such experiments, there is preliminary evidence that a convex

spacing of beta values may be preferable when sampling across the entire state of North Carolina. This hypothesis arose after running several sampling runs using constant beta spacing which resulted in mixing plots which revealed heavy swapping around colder temperature betas and from visualizations provided from trace plots of the energy of redistricting plans where separate beta value are located which showed redundant overlap in the energy distributions for lower temperature beta values.

In addition to these efforts, much of the work performed by the authors of this paper has pertained to the exploration of data visualizations which allow for the assessment of parallel tempering such as the calculation of frequencies for individual beta swaps (i.e. the swap frequency of β_2 and β_3) as well as the calculation of specific energy caused by the various different sub-functions of $J(\xi)$ (i.e. what the change in population deviation energy was as a result of a particular beta swap), however the results of these efforts have been inconclusive.

F. Future Goals of Parallel Tempering

Although the computational aspects of parallel tempering on a multiparameter space (i.e. parallel tempering of county weights as well as beta values) has been implemented, extensive testing of such parallel tempering has yet to be performed as well as the development of methods to analyze the results of such methods. Much of the previous work of the authors of this paper has been towards the understanding of the parallel tempering scheme in general and the implementation of this method. It is the goals of these authors to further develop methods of evaluating parallel tempering, with one unexplored areas being the analysis of redistricting plans produced by parallel tempering by developing reliable metrics to measure the distance between redistricting plans.

This being said, having already built of an analytically took kit for the analysis of parallel tempering runs, future efforts will focus on using these methods to analyze future parallel tempering runs in order to develop a better understanding of how parallel tempering may be used explore Φ, the space of possible redistricting plans of North Carolina.

IV. Stratified Sampling

A. Overview of Procedure

The Umbrella Sampling method is a variant of the stratified sampling approach originally proposed by Torrie Valleau (1977) and has found widespread use and success in computational chemistry problems where long posterior tails play an essential role (Boczko Brooks 1995; Berneche Roux 2001). Similar to the map sampling problem we are interested in, this class of application suffers from high dimensionality of sample space, multi-modal $\pi(x)$ and a particular interest in low-probability nodes. As a result both domains experience high computational cost of the $\pi(x)$ evaluation and slow convergence of the MCMC estimate.

The Umbrella Sampling procedure seeks to address these challenges by dividing up or stratifying the sample space into many smaller MCMC sampling problems. Overlapping window functions, or umbrellas $\psi_i(x)$, are defined which confine the MCMC walks to their corresponding distributions, $\pi_i(x) \propto \psi_i(x)\pi(x)$. If a selected window contains a high energy region sampling is necessarily confined to it, which enables much more efficient coverage of low probability areas and ensure the discovery of widely separated peaks in multi-modal landscapes. As such, similar to parallel tempering, it is designed to sample complex space more uniformly.

After sampling these diverse regions of state space independently, they must be combine to approximate the whole of the region sampled (Dinner et al. 2017). To do so, we use an implementation of the Eigenvalue Method for Umbrella Sampling (EMUS) described by [4] and originally proposed by [5] [6]. EMUS is an iterative method for computing the relative weights of each sample umbrella $\pi_i(x)$ and combining them to compute the global $\pi(x)$, where

$$\pi_i(x) := \frac{1}{z_i}\psi_i(x)\pi(x)$$

with $\psi_i(x)$ being the umbrellas or strata for each individual distribution, and z_i is the normalizing constant that is needed to ensure that the $\sum \pi_i(x) = 1$

Algorithm 2 Umbrella Sampling

Target Distribution $\pi(x)$

Define $\pi_i(x)$ by centroid c_i, radius r, and metric space $(\pi(x), \rho)$

for each $\pi_i(x) \in \pi(x)$ **do**:
  Sample $\xi \in \pi_i(x)$
  if $\rho(c_i, \xi) > r$ **then**
    Store ξ in $\pi_i(x)$.exitPoints
  else
    Store ξ in $\hat{\pi}_i(x)$
  end if
end for

Let F be a $i \times i$ matrix

for $\xi_i \in \pi_i(x)$.exitPoints $\forall \pi_i \in \pi(x)$ **do**:
  for each $\pi_j(x) \in \pi(x)$ **do**
    if $\rho(\xi_i, c_j) < r$ **then**
      $F[i][j] += 1$
    end if
  end for
end for

Let $z =$ Eigenvector of F
Each z_i is normalization constant for $\pi_i(x)$

$$\hat{\pi}(x) = \sum z_i \cdot \hat{\pi}_i(x)$$

return $\hat{\pi}(x)$

B. Implementation

Umbrella Sampling does not make any specifications regarding how distance between points in space should be defined, how the window function ψ_i should confine each umbrella π_i, or how to select each new umbrella's centroid C_i. Here we outline the domain specific Metric Space $(\pi(x), \rho)$, window function $\psi_i(C_i, \xi)$, and the selection procedure for C_i.

We define a simple the euclidean distance for ρ in $\mathbb{R}^{13}$ (each dimension representing a district's latitude and longitude) and compute ρ between a windows centroid map C_i and some proposed map's ξ embedding in $\mathbb{R}^{13}$ to define our window function ψ_i.

$$\rho(C_i, \xi) := \sum_{j=0}^{13} (\xi_{xj} - C_{xj})^2 + (\xi_{yj} - C_{yj})^2$$

$$\psi_i(C_i, \xi) := \begin{cases} \text{accept} & \rho(C_i, \xi) < r \\ \text{reject} & \rho(C_i, \xi) \geq r \end{cases}$$

Each dimension in a plan's centroid embedding in $\mathbb{R}^{13}$. is computed by getting the geographic latitude x_j and longitude y_j of each district's composite precincts.

This function $\psi_i(\xi)$ rejects proposed redistricting plans if their euclidean distance ρ from the umbrellas centroid fall beyond some arbitrary radius r. In setting this radius the goal should be to capture a sufficient number of redistricting plans, such that other compliant maps are captured but also not so large that the MCMC does not frequently attempt to step beyond the radius.

In this implementation an umbrella's centroid C_i is defined by a redistricting plan that was previously attempted, but did not satisfy ψ_{i-1}. This redistricting plan, however, satisfies all other step acceptance criteria (e.g. contiguity). The most recently rejected redistricting plan after the previous umbrella terminates therefore becomes the centroid C_i for the next sampling umbrella and a new sampling procedure confined by ψ_i begins. When instantiating the first umbrella, we begin with a compliant seed centroid which takes the form of either the North Carolina Judge's Map, the implemented North Carolina 2012 Map, or the implemented North Carolina 2016 Map.

C. Distance and Radius Selection

In Figure 3, we show an exploration of distance space surrounding several seed centroid selections. First and foremost this data suggests that the state space is non-uniform with regard to ρ and varies based on what region of state space is being explored. This underscores the influence of radius selection on how effectively an umbrella can sample. It suggests that radius selection should likely be dynamically computed instead of fixed for all umbrellas to account for variations in state space. If the State Space surrounding a centroid is extremely sparse with redistircting plans in $\mathbb{R}^{13}$, the radius should be adjusted to capture more maps. If the State Space surrounding a centroid is instead dense with redistircting plans in $\mathbb{R}^{13}$, the radius should be contracted

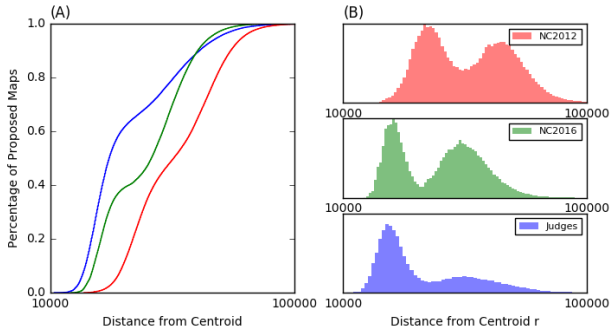

Fig. 4. (A) Cumulative Density Functions (cdfs) that display the percentage of proposed maps that fall within radius *r* during the first 800,000 steps. (B) Each histogram displays the distribution of these steps as they step away from the seed centroid. Each color represents a unique seed centroid, (red) Uses the implemented North Carolina 2012 map as the seed centroid from which 800,000 consecutive steps are proposed. (green) Uses the implemented North Carolina 2016 map and (blue) uses the bipartisan Judges map. All steps, both those rejected and accepted are included in the CDFs and histograms.

to allow for more thorough exploration. The nature of the histograms in Figure 3B also suggest that the minimum step distance for ρ exists on the order of 10,000 and highlights a tendency to take many steps very close to low-energy compliant maps (the centroids in all displayed simulations are compliant). The bi-modal nature of each suggests that a second low-energy state was found just beyond the compliant centroid in each simulation. Further, if the simulation was allowed to continue to run, additional peaks more distant from the centroid would likely appear in the histograms. These additional peaks would theoretically represent additional low-energy states. The question therefore becomes how many of these low-energy states would we like to capture within a specified umbrella radius. This consideration must then be balanced with a desirable non-zero percentage of compliant redistrictings being rejected by ψ . Discerning the balance between these two considerations should be the task of future inquiry.

D. Centroid Selection

Although umbrella sampling is highly parallelizable, it is not possible to sample all $\pi_i(x)$ at once. Thus, we must decide which $\pi_i(x)$ to prioritize and sample first. Since this decision must be made prior to actually sampling a $\pi_i(x)$, we must make use of other information. A possible solution is to pick a map, ξ, as a centroid if it is surrounded by other low energy or compliant maps. This however could have negative consequences, as one of the advantages of umbrella sampling is that it forces exploration of higher energy areas. Other naive approaches include allocating centroids in a first-come-first-serve approach (our current implementation), selecting centroids with a probability proportional to its energy, or selecting purely at random.

V. Parallel-Umbrella Sampling

An area of further research is to combine parallel tempering and umbrella sampling into a single algorithm. The un-

derlying structure is extremely similar to umbrella sampling with its multiple strata. In umbrella sampling, each strata is given the same β value. In parallel-umbrella sampling, each stratum $\pi_i(x)$ would have a β_i. In addition to proposing moves to a new $\xi \in \pi_i(x)$, moves could also be proposed to swap b_i and b_j in the same manner as parallel tempering. This gives the same benefits as parallel tempering, namely that it allows for exploration of multiple local minimums while allowing jumps over large energy barriers, but the benefit is now also shared in each stratum. This procedure would have benefits over the ordinary umbrella sampling if within a stratum there is high variance in energy levels.

VI. Conclusion

In this report, we document the theoretical and material factors which influence the implementation of various methods used to sample congressional redistricting plans of North Carolina, in particular, we focus on parallel tempering and umbrella sampling. Our investigations have a promising future as they have developed a variety of analytical techniques to evaluate the efficiency of various sampling procedures. Specifically, the development of multi-parameter parallel tempering over beta values as well as county splitting weights may allow for better mixing than parallel tempering over beta values alone may afford. In addition to this accomplishment, we successfully implemented a basic umbrella sampling procedure for the first time in this domain. While it effectively creates, explores, and merges umbrella samples, much future work is required to improve radius and centroid selection. Preliminary data presented here also provides insight into the nature of how the state space is distributed with regard to our selected distance function ρ which has implications for the definition of ψ. Much additional exploration into alternative notions of distance (e.g. $\rho = |J(\xi_i) - J(\xi_j)|$ or $\rho = |J_{pop}(\xi_i) - J_{pop}(\xi_j)|$, etc.) is warranted as the state space is highly complex and the definition of ρ fundamentally changes how the space will be sampled. Further work must also focus on parallelizing the umbrella sampling procedure so that it can be run at scale.

ACKNOWLEDGMENT

The Duke Math Department, the Information Initiative at Duke (iID), Data+ undergraduate research programs, and Duke Bass Connections for financial and material support. The Quantifying Gerrymandering Team has significantly influenced this work. Rahul Ramesh for parallel tempering implementation. Greg Herschlag and John Mattingly for immense theoretical, computational, and emotional guidance.

Algorithm 3 Iteratively Finding Beta Values

β_h = 0.1, β_c = 0.1 * spacingFactor
$\beta tuples \sim \{[\beta_c, \beta_h], [\beta_c, \beta_h], \dots\}$

After each β tuple has proposed x number of swaps:
frequencies $\sim \{frequency_1, frequency_2, \dots, frequency_n\}$

if $30\% \leq allfrequencies \leq 50\%$ **then**
  move on
else
  for frequency in frequencies **do**
    if $frequency < 30\%$ **then**
      β_c-=$(\beta_h - \beta_c) \cdot 0.09$
    end if
    if $frequency > 50\%$ **then**
      β_c+=$(\beta_h - \beta_c) \cdot 0.2$
    end if
  end for
end if

if move on **then**
  update all $\beta_c$to lowest β_h
  update all $\beta_h$to $\beta_c \cdot$ spacingFactor
end if

Algorithm 4 Diversity Evaluation

```
diversityScore, totalDistance
for mapA:maps do
    list centroidsA
    for mapB:maps excluding mapA do
        list centroidsB
        for centroid:centroidsA do
            d = distance to closest centroid in centroidsB
            totalDistance += d
        end for
    end for
    diversityScore += totalDistance
end for
```

References

[1] B. Fifield, M. Higgins, K. Imai, and A. Tarr (2018) A New Automated Redistricting Simulator using Markov Chain Monte Carlo.
[2] G. Herschlag et. al (2018), Quantifying Gerrymandering in North Carolina. arXiv:1801.03783
[3] Warrington GS (2017) Quantifying gerrymandering using the vote distribution. arXiv:1705.09393.
[4] Charles Matthews et. al (2016), Umbrella sampling: a powerful method to sample tails of distributions distribution arXiv: 1712.05024.
[5] Thiede E. H., et al. (1977) Journal of Chemical Physics, 145, 084115 Torrie G., Valleau J., 1977, Journal of Computational Physics, 23,187
[6] Dinner A. R., Thiede E., Van Koten B., Weare J. (2017), distribution: ArXiv:1705.08445
[7] Chambers, Christopher P. and Miller, Alan D. and Sobel, Joel, Flaws in the Efficiency Gap (October 2, 2017). Journal of Law and Politics, Vol. 33, No. 1, 1. Available at SSRN: https://ssrn.com/abstract=3046797