Nonparametric Detection of Gerrymandering in Multiparty Plurality Elections

Dariusz Stolicki[1] , **Wojciech Słomczyński**[1] and **Stanisław Szufa**[2,3]
[1]Jagiellonian University, Center for Quantitative Political Science, Kraków
[2]AGH University, Kraków
[3]CNRS, LAMSADE, Université Paris Dauphine - PSL
{wojciech.slomczynski, dariusz.stolicki}@uj.edu.pl, s.szufa@gmail.com

Abstract

Partisan gerrymandering, i.e., manipulation of electoral district boundaries for political advantage, is one of the major challenges to election integrity in modern day democracies. Yet most of the existing methods for detecting partisan gerrymandering are narrowly tailored toward fully contested two-party elections, and fail if there are more parties or if the number of candidates per district varies. We propose a new method, applying nonparametric statistical learning to detect anomalies in the relation between (aggregate) votes and (aggregate) seats. Unlike in most of the existing methods, we propose to learn the standard of fairness in districting from empirical data rather than assume one a priori. Finally, we test the proposed methods against experimental data as well as real-life data from 17 countries employing the plurality (FPTP) system.

1 Introduction

Most of the traditional methods developed for detecting gerrymandering in first-past-the-post electoral systems assume that there are only political parties really contesting the election, or, at least, that the party system is regular in the sense that all parties field candidates in every district. This is certainly a very reasonable assumption in many cases: under a well-known empirical regularity known as *Durverger's law* FPTP tends to be correlated with the emergence of two-party systems. Moreover, many of the authors working on gerrymandering detection are motivated by U.S. legislative elections (state and federal), where the regular two-party pattern of competition prevails. However, in many other systems using FPTP we discover significant deviations from such patterns in the form of regional parties, strong independent candidates, minor parties that forgo campaigning in some districts, etc. In the face of such deviations, many of the traditional methods fail completely. Our objective, therefore, is to develop a method of detecting gerrymandering that can be applied to such partially-contested multiparty election.

1.1 Contribution

Our main contribution consists of the development of a *nonparametric method for detecting gerrymandering* in *partially-contested multiparty elections*. By *nonparametric* we mean that, unlike most of the traditional statistical methods, the proposed method is free of assumptions about the probability distribution from which observed data points are drawn or the latent mechanism through which such data is generated. Instead we use statistical learning to identify regularities on the basis of the available empirical data.

By a *partially-contested multiparty election* we mean any FPTP election where at least some candidates are affiliated into one or more *political parties* (after all, if every district-level election is completely independent and candidates cannot be affiliated into blocks, the very concept of gerrymandering as traditionally defined is meaningless), but *for every party there is at most one affiliated candidate in every district* (so there is no intra-party competition). We treat independent candidates as singleton parties.

In particular, we permit the following deviations from the two-party competition pattern: (1) the number of parties can differ from two, (2) the number of candidates within each district can differ from two, (3) a party can run candidates in any number of electoral districts, (4) the set of parties contesting the election varies from one district to another.

Another area in which our approach differs from traditional methods for detecting gerrymandering is that they have been tailored towards *testing a large ensemble of elections* (not necessarily from the same jurisdiction) rather than a single election. In particular, the proposed methods, like all statistical learning methods, require the researcher who wants to use them to have a large *training set of elections* that they believe to be sufficiently similar insofar as the translation of votes into seats is concerned. If there is a large ensemble of elections being tested, they might form such a training set itself. There is no requirement that the training set and the tested set be disjoint as long as we can assume that gerrymandering is not ubiquitous and we are only concerned with assessing individual test cases, since for each test case we can train the model on the remainder of the ensemble.

1.2 Prior Work

Among the methods of detecting gerrymandering that focus on the political characteristics of the districting plan (e.g., its impact on seats-votes translation or district-level vote distribution) the earliest focused on measuring how actual elections results deviate from a theoretically or empiri-

cally determined *seats-votes curve* [Kendall and Stuart, 1950; March, 1957; Taagepera, 1973; Grofman, 1983; Browning and King, 1987; Gilligan and Matsusaka, 1999]). There is a consensus in the literature that a two-party seats-votes relation is usually described by a modified power law:

$$\frac{s_i}{1-s_i} = \beta_i \left(\frac{v_i}{1-v_i} \right)^{\rho}, \tag{1.1}$$

where s_i and v_i are, respectively, seat- and vote-share of the i-th party, β_i is party-dependent, and ρ is a constant [Tufte, 1973; Grofman, 1983]. However, only few authors have considered multiparty elections [Taagepera, 1986; King, 1990; Linzer, 2012], and their results are mostly heuristic.

The state-of-the-art approach to detecting gerrymandering is the *partisan symmetry* method. The general concept was first proposed by Niemi and Deegan [1978], who noted that an election should not be regarded as gerrymandered if it deviates from a model seats-votes curve as long as the deviation is the same for each party, i.e., each party has the same seats-votes curve. The main challenge here lies in obtaining that curve from a single realization. The original idea has been to extrapolate by assuming a *uniform partisan swing*, i.e., that as the aggregate vote share of a party changes, its district-level vote shares increase or decrease uniformly and independently of their original levels [Butler, 1947]. A more sophisticated extrapolation method has been developed by Gelman and King [1990a; 1990b; 1994]. Yet neither of these two methods can account for multiple parties absent very restrictive (and unrealistic) assumptions.

The third approach is the *efficiency gap* method proposed by McGhee [2014] and further developed in Stephanopoulos and McGhee [2015]. It is based on the assumption that in an unbiased election all contending parties should waste the same number of votes. While *prima facie* attractive, this assumption is actually highly problematic because it requires the electoral system to match a very specific seats-votes curve [McGann *et al.*, 2015, p. 296], see also [Bernstein and Duchin, 2017]. In this respect it represents a methodological step backwards, making it again impossible to distinguish asymmetry from responsiveness. More importantly for us, the method fails to account for multiple parties.

Finally, there are several method designed to identify anomalies in the vote distribution indicative of standard gerrymandering techniques like packing and cracking. These include the *mean-median difference* test proposed by McDonald *et al.* [2011], which measures the skewness of the vote distribution; the *multimodality* test put forward by Erikson [1972]; the *declination coefficient* introduced by Warrington [2018] and measuring the change in the shape of the cumulative distribution function of vote shares at $1/2$; and the *lopsided winds* method of testing whether the difference between the winners' vote shares in districts won by the first and the second party is statistically significant [Wang, 2016]. Again, virtually of all those methods assume a two-party system. For instance, natural marginal vote share distributions in multiparty systems (e.g., beta or log-normal distribution) are necessarily skewed. Similar assumptions underlie the declination ratio and the lopsided wins test. Finally, the multimodality test assumes a constant number of competitors.

2 Preliminaries

Gerrymandering is usually defined as manipulation of electoral district boundaries aimed at achieving a political benefit. Hence, *intentionality* is inherent in the very concept. However, identical results can also arise non-intentionally, as geographic concentration of one party's electorate in small areas (major cities, regions) can produce similar effects to intentional packing. We use the term 'electoral bias' to refer to such 'natural gerrymandering'.

Our basic idea is to treat gerrymandering and electoral bias as *statistical anomalies* in the *translation of votes into seats*. Identification of such anomalies requires a reference point, either theoretical, such as a theoretical model of district-level vote distribution, or empirical, such as a large set of other elections that can be expected to have come from the same statistical population. As the former approach is burdened with the risk that the theoretical model deviates from the empirical reality, in this paper we focus on the latter.

One major limitation of our method lies in its *inability to distinguish gerrymandering from natural electoral bias*. This limitation is shared, however, with virtually all methods in which the evidence for gerrymandering is sought in analyzing voting patterns. For many applications that may be enough, since for the end users it might not matter whether the bias in the electoral system is artificial or natural. If the distinction does matter, the proposed method can still be useful to identify cases requiring more in-depth investigation.

There are three basic assumptions underlying our method. One is that we have a *training set* of elections that come from the same statistical population as the election we are studying. Another one is that electoral bias is *an exception rather than a rule*. Thus, we assume that a substantial majority of the training set elections are free from bias. The third assumption is that while district-level results can be biased, *aggregate electoral results* (e.g., vote shares) never are.

2.1 Notation

Let us introduce some basic notation to be used in this paper:

sequence For $k \in \mathbb{N}$, let $[k] = \{1, \ldots, k\}$.

unit simplex For $m \in \mathbb{N}_+$, we denote the unit simplex by $\{\mathbf{x} \in \mathbb{R}^m_+ : \|\mathbf{x}\|_1 = 1\}$ by Δ_m.

k-th largest / smallest coordinate For $m \in \mathbb{N}_+$, $\mathbf{x} \in \mathbb{R}^m$, and $k \in [m]$, we denote the (k)-th largest coordinate of $\mathbf{x}$ by $x^{\downarrow}_k$, and the (k)-th smallest one by $x^{\uparrow}_k$.

set of districts We denote the set of districts by $D := [c]$.

set of parties We denote the set of parties by $P := [m]$.

set of contested districts For $i \in P$, we denote the set of districts in which the i-th party runs a candidate by D_i.

set of contesting parties For $k \in D$, we denote the set of parties that run a candidate in the k-th district by P_k.

district size For $k \in D$, we denote the number of voters cast in the k-th district by w_k.

district-level vote share For $i \in P$ and $k \in D$, we denote the district-level vote share of the i-th party's candidate in the k-th district, i.e., their number of votes divided by w_k, by v^k_i. If there is no candidate, we assume $v^k_i = 0$.

district-level seat share For $i \in P$ and $k \in D$, let s_i^k equal 1 for the winning candidate, and 0 for others.

aggregate vote share For $i \in P$, we denote the aggregate vote share of the i-th party, i.e., the sum of all votes cast for that party's candidates, $\sum_{k\in D} v_i^k w_k$, divided by the number of all votes cast in the districts contested by that party, $\sum_{k\in D_i} w_k$, by v_i. Note that v_i differs from mean vote share over districts by being weighted by w_k.

aggregate seat share Let $s_i := \left(\sum_{k\in D_i} s_i^k\right)/|D_i|$ be the aggregate seat share of the i-th party, $i \in P$.

2.2 Seats-Votes Functions

As noted above, we define electoral bias (and thus gerrymandering) in terms of the relations between votes to seats. Mathematically, such relation is represented by a seats-votes function, i.e., a measure-valued function that maps an aggregate vote share to a probability measure on the domain of possible aggregate seat shares:

Definition 2.1 (Seats-Votes Function)**.** *A* seats-votes function[1] *is an increasing function $f : [0,1] \longrightarrow \mathcal{P}([0,1])$, where $\mathcal{P}(X)$ denotes the set of Borel probability measures on some measurable set X, and monotonicity is defined as by Burkill [1984] in terms of the partial order on $\mathcal{P}([0,1])$ induced by the stochastic ordering of random variables.*

Definition 2.2 (Seats-Votes Curve)**.** *For a fixed seats-votes function, a* seats-votes curve *is a function $s : [0,1] \longrightarrow [0,1]$ that maps an aggregate vote share to the expectation of its image under that seats-votes function.*

Why do we define seats-votes relations in probabilistic rather than deterministic terms? The answer becomes clear once we note that in plurality voting systems with districts the number of seats obtained by a party depends directly not on its aggregate vote share, but on district vote shares of itself and of its competitors (or, to be more precise, on the first order statistic of the latter, i.e., the district vote share of the largest competitor in a given district). Hence, except for the trivial case of a single district and two parties, aggregate vote share carries insufficient information to uniquely determine the resulting seat share. Thus, speaking of a seats-votes function only makes sense if we treat the two undetermined factors – *allocation of the party vote across districts* and *division of the remaining votes among competitors* – as random.

Remark 2.1. Consider seats-votes curves in multi-party elections. If we assume that they are anonymous (i.e., identical for all parties), non-decreasing, and surjective, perfect proportionality ($s_i = v_i$ for each $i \in P$) is the only one that does not depend on the distribution of competitors' votes [Boratyn *et al.*, 2022, Theorem 1].

3 Overview of the Method

Our basic idea is to detect electoral bias by comparing observed seat shares with the estimated value of the seats-votes function for a given party at its observed vote share. Seat share values corresponding to the tails of the seat distribution are then considered indicative of bias. However, gerrymandering arises from the manipulation of district boundaries, and is thus exclusively associated with the first source of seat share randomness. Thus, we actually need to compare the observed seat-share with the value of the seats-votes function conditional on the *competition pattern*, i.e., the number of competitors and the distribution of their votes [Calvo, 2009; Manow, 2011; Calvo and Rodden, 2015].

3.1 Seat Thresholds

It would be convenient if we were able to describe the competition pattern by a single numerical parameter. Our objective here is to find a measure of the 'difficulty' of winning a seat given the number of competitors and the distribution of their vote shares (renormalized so as to sum to 1). A natural choice would be the *seat threshold*:

Definition 3.1 (Seat Threshold)**.** *Fix $i \in [m]$, and assume that renormalized vote shares of the competitors of the i-th candidate equal some random variable $\mathbf{Z}$ distributed according to some probability measure on $\Delta_{m,-i}$. A* seat threshold *of the i-th candidate is such $t_i \in [1/m, 1/2]$ that $Pr(S_i = 1|v_i) > 1/2$ for every $v_i > t_i$, i.e., the probability that the i-th candidate wins a seat with vote share equal v_i exceeds $1/2$.*

Observation 3.1. *It is easy to see that the probability of winning a seat, $Pr(S_i = 1|v)$, equals $1 - F_{Z_1^{\downarrow}}(v/(1-v))$, where $F_{Z_1^{\downarrow}}$ is the cumulative distribution function of the renormalized vote share of the largest competitor.*

For practical applications, we need to approximate the seat threshold in cases where we do not have any knowledge of the distribution of the competitors' vote shares, but only a single realization thereof. We therefore need a statistic that is both a stable estimator of the distribution parameters and highly correlated with the value of the largest order statistic. We posit that the best candidates for such statistics are *measures of vote diversity* among competitors, and use a Monte Carlo simulation to test a number of such measures.

Observation 3.2. *Let $\alpha \sim \mathrm{Gamma}(1,1)$, and let $\mathbf{V} \sim \mathrm{Dir}(\{\alpha\}^m)$, where $m = 3, \ldots, 12$ and $\mathrm{Dir}(\boldsymbol{\alpha})$ is the Dirichlet distribution with parameter vector $\boldsymbol{\alpha}$. For a sample of 2^{16} realizations of $\mathbf{V}$ we have calculated Spearman's correlation coefficients [Spearman, 1904] for:*

1. *α,*
2. *$\mathbf{V}_1^{\downarrow}$, i.e., maximum of the coordinates,*
3. *$\mathbf{V}_1^{\uparrow}$, i.e., minimum of the coordinates,*
4. *median coordinate V_{med},*
5. Shannon entropy *[Shannon, 1948],*

$$\mathrm{H}(\mathbf{V}) := -\sum_{i=1}^{m} V_i \log V_i,$$

6. Herfindahl–Hirschman–Simpson index *[Hirschman, 1945; Simpson, 1949; Herfindahl, 1950],*

$$\Phi(\mathbf{V}) := \sum_{i=1}^{m} V_i^2,$$

[1] It would be more natural to speak of *votes-seats functions*, since vote shares are clearly arguments and seat shares are clearly outputs, but the term *seats-votes function* is already widely used in electoral studies literature.

7. Gini coefficient *of the coordinates,*
8. Bhattacharyya angle *[Bhattacharyya, 1943] between V and the barycenter of the simplex,*

$$B(\mathbf{V}) := \arccos \sum_{i=1}^{m} \sqrt{V_i/m}.$$

The results for $m = 3$ are given in Table 1, while those for $m = 6$ and 12 – in the full version of the paper [Słomczyński et al.*, 2024].*

	α	$V_1^{\downarrow}$	$V_1^{\uparrow}$	V_{med}	$\text{H}(\mathbf{V})$	$\Phi(\mathbf{V})$	**Gini**	$B(\mathbf{V})$
α	1.0	-.51	.58	.25	.58	-.56	-.57	-.59
$V_1^{\downarrow}$	-.51	1.0	-.81	-.73	-.94	.97	.97	.92
$V_1^{\uparrow}$	.58	-.81	1.0	.23	.95	-.92	-.93	-.97
V_{med}	.25	-.73	.23	1.0	.49	-.56	-.55	-.44
$\text{H}(\mathbf{V})$	.58	-.94	.95	.49	1.0	-.99	-.99	-.99
$\Phi(\mathbf{V})$	-.56	.97	-.92	-.56	-.99	1.0	.99	.99
Gini	-.57	.97	-.93	-.55	-.99	.99	1.0	.99
$B(\mathbf{V})$	-.59	.92	-.97	-.44	-.99	.99	.99	1.0

Table 1: Correlation Matrix for $m = 3$.

The *Herfindahl–Hirschman–Simpson index* is consistently the one that best correlates with the maximal coordinate while also being a reasonably good estimate of the distribution parameters. Accordingly, in our procedure for estimating the seat threshold we use its monotonic transform, known as the *effective number of competitors* [Laakso and Taagepera, 1979; Taagepera and Grofman, 1981]:

Definition 3.2 (Effective Number of Competitors)**.** *For $i \in [m]$ and $k \in D_i$ the* effective number of competitors *of the i-th candidate in the k-th district equals:*

$$\varphi_i^k := \left(\sum_{j=1, j\neq i}^{m} z_j^2 \right)^{-1}, \tag{3.1}$$

where $\mathbf{z} \in \Delta_{m,-i}$ is a vector of the vote shares of that candidate's competitors normalized so that $\sum_{j=1,j\neq i}^{m} z_j = 1$.

We shall see that the vote share and the number and effective number of competitors enable us to accurately classify candidates as winning and losing (see Figure 1 and Table 2).

Observation 3.3. *Clearly, with three candidates, i.e., two competitors, the classifier is exact (modulo ties), as the effective number of competitors uniquely determines the share of the larger one in their aggregate vote share:*

$$\max\{z_{j_1}, z_{j_2}\} = \frac{1}{2}\left(1 + \sqrt{\frac{2}{\varphi_i} - 1}\right). \tag{3.2}$$

Then the decision boundary is the set of points satisfying:

$$\varphi = \frac{1 - 2v + v^2}{1 - 4v_i + 5v^2}. \tag{3.3}$$

For the case of more than three candidates, there is no exact classifier, whence we use an approximate one:

Model 3.1 (Decision Boundary for $m > 3$)**.** *For $m > 3$, the decision boundary is determined on the basis of the data using a support vector machine-based classifier [Boser* et al.*, 1992; Cortes and Vapnik, 1995] with a third-order polynomial kernel, and then approximated by a strictly decreasing B-spline of degree 3, with boundary nodes at $1/m$ and $1/2$ and interior nodes fitted using cross-validation.*

A definition of the effective seat threshold follows from the one of the decision boundary:

Definition 3.3 (Effective Seat Threshold)**.** *The* effective seat threshold *of the i-th party in the k-th district, t_i^k, is the value of the v coordinate of the decision boundary at the point where $m = |P_k|$ and $\phi = \phi_i^k$.*

Definition 3.4 (Effective Seat Threshold Classifier)**.** *An effective seat threshold classifier is a function $\mathfrak{s} : [0, 1] \times \mathbb{N} \times [1, \infty) \to \mathbb{B}$ that maps a triple (v, m, φ) to 0 iff the probability of winning with vote share v, $m - 1$ competitors, and φ effective competitors is below $1/2$.*

m	R	m	R	m	R
3	.0035	7	.0136	11	.0142
4	.0137	8	.0068	12	.0186
5	.0152	9	.0073	13	.0152
6	.0137	10	.0067	14	.0171

Table 2: Effective seat threshold classifier error, R.

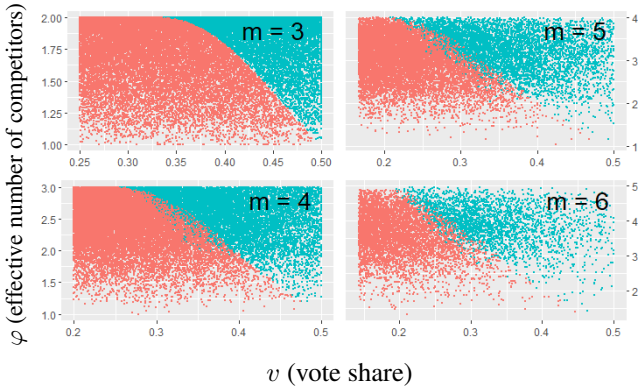

Figure 1: Effective seat thresholds for $m = 3, 4, 5, 6$. Blue points indicate successful candidates, while red points – unsuccessful candidates.

By averaging effective seat thresholds over districts, we finally arrive at our measure of the difficulty of winning a seat:

Definition 3.5 (Mean Effective Seat Threshold)**.** Mean effective seat threshold*, t_i, is given by $t_i := \langle t_i^k \rangle_{k \in D_i}$.*

3.2 Estimating Seats-Votes Functions

A natural approach to estimating seats-votes functions would be to start with some probabilistic model of inter-party vote distribution, preferably derived axiomatically from first principles, then use it to calculate the seat threshold, and use a

probabilistic model of inter-district vote distribution to calculate the probability of district vote share exceeding the seat threshold. Finally, either by convolving binomial distributions (for small values of c) or by the central limit theorem (for large values of c) we would obtain the seat distribution.

Unfortunately, any approach based on a theoretical seats-votes function has an unavoidable weakness: a systematic deviation from the expectation of such a function might just as easily arise from electoral bias as from incongruities between the theoretical distributional assumptions and the empirical reality. To avoid this issue we derive our model seats-votes function solely from the reference election dataset with minimal theoretical assumptions[2] by using the kernel regression method [Nadaraya, 1965; Watson, 1964]. Its general idea is to estimate the conditional expectation of a random variable at a point in the condition space by averaging its realizations at neighboring points with distance-decreasing weights.

Model 3.2 (Locally-Constant Kernel Regression)**.** *Let $S \in \mathbb{R}$ be a random* response variable*, and let $\mathbf{X} \in \mathfrak{F}$, where $\mathfrak{F}$ is some linear* feature space *and $D := \dim \mathfrak{F}$, be a vector of* predictor variables*. Assume we have a vector of N realizations of S, $\mathbf{s}$, and an $N \times \dim \mathfrak{F}$ matrix of realizations of $\mathbf{X}$, $\mathbf{x}$. We denote its j-th row by $\mathbf{x}_j$. Then the* locally-linear kernel regression estimate *of the conditional expectation of S given a vector of predictors $\mathbf{x}_0 \in \mathfrak{F}$ is given by:*

$$\mathbb{E}(S|\mathbf{x}_0) = \frac{\sum_{j=1}^{N} s_j K\left((\mathbf{x}_j - \mathbf{x}_0)\mathbf{h}_{\mathbf{x}_0,\mathbf{x}_j}\right)}{\sum_{j=1}^{N} K\left((\mathbf{x}_j - \mathbf{x}_0)\mathbf{h}_{\mathbf{x}_0,\mathbf{x}_j}\right)}, \tag{3.4}$$

where N is the number of observations (in our case, sum of the number of parties over all elections in our set of elections), K is a second-order kernel*, and $\mathbf{h}_{\mathbf{x}_0,\mathbf{x}_j} \in \mathbb{R}_+^D$ is a* bandwidth *parameter for the pair $(\mathbf{x}_0, \mathbf{x}_j)$. In other words, we average the values of s over all parties with weights determined by the value of the kernel at $(\mathbf{x}_j - \mathbf{x}_0)\mathbf{h}_{\mathbf{x}_0,\mathbf{x}_j}$.*

The above general model can be used with different kernel and bandwidth definitions. We use the following ones:

Definition 3.6 (Kernel)**.** *Let K be the D-*variate Gaussian kernel*:*

$$K(\mathbf{x}) := (2\pi)^{-k/2} \exp\left(-\frac{1}{2}\|\mathbf{x}\|_2\right) \tag{3.5}$$

Definition 3.7 (Multivariate Adaptive Nearest-Neighbor Bandwidth [Breiman *et al.*, 1977; Abramson, 1982; Silverman, 1986; Schucany, 1995])**.** *For $\mathbf{x}_0, \mathbf{x}_j \in \mathfrak{F}$, the i-th coordinate of the* multivariate adaptive nearest-neighbor bandwidth*, $i = 1, \ldots, D$, is given by:*

$$(h_{\mathbf{x}_0,\mathbf{x}_j})_i = h_{0,i}|x_{0,i} - x_{N_k^i(\mathbf{x}_j),i}|, \tag{3.6}$$

where $N_k^i(\mathbf{x})$ is the index of the k-th nearest neighbor of $\mathbf{x}$ along the i-th dimension of the feature space under the absolute difference metric, and $\mathbf{h}_0 \in \mathbb{R}_+^D$ is a scaling vector.

The choice of a Gaussian kernel is a standard one, and the use of adaptive bandwidth is motivated by nonuniformities in data density in both feature and response domains.

[2] In particular, we assume the seat shares to be distributed according to some absolutely continuous probability measure on $[0, 1]$.

We still need to choose two hyperparameters of the model: The scaling vector $\mathbf{h}_0$ and the nearest-neighbor parameter k. This we do using leave-one-out cross-validation [Li and Racine, 2004; Härdle *et al.*, 1988] with the objective function defined as the Kullback-Leibler [Kullback and Leibler, 1951] divergence between the predicted and actual value vectors, together with an optimization algorithm by Hurvich *et al.* [1998] which penalizes high-variance bandwidths (with variance measured as the trace of the parameter matrix) similarly to the Akaike information criterion [Akaike, 1974].

3.3 Deviation from the Seats-Votes Function

By this point, we have estimated a party's expected seat share given its aggregate vote share and the competition patterns in the districts it contests. But what we actually need is a measure of how much the actual seat share deviates from that expectation. A natural choice would be the difference of the two. It is, however, inappropriate, as seat shares only assume values within a bounded interval $[0, 1]$ and there is no reason to expect seat share distributions to be even approximately symmetric around the mean.

We therefore use another measure of deviation: the probability that a seat share deviating from the median more than the empirical seat share could have occurred randomly. Note how this quantity is analogous to the p-value used in statistical hypothesis testing.

Definition 3.8. *Electoral Bias p-Value. Let s_i be an empirical seat share and let μ be the conditional distribution of the aggregate seat share given the empirical aggregate vote share and the empirical mean effective seat threshold. Then the* electoral bias p-value *is given by:*

$$\pi_i := \min\{\mu((0, s_i)), \mu((s_i, 1))\} = \min\{F(s_i), 1 - F(s_i)\},$$

where F is the cumulative distribution function of μ.

We thus need not a regression estimator, but a conditional cumulative distribution function estimator. One approach would be to estimate the conditional density of S_i [Rosenblatt, 1956; Parzen, 1962] and integrate it numerically. This method, however, is prone to potential numerical errors. We therefore use another approach, relying on the fact that a conditional cumulative distribution function is defined in terms of the conditional expectation, and therefore its estimation can be treated as a special case of the kernel regression problem.

There remains one final problem: When comparing parties contesting different number of districts, we need an adjustment for the fact that the probability of getting an extreme value depends on that number (decreasing exponentially as the number of contested districts increases). To avoid that problem, the kernel model for parties with exactly k districts, $k \in [c]$, is trained only on parties with as many or fewer contested districts. If the distribution of the number of contested districts has a tail, it is optimal to adopt a cutoff point k_0 such that for the set of parties contesting k_0 or more districts each party is compared with a model trained on that set.

3.4 Aggregation

The final step is the aggregation of party-level indices into a single election-level index of electoral bias. We would like

our aggregation function to: (1) assign greater weight to major parties than to minor parties; (2) be sensitive to very low p-values and less sensitive to even substantial differences in large p-values; and (3) be comparable among elections, i.e., independent of the number of parties and districts. We use the *weighted geometric mean*:

$$\pi := \exp\left(\sum_{i=1}^{m} w_i \log \pi_i\right), \tag{3.7}$$

where w_i is the number of votes cast for the i-th party divided by the number of all valid votes cast in the election (including those cast in districts not contested by the i-th party).

4 Experimental Test

Testing our proposed method is nontrivial, as we do not have a reliable empirical benchmark: while for two-party elections, such as U.S. congressional elections, there is a wealth of alternative gerrymandering indices, this is emphatically not the case for multiparty elections. Nor there are other data sources that would enable us to label some multiparty elections as gerrymandered apart from anecdotal evidence in the literature which is in most cases either methodologically unreliable, being based on highly circumstantial and indirect evidence, or limited to the most extreme of cases.

We address this problem by testing our method on a large sample of simulated elections, consisting both of 'fair' districting plans, drawn at random with a distribution intended to approximate the uniform distribution on the set of all admissible plans, and of 'unfair' plans generated by an optimalization algorithm designed to maximize one party's vote share. To ensure that the simulated elections resemble real-life ones, we use empirical data on voting patterns and electoral geography from France, Germany, and Poland.

4.1 Experimental Setup

Our baseline dataset consists of data from 6 French legislative elections (1993–2017), 7 German legislative elections (1998–2021), 9 Polish legislative elections (1991–2019), and one municipal election in a major European city. For each election, our input data consists of a neighborhood graph of territorial units (over 35,000 *communes* / 1,800 *cantons* in the French case, nearly 300 *Wahlkreise* in the German case, close to 2,500 municipalities in the Polish case, and 456 in the municipal case), where each unit k is associated with a tuple $(\mathbf{v}_k, w_k)$, $\mathbf{v}_k \in \Delta_m$ is a vote share vector, and $w_k \in \mathbb{R}_+$ is a population weight.

Remark 4.1. In real life, there can be significant deviations in unit population, but because we are unable to subdivide the largest units (lacking lower-granularity data), we ignore those deviations, instead assigning population numbers to units by transforming their empirical distribution to a lognormal distribution with parameters $\mu = 0$ and $\sigma = 1$.

We use algorithms described in the following two subsections to draw 192 fair districting plans, 64 unfair plans drawn to advantage the largest party, and 64 unfair plans drawn to advantage the second largest party. The output of each districting algorithm consists of a *districting plan*, i.e., a partition of the neighborhood graph into connected components (*districts*) of approximately equal population weight (we permit 20% deviation from the mean). We fix the number of districts to be created at 80 in the French and Polish case, and 50 in the German and municipal cases.

4.2 Algorithms for Generating Fair Plans

To generate fair districting plans we use three Monte Carlo algorithms. The Monte Carlo algorithms are: the Sequential Monte Carlo algorithm by McCartan and Imai [2023], based on drawing random spanning trees and then semi-randomly eliminating edges to partition the graph; the Merge-Split Markov Chain Monte Carlo algorithm by Autrey *et al.* [2019], based on randomly recombining and repartitioning spanning trees; and the FLIP MCMC algorithm by Fifield *et al.* [2015; 2020], using the Swendsen-Wang algorithm [Swendsen and Wang, 1987], as modified by Barbu and Zhu [2005], to randomly walk the graph of solutions.[3] For all three algorithms the stationary distribution is the uniform distribution on the set of admissible partitions. For all three Monte Carlo algorithms, we generate the starting partition using a heuristic Random Seed Growing algorithm by Chen and Rodden [2013; 2015].

4.3 Algorithm for Generating Unfair Plans

To generate unfair districting plans we used an integer linear programming algorithm introduced in [Flis *et al.*, 2023, Ch. 3.7]. We consider all connected components of the input graph, $K_1, \ldots, K_d$, satisfying

$$\left|\left(c\sum_{k\in K_j} w_k\right)\Big/\left(\sum_{k=1}^{c} w_k\right)\right| \in \left(1-\frac{1}{5}, 1+\frac{1}{5}\right) \tag{4.1}$$

for every $j \in [d]$, and solve the following optimization problem for the party to be advantaged, $i \in [m]$:

Problem 4.1. ***For***

$$\boldsymbol{\xi} \in \mathbb{B}^d$$

maximize

$$\sum_{j=1}^{d} \xi_j s_i \left(\sum_{k\in K_j} \mathbf{v}_k w_k\right) \tag{4.2}$$

subject to

$$\sum\nolimits_{j=1}^{d} \xi_j = c, \tag{4.3}$$

$$\sum\nolimits_{j=1}^{d} \mathbf{1}_{K_j}(k) = 1 \textit{ for every } k \in [c], \tag{4.4}$$

where $s_i : \Delta_m \longrightarrow \mathbb{B}$ equals 1 iff the i-th coordinate of the argument is the largest one.

In practice, it is infeasible to enumerate all possible districts with hundreds of precincts. We therefore first artificially combine leaf nodes, small precincts, and similar precincts until the number of precincts is reduced below 320. Only then

[3] This algorithm has a better rate of convergence than classical Metropolis-Hastings, but obtaining satisfactory performance still required additional heuristics like simulated annealing [Marinari and Parisi, 1992; Geyer and Thompson, 1995].

we run the ILP solver and recover the full solution by replacing combined precincts with their original elements. We then run a local neighborhood search to find a local maximum. By randomizing the combination process, we are able to obtain diverse gerrymandering plans of different quality (in terms of advantage to the favored party).

4.4 Results

On the experimental dataset, our method achieves *precision* of .896 and *recall* of .912.

5 Empirical Test

We have tested our method on data from four sets of elections:

1. $\mathcal{D}_{14}$, 2014 Polish municipal elections (2412 instances),
2. $\mathcal{D}_{18}$, 2018 Polish municipal elections (2145 instances),
3. $\mathcal{D}_{U}$, U.S. House elections, 1900-2022 (2848 instances),
4. $\mathcal{D}_{N}$, national legislative elections from 15 countries (206 instances) [Kollman *et al.*, 2023].

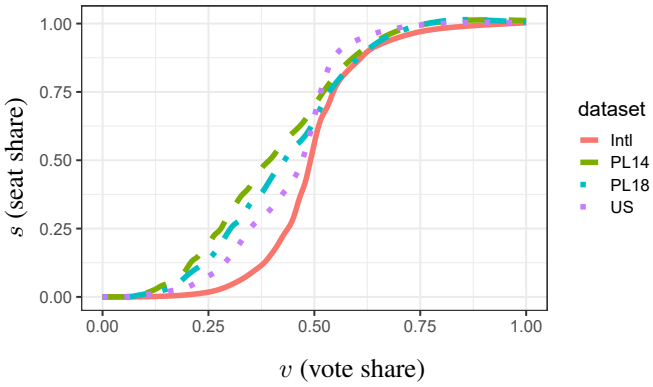

Figure 2: Nonparametric seats-votes curves for the four training sets.

Finding an appropriate baseline for comparison, however, is quite difficult. Even if we were able to easily model the expected scale of random electoral bias, it would still be impossible to determine whether deviation from it is caused by the deficiencies of our method or by actual instances of gerrymandering. Hence, a more appropriate test would be to analyze whether our measure agrees with other methods for detecting gerrymandering used in the literature. In [Stolicki, 2024], one of us tests our method against eight other ones, including partisan bias (two variants), efficiency gap, mean-median difference, and declination coefficient, as well as methods based on express distributional assumptions [Stolicki *et al.*, 2019], Monte Carlo simulations [McCartan and Imai, 2023], and district geometry [Niemi *et al.*, 1990], using data from the 2022 congressional elections. The kernel method exhibits the highest Pearson's correlation coefficient with the normalized aggregate out of all nine, 0.623, suggesting that it best captures the core concept of gerrymandering.

6 Conclusions

We introduce a new method for detecting gerrymandering solely on the basis of aggregate seat and vote counts. It has

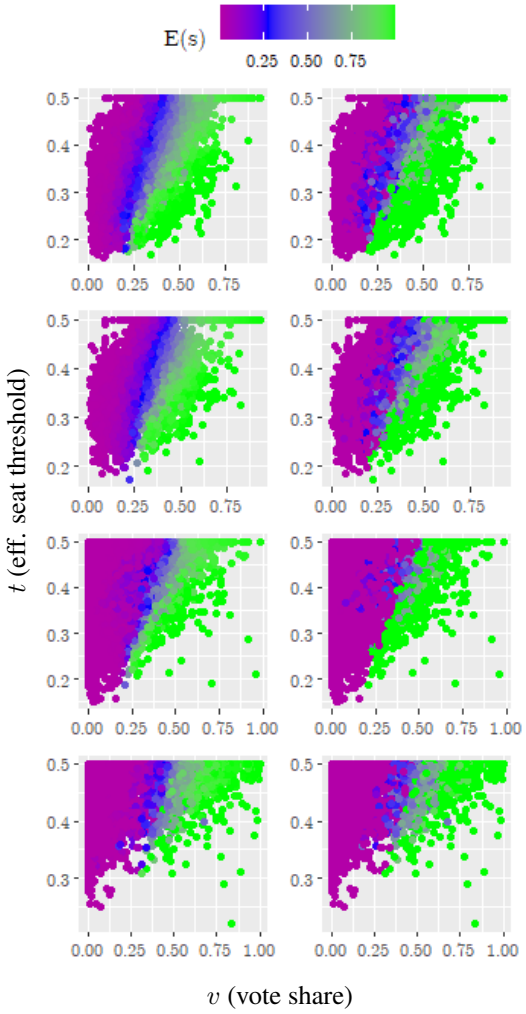

Figure 3: The expected (left) and empirical (right) seat shares as functions of the vote share and the effective seat threshold. From the top: PL2014, PL2018, international, and U.S. elections.

two significant advantages over existing methods: freedom from restrictive theoretical assumptions regarding preference distribution, voter clustering patterns, or seats-votes relations, as the standard of electoral fairness is not derived from such assumptions, but learned from available empirical data; and suitability for use in multi-party elections, including those where independent candidates play a significant role and where different parties contest different districts.

We have tested our method primarily on the basis of experimental (simulation) data, where it performed very well. We have also demonstrated that the application of our method to real-life data is feasible, and that for U.S. elections leads to results that are in agreement with other existing methods.

Future work will focus primarily on developing better benchmarks, testing the method against more simulations based on real-life data as well as purely random elections. We will also seek to explain disagreements between the proposed method and other methods found in the literature.

Acknowledgements

This research has been funded under the European Union's Horizon 2020 research and innovation program, European Research Council (ERC) grant agreement no. 101002854; Polish National Center for Science grants no. 2019/35/B/HS5/03949 and 2014/13/N/HS5/01110; the Jagiellonian University Excellence Inititative, DigiWorld research area, QuantPol Center project; and the French government under management of Agence Nationale de la Recherche as part of the "Investissements d'avenir" program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute).

References

[Abramson, 1982] I. S. Abramson. On Bandwidth Variation in Kernel Estimates – A Square Root Law. *Annals of Statistics*, 10(4):1217–1223, December 1982.

[Akaike, 1974] H. Akaike. A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, December 1974.

[Autrey *et al.*, 2019] E. Autrey, D. Carter, G. Herschlag, Z. Hunter, and J. C. Mattingly. Metropolized Forest Recombination for Monte Carlo Sampling of Graph Partitions. arXiv: 1911.01503 [cs.DS], 2019.

[Barbu and Zhu, 2005] A. Barbu and S.-C. Zhu. Generalizing Swendsen-Wang to Sampling Arbitrary Posterior Probabilities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1239–1253, August 2005.

[Bernstein and Duchin, 2017] M. Bernstein and M. Duchin. A Formula Goes to Court: Partisan Gerrymandering and the Efficiency Gap. *Notices of the American Mathematical Society*, 64(09):1020–1024, October 2017.

[Bhattacharyya, 1943] A. K. Bhattacharyya. On a Measure of Divergence between Two Statistical Populations Defined by Their Probability Distributions. *Bulletin of the Calcutta Mathematical Society*, 35:99–109, 1943.

[Boratyn *et al.*, 2022] D. Boratyn, W. Słomczyński, D. Stolicki, and S. Szufa. Spoiler Susceptibility in Multi-District Party Elections, February 2022.

[Boser *et al.*, 1992] B. E. Boser, I. M. Guyon, and V. N. Vapnik. A Training Algorithm for Optimal Margin Classifiers. In *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, pages 144–152, New York, NY, USA, July 1992. ACM.

[Breiman *et al.*, 1977] L. Breiman, W. Meisel, and E. Purcell. Variable Kernel Estimates of Multivariate Densities. *Technometrics*, 19(2):135–144, May 1977.

[Browning and King, 1987] R. X. Browning and G. King. Seats, Votes, and Gerrymandering: Estimating Representation and Bias in State Legislative Redistricting. *Law & Policy*, 9(3):305–322, July 1987.

[Burkill, 1984] H. Burkill. Monotonic Functions on Partially Ordered Sets. *Journal of Combinatorial Theory, Series A*, 37(3):248–256, November 1984.

[Butler, 1947] D. E. Butler. Appendix III: The Relation of Seats to Votes. In R. B. McCallum and A. Readman, editors, *The British General Election of 1945*, pages 277–292. G. Cumberlege, London, 1947.

[Calvo and Rodden, 2015] E. Calvo and J. Rodden. The Achilles Heel of Plurality Systems: Geography and Representation in Multiparty Democracies. *American Journal of Political Science*, 59(4):789–805, October 2015.

[Calvo, 2009] E. Calvo. The Competitive Road to Proportional Representation: Partisan Biases and Electoral Regime Change under Increasing Party Competition. *World Politics*, 61(2):254–295, April 2009.

[Chen and Rodden, 2013] J. Chen and J. Rodden. Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures. *Quarterly Journal of Political Science*, 8(3):239–269, June 2013.

[Chen and Rodden, 2015] J. Chen and J. Rodden. Cutting Through the Thicket: Redistricting Simulations and the Detection of Partisan Gerrymanders. *Election Law Journal: Rules, Politics, and Policy*, 14(4):331–345, December 2015.

[Cortes and Vapnik, 1995] C. Cortes and V. Vapnik. Support-Vector Networks. *Machine Learning*, 20(3):273–297, September 1995.

[Erikson, 1972] R. S. Erikson. Malapportionment, Gerrymandering, and Party Fortunes in Congressional Elections. *American Political Science Review*, 66(4):1234–1245, 1972.

[Fifield *et al.*, 2015] B. Fifield, M. Higgins, and K. Imai. A New Automated Redistricting Simulator Using Markov Chain Monte Carlo. Working Paper, Princeton University, Princeton, NJ, 2015.

[Fifield *et al.*, 2020] B. Fifield, M. Higgins, K. Imai, and A. Tarr. Automated Redistricting Simulation Using Markov Chain Monte Carlo. *Journal of Computational and Graphical Statistics*, 29(4):715–728, October 2020.

[Flis *et al.*, 2023] J. Flis, W. Słomczyński, D. Stolicki, S. Szufa, and J. Sokołowski. *Skrzywione szranki. Na tropie gerrymanderingu w polskich wyborach do rad gmin*. Wydawnictwo UJ, Kraków, 2023.

[Gelman and King, 1990a] A. Gelman and G. King. Estimating Incumbency Advantage without Bias. *American Journal of Political Science*, 34(4):1142–1164, November 1990.

[Gelman and King, 1990b] A. Gelman and G. King. Estimating the Electoral Consequences of Legislative Redistricting. *Journal of the American Statistical Association*, 85(410):274–282, 1990.

[Gelman and King, 1994] A. Gelman and G. King. A Unified Method of Evaluating Electoral Systems and Redistricting Plans. *American Journal of Political Science*, 38(2):514–554, 1994.

[Geyer and Thompson, 1995] C. J. Geyer and E. A. Thompson. Annealing Markov Chain Monte Carlo with Applications to Ancestral Inference. *Journal of the American Statistical Association*, 90(431):909–920, September 1995.

[Gilligan and Matsusaka, 1999] T. W. Gilligan and J. G. Matsusaka. Structural Constraints on Partisan Bias Under the Efficient Gerrymander. *Public Choice*, 100(1):65–84, July 1999.

[Grofman, 1983] B. Grofman. Measures of Bias and Proportionality in Seats-Votes Relationships. *Political Methodology*, 9(3):295–327, 1983.

[Herfindahl, 1950] O. C. Herfindahl. *Concentration in the U.S. Steel Industry*. PhD thesis, Columbia University, New York, NY, 1950.

[Hirschman, 1945] A. O. Hirschman. *National Power and the Structure of Foreign Trade*. University of California Press, Berkeley, CA, 1945.

[Hurvich *et al.*, 1998] C. M. Hurvich, J. S. Simonoff, and C.-L. Tsai. Smoothing Parameter Selection in Nonparametric Regression Using an Improved Akaike Information Criterion. *Journal*

of the Royal Statistical Society: Series B, 60(2):271–293, 1998.

[Härdle *et al.*, 1988] W. Härdle, P. Hall, and J. S. Marron. How Far are Automatically Chosen Regression Smoothing Parameters from their Optimum? *Journal of the American Statistical Association*, 83(401):86–95, March 1988.

[Kendall and Stuart, 1950] M. G. Kendall and A. Stuart. The Law of the Cubic Proportion in Election Results. *British Journal of Sociology*, 1(3):183–196, September 1950.

[King, 1990] G. King. Electoral Responsiveness and Partisan Bias in Multiparty Democracies. *Legislative Studies Quarterly*, 15(2):159–181, 1990.

[Kollman *et al.*, 2023] K. Kollman, A. Hicken, D. Caramani, D. Backer, and D. Lublin. Constituency-Level Elections Archive. Dataset, Center for Political Studies, University of Michigan, Ann Arbor , MI, 2023.

[Kullback and Leibler, 1951] S. Kullback and R. A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[Laakso and Taagepera, 1979] M. Laakso and R. Taagepera. "Effective" Number of Parties: A Measure with Application to West Europe. *Comparative Political Studies*, 12(1):3–27, April 1979.

[Li and Racine, 2004] Q. Li and J. Racine. Cross-Validated Local Linear Nonparametric Regression. *Statistica Sinica*, 14:485–512, 2004.

[Linzer, 2012] D. A. Linzer. The Relationship between Seats and Votes in Multiparty Systems. *Political Analysis*, 20(3):400–416, 2012.

[Manow, 2011] P. Manow. The Cube Rule in a Mixed Electoral System: Disproportionality in German Bundestag Elections. *West European Politics*, 34(4):773–794, July 2011.

[March, 1957] J. G. March. Party Legislative Representation as a Function of Election Results. *Public Opinion Quarterly*, 21(4):521–542, 1957.

[Marinari and Parisi, 1992] E. Marinari and G. Parisi. Simulated Tempering: A New Monte Carlo Scheme. *Europhysics Letters (EPL)*, 19(6):451–458, July 1992.

[McCartan and Imai, 2023] C. McCartan and K. Imai. Sequential Monte Carlo for Sampling Balanced and Compact Redistricting Plans. *The Annals of Applied Statistics*, 17(4):3300–3323, December 2023.

[McDonald *et al.*, 2011] M. D. McDonald, J. Krasno, and R. E. Best. An Objective and Simple Measure of Gerrymandering: A Demonstration from New York State. In *2011 Annual Meeting of the Midwest Political Science Association*, Chicago, IL, 2011.

[McGann *et al.*, 2015] A. J. McGann, C. A. Smith, M. Latner, and J. A. Keena. A Discernable and Manageable Standard for Partisan Gerrymandering. *Election Law Journal*, 14(4):295–311, November 2015.

[McGhee, 2014] E. McGhee. Measuring Partisan Bias in Single-Member District Electoral Systems. *Legislative Studies Quarterly*, 39(1):55–85, February 2014.

[Nadaraya, 1965] E. A. Nadaraya. On Non-Parametric Estimates of Density Functions and Regression Curves. *Theory of Probability & Its Applications*, 10(1):186–190, January 1965.

[Niemi and Deegan, 1978] R. G. Niemi and J. Deegan. A Theory of Political Districting. *American Political Science Review*, 72(4):1304–1323, December 1978.

[Niemi *et al.*, 1990] R. G. Niemi, B. Grofman, C. Carlucci, and T. Hofeller. Measuring Compactness and the Role of a Compactness Standard in a Test for Partisan and Racial Gerrymandering. *Journal of Politics*, 52(4):1155–1181, 1990.

[Parzen, 1962] E. Parzen. On Estimation of a Probability Density Function and Mode. *Annals of Mathematical Statistics*, 33(3):1065–1076, September 1962.

[Rosenblatt, 1956] M. Rosenblatt. Remarks on Some Nonparametric Estimates of a Density Function. *Annals of Mathematical Statistics*, 27(3):832–837, September 1956.

[Schucany, 1995] W. R. Schucany. Adaptive Bandwidth Choice for Kernel Regression. *Journal of the American Statistical Association*, 90(430):535–540, June 1995.

[Shannon, 1948] C. E. Shannon. A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27(3):379–423, July 1948.

[Silverman, 1986] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, New York, 1986.

[Simpson, 1949] E. H. Simpson. Measurement of Diversity. *Nature*, 163(4148):688–688, April 1949.

[Spearman, 1904] C. Spearman. The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*, 15(1):72–101, 1904.

[Stephanopoulos and McGhee, 2015] N. O. Stephanopoulos and E. M. McGhee. Partisan Gerrymandering and the Efficiency Gap. *University of Chicago Law Review*, 82(2):831–900, 2015.

[Stolicki *et al.*, 2019] D. Stolicki, W. Słomczyński, and J. Flis. A Probabilistic Model for Detecting Gerrymandering in Partially-Contested Multiparty Elections. In *Transactions on Computational Collective Intelligence XXXIV*, volume 11890 of *Lecture Notes in Computer Science*, pages 1–20, Berlin, Heidelberg, 2019. Springer.

[Stolicki, 2024] D. Stolicki. Redistricting and Gerrymandering after the 2020–2021 Census. In R. Duda and M. Turek, editors, *The Crossroads Elections: European Perspectives on the 2022 U.S. Midterm Elections*, pages 76–95. Routledge, New York, 1 edition, March 2024.

[Swendsen and Wang, 1987] R. H. Swendsen and J.-S. Wang. Nonuniversal Critical Dynamics in Monte Carlo Simulations. *Physical Review Letters*, 58(2):86–88, January 1987.

[Słomczyński *et al.*, 2024] W. Słomczyński, D. Stolicki, and S. Szufa. Nonparametric Detection of Gerrymandering in Multiparty Elections. arXiv: 2306.03085 [stat.AP], May 2024.

[Taagepera and Grofman, 1981] R. Taagepera and B. Grofman. Effective Size and Number of Components. *Sociological Methods & Research*, 10(1):63–81, August 1981.

[Taagepera, 1973] R. Taagepera. Seats and Votes: A Generalization of the Cube Law of Elections. *Social Science Research*, 2(3):257–275, September 1973.

[Taagepera, 1986] R. Taagepera. Reformulating the Cube Law for Proportional Representation Elections. *American Political Science Review*, 80(2):489–504, June 1986.

[Tufte, 1973] E. R. Tufte. The Relationship between Seats and Votes in Two-Party Systems. *American Political Science Review*, 67(2):540–554, June 1973.

[Wang, 2016] S. S.-H. Wang. Three Tests for Practical Evaluation of Partisan Gerrymandering. *Stanford Law Review*, 68(6):1263–1321, 2016.

[Warrington, 2018] G. S. Warrington. Quantifying Gerrymandering Using the Vote Distribution. *Election Law Journal*, 17(1):39–57, March 2018.

[Watson, 1964] G. S. Watson. Smooth Regression Analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, 26(4):359–372, 1964.