Big City vs. the Great Outdoors: Voter Distribution and How it Affects Gerrymandering

Allan Borodin, Omer Lev, Nisarg Shah and Tyrone Strangway

Abstract

Gerrymandering is the process by which parties manipulate boundaries of electoral districts in order to maximize the number of districts they can win. Demographic trends show an increasingly strong correlation between residence and party affiliation; some party's supporters congregate in cities, while others stay in more rural areas. We investigate both theoretically and empirically the effect of this trend on a party's ability to gerrymander in a two-party model ("urban party" and "rural party"). Along the way, we propose a definition of the gerrymandering power of a party, and an algorithmic approach for near-optimal gerrymandering in large instances.

Our results suggest that beyond a fairly small concentration of urban party's voters, the gerrymandering power of a party depends almost entirely on the level of concentration, and not on the party's share of the population. As partisan separation grows, the gerrymandering power of both parties converge so that each party can gerrymander to get only slightly more than what its voting share warrants, bringing about, ultimately, a more representative outcome. Moreover, there seems to be an asymmetry between the gerrymandering power of the parties, with the rural party being more capable of gerrymandering.

1 Introduction

The question of how to aggregate various peoples' preferences and choose a single option has existed for millennia and is a fundamental issue in social choice theory and practice. Since the early days when humans institutionalized their decision making, manipulation came along with it [31], as people looked for ways to change the outcome more to their liking, and leaders looked for ways to reduce their rivals' power [19].

Jumping ahead several centuries, during the late middle ages various European kingdoms established assemblies which included representatives from various areas of their country: the English parliament, the French Estates General, and others. In those assemblies which survived to become significant policy bodies, the issue of which areas received representation became crucial. In Britain, for example, the parliamentary districts still reflected medieval population distribution until the Great Reform Bill of 1832. But just as Britain was trying to solve its district allocation problem, across the Atlantic, American politicians were realizing the potential power that comes with the ability to divide a state into its districts. In 1812, Massachusetts governor Elbridge Gerry gave his name to the practice of gerrymandering – creating oddly shaped districts for political gain.

Dividing a geographic area (e.g., a state into electoral districts; municipality lines for allocating municipal taxes) into subareas subject to some constraints is a problem for which there are many possible solutions, and one of them needs be selected. Gerrymandering is a control problem, in which the agent in charge of the division optimizes it for their personal preferred outcome, even when there is a more natural one. From here on we will use political nomenclature, but note that all the results and observations apply to dividing any geographic area with resources (partisan voters are only one such possibility) between agents in a way that can be fair, but can also be highly biased.

There has been much research concerning gerrymandering in the United States, and since the 1965 Voting Rights Act particular attention has been given to minority representation issues. This interest has accelerated in the past two decades (and even more so in todays' political atmosphere),

with much effort devoted to sharp denunciations of gerrymandering, and arguing that it is a significant danger to the US political wellbeing [21, 24]. However, despite substantial effort, it is still not clear what constitutes a "good" or "fair" district map [35]. Is it district compactness? Is it population homogeneity [36]? But the main criticism of gerrymandering seems to be that it is producing results that are unrepresentative of the overall population's desires and preferences [25].

In parallel to this debate, a different dynamic is taking place in the United States: for a variety of reasons, individuals are choosing to reside in such a way that they end up living near people with similar party affiliation. In particular, supporters of one party cluster in urban areas, while rural areas are fast becoming the domain of their political opponents [4]. Many commentators are mixing this effect with the negative effect of gerrymandering [11, 28].

We examine the relation between gerrymandering and partisanship distribution. We explore several theoretical insights, and construct a simulation tool and an algorithm to find highly gerrymandered district divisions. We propose a novel metric, *gerrymandering power*, which measures how much gerrymandering can make a party powerful beyond its vote share in the population. We work with a synthetic grid map in order to focus more on the effect of a party's vote share and the distribution of its voters on its gerrymandering power.

Contrary to common opinion [34], our results suggest that as partisan separation grows and urban voters cluster, the opposite party's ability to gerrymander *drops*. Surprisingly, once a certain (fairly low) urban party concentration is exceeded, a party's gerrymandering power seems to depend almost entirely on the density of the urban center and not on its share of the population. As partisan separation grows, the gerrymandering power of both parties converges so that the parties are limited in their ability to gain much more than what their share of the population warrants, bringing about, ultimately, a more representative outcome. We also observe complex effects in close elections with moderate concentration levels. Moreover, our results also suggest a *basic asymmetry between the gerrymandering power of the urban and rural parties*, wherein the rural party has a stronger gerrymandering ability.

2 Related Work

Research on gerrymandering has been done, throughout the years, from sociological vantage points [23], historical ones [10, 5], and, particularly since the 1965 Voting Rights Act, legal ones [29, 18, 15]. Naturally, however, it has been mainly explored in the political science arena [12], primarily based on analysis of past elections [16, 32, 13] – trying to figure out if it occurred, and trying to calculate some measure of its effects. In the past few years the computational social choice community has also taken interest in this topic, on issues such as worst case analysis of how districts effect voting rules [3], and the computational complexity of gerrymandering [22, 33]. Recently, Pegden et al. [26] suggested a "cut and choose"-like mechanism to divide districts in practice.

The work of Lewenberg et al. [22] is closest to ours, they also provide an algorithm for gerrymandering over a graph. Unlike our algorithm, their greedy algorithm produces districts which may differ in population by up to 6500%. However, most US congressional districts must be within 1% of each other, and that algorithm will fail when trying to reach this constraint.

Finding an optimal gerrymandering division of a geographical area has long been viewed as a planar graph related problem, in which precincts (which are, practically, our undivisible smallest unit) are nodes in the graph, and one looks for cuts in the graph that will result in sub-graphs with particular properties (e.g., contiguity). Dyer and Frieze [9] hypothesized that even the complexity of finding a division of the graph to equal-sized connected parts (akin to contiguous, equal population districts) is NP-hard, and had several related results which seem to indicate that this is, indeed, the case. Further papers have further tried to attack this problem (e.g., [37]), but without significant breakthrough. Apollonio et al. [2] limited themselves to the grid, as we do, and found some bounds on gerrymandering there.

The discussion of algorithmically finding optimal districts has been with us since it became feasible to consider such an option in the '60s (see summary in Altman [1]), and work on it has gone hand-in-hand with considering what are metrics to measure gerrymandering, and avoiding such settings (see [35, 17] on the various metrics that have been suggested). Puppe and Tasnádi [27] axiomatize a districting division that strives to optimize gerrymandering for one of the parties. More practically, Fifield et al. [14], tried to produce a random sample of district maps under some constraints, suggest a method that takes an existing partition of the graph, and slowly changes it, as it slowly "swaps" precincts bordering on the dividing line between districts. To achieve a similar goal, Chen and Cottrell [6] take a more classic local search approach.

The observation that voter distribution is not uniform, and instead follows a clustering of one party into dense cities, was made prominent by the book "The Big Sort" [4]. Following research has corroborated this observation [8, 7].

3 Model

We use a graph-theoretic formulation of the districting problem. In our formulation, a state is represented by a graph G, where the vertex set $V(G)$ contains a vertex for every precinct, and the edge set $E(G)$ contains an edge between every pair of precincts that share a physical boundary. Because the map of a state is two-dimensional, we assume that G is planar. Let n_v denote the number of voters in vertex v. For simplicity, we assume that voters are divided between two major political parties, P_1 and P_2. For $P \in \{P_1, P_2\}$, let n_v^P denote the number of voters of party P in vertex v, and let $N^P = \sum_{v \in V(G)} n_v^P$ denote the total number of voters of party P. Let $N = N^{P_1} + N^{P_2}$ denote the total number of voters. We use $\alpha^{P_1} = N^{P_1}/N$ and $\alpha^{P_2} = N^{P_2}/N$ to denote the proportional vote shares of the two parties.

Given a desired number of districts $K \in \mathbb{N}$, the *districting problem* is to partition the graph G into K vertex-disjoint subgraphs $G_1, \ldots, G_K$ (called *districts*) that satisfy a number of constraints. In this work, we focus on two constraints that exist widely in practice.

1. *Contiguity.* For each district $k \in [K]$, G_k must be a connected subgraph of G.

2. *Equal Population.* The total number of voters in each district should be approximately equal. Formally, given a tolerance level δ, we need that for each $k \in [K]$,

$$1 - \delta \leq \frac{\sum_{v \in V(G_k)} n_v}{N/K} \leq 1 + \delta.$$

We say that a districting is valid if it satisfies both these constraints. Let $\mathcal{R}$ denote the set of valid districtings. There are additional criteria that districting should satisfy such as compactness of districts, preservation of existing political communities, and racial fairness. However, we overlook these criteria in this work, as there is still work to be done on formulating a consensus on their quantitative definitions. Given a districting $R \in \mathcal{R}$, we say that party P wins district k if it has a majority in the district: $\sum_{v \in V(G_k)} n_v^P > (1/2) \cdot \sum_{v \in V(G_k)} n_v$[1]. Let $K^P(R)$ denote the number of districts won by party P in districting R, and let $\sigma^P(R) = K^P(R)/K$.

There may of course be many solutions to the districting problem satisfying the contiguity and equal population constraints. The goal of *gerrymandering* is to find the districting that maximally favors one party. In this work, we focus on partisan gerrymandering (henceforth, simply gerrymandering) where the goal is to maximally favor a given party. Partisan fairness would require choosing a districting R in which $\sigma^P(R)$ is as close to α^P as possible. We define the *gerrymandering power* of party P to be $\max_{R \in \mathcal{R}} \sigma^P(R) - \alpha^P$, i.e., the maximum boost the party can get through gerrymandering above their proportional share of the districts. Note that negative gerrymandering power

---

[1] We break ties in favour of the party which controls the districting procedure

implies that the voters are distributed in such a way that the party falls short of its proportional share of the districts even with maximum gerrymandering.

In this paper, we make some simplifying assumptions. First, we assume that graph G is an $n \times n$ grid. Grids are among the simplest planar graphs that still present non-trivial challenges. Second, we assume that each vertex of the grid has an equal number of voters: let $n_v = T$ for each $v \in V(G)$, for a sufficiently large constant T. Third, we mandate that all districts be of equal size, i.e., we set $\delta = 0$ in the equal population constraint. Finally, we assume voter preferences to be fixed. While these assumptions drag our model a bit farther from reality, they allow us to focus on the dependence of a party's gerrymandering power on its vote share and the geographic distribution of its voters. As we discuss in Section 7, we believe our observations would not change qualitatively when moving to general graphs as the key insights we derive in Sections 4 and 6 are applicable to general graphs.

4   A Worst-Case Viewpoint

Our goal is to study the effect of voters' geographic distribution on the gerrymandering power of the parties. In this section, we take a worst-case point of view: *How does the gerrymandering power of a party change with its vote share when its voters are distributed in the worst possible way?* Formally, given a party P and its vote share α^P, we want to analyze the maximum fraction of districts the party can win in the worst case choice of $\{n_v^P\}_{v \in V(G)}$ that satisfies $0 \le n_v^P \le T$ for each $v \in V(G)$ and $\sum_{v \in V(G)} n_v^P = \alpha^P \cdot N$. For the grid graph, $N = n^2 \cdot T$ is the total number of voters.

We begin by making an interesting observation in the large-graph limit. Imagine the $n \times n$ grid embedded in a bounded convex region. As $n \to \infty$, one can treat the graph as a continuous convex region in $\mathbb{R}^2$ endowed with two measures μ^{P_1} and μ^{P_2} that represent how the voters of the two parties are distributed across the region. In this case, we show that there is a sharp transition in which a party can win every district or no district depending on whether it has a majority or a minority vote share.

The idea of the proof is as follows. When $\alpha^P < 1/2$, party P clearly wins no districts if its voters are uniformly spread, i.e., if it has α^P fraction of the voters in each individual vertex. When $\alpha^P \ge 1/2$, we invoke a generalization of the popular Ham Sandwich Theorem [30, 20], which states that given d measures in $\mathbb{R}^d$, there exist K interior-disjoint convex partitions that divide each measure equally. Applying this to measures μ^{P_1} and μ^{P_2} in $\mathbb{R}^2$, we get a valid districting in which party P has a majority in every district.

Theorem 1. *Suppose an n x n grid is embedded into a bounded convex region. As $n \to \infty$, for every $K \in \mathbb{N}$, party P (which controls the districting) can guarantee winning every district if its vote share is $\alpha^P \ge 1/2$, and wins no districts in the worst case if its vote share is $\alpha^P < 1/2$.*

When n is finite and $\alpha^P < 1/2$, uniform voter distribution still remains a worst case for party P regardless of the number of districts K, and prevents the party from winning any district. However, the case of $\alpha^P \ge 1/2$ becomes more fine-grained. For a constant K, increasing the graph size (i.e., increasing n) gives the party more gerrymandering power. We illustrate this using the case of two districts ($K = 2$). For $n = 2$ (i.e., in a 2×2 grid), it is easy to show that a party needs 75% vote share to win both districts in the worst case.

Proposition 1. *For $n = K = 2$, the following holds for party P in the worst case.*

1. *If $\alpha^P \ge 3/4$, the party wins both districts.*

2. *If $1/2 \le \alpha^P < 3/4$, the party wins a single district.*
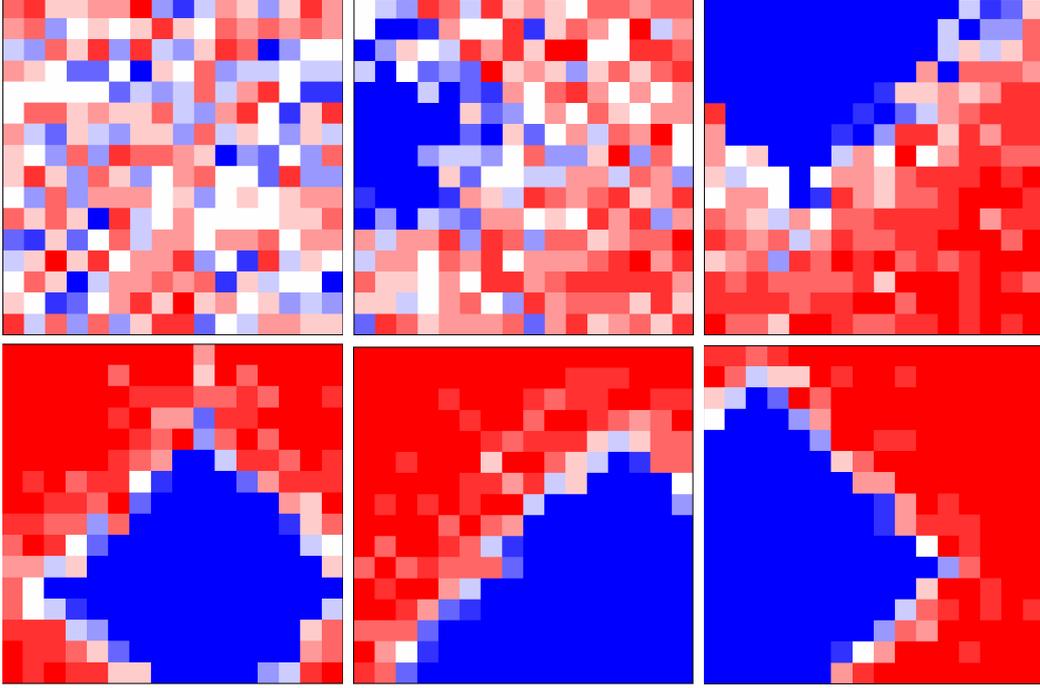
3. *If $\alpha^P < 1/2$, the party wins no districts.*

Figure 1: Partisanship distribution for $\alpha^U = 0.45$ and various values of ϕ ($\phi = 0, 1, 3$ in the top row from left to right and $\phi = 5, 7, 9$ in the bottom row from left to right). Blue/red represents a majority of U/R voters, and colour intensity increases with the majority strength.

Proof. Let us number the precincts in the 2×2 grid as $\{1, 2, 3, 4\}$ so that the number increases from left to right and from top to bottom. For $i \in \{1, 2, 3, 4\}$ and $P \in \{P_1, P_2\}$, let n_i^P denote the number of voters of party P in precinct i. Suppose party P is gerrymandering. Note that there are two ways to district: either each row is a district, or each column is a district.

Suppose $\alpha^P \geq 3/4$, i.e., $\sum_{i=1}^4 n_i^P \geq 3N/4$. We want to show that party P holds a weak majority either in each row or in each column. Suppose this is not the case. Without loss of generality, suppose it holds a minority in the top row ($n_1^P + n_2^P < N/4$) and in the left column ($n_1^P + n_3^P < N/4$). Summing the two equations, and subtracting from the total vote share of party P, we obtain $n_4^P - n_1^P > N/4$, which is impossible since each precinct has $N/4$ voters.

Next, suppose $1/2 \leq \alpha^P < 3/4$. It is easy to see that regardless of the districting, party P must win at least one district because $\alpha^P \geq 1/2$. For any $\epsilon > 0$, we want to show an instance with $\alpha^P = 3/4 - \epsilon$ in which party P cannot win both districts. One such instance is given by $n_1^P = n_4^P = N(1/4 - \epsilon/2), n_2^P = N/4$, and $n_3^P = 0$.

Finally, let $\alpha^P < 1/2$. If the voters are uniformly spread, the party trivially cannot win a majority in any district, regardless of the districting. $\square$

However, as n increases, we can show that the required vote share for winning both districts quickly converges to the 50% limit indicated by Theorem 1. In the next result, we only consider even n because creating two districts of equal size is impossible when n is odd.

Theorem 2. *For even n and $K = 2$, a party can find a districting where it wins both distrcits if its vote share is at least $1/2 + 1/n$.*

Proof. Consider an $n \times n$ grid. Suppose party P has vote share $\alpha^P \geq 1/2 + 1/n$. We want to show that there exists a valid districting in which the party wins both districts.

To take care of the contiguity and equal population constraints, let us impose a specific structure on the districting. We assign the top row consisting of n vertices to district 1, and the bottom row consisting of n vertices to district 2. This leaves n columns of height $n-2$ each, which we call *strips*. Note that every solution in which $n/2$ strips are assigned to each district gives a valid districting. We want to show that one such assignment results in party P winning both districts.

Suppose this is not true. Consider the assignment that maximizes the minimum vote share of party P across the two districts. Without loss of generality, suppose party P wins district 1, but loses district 2. Let n_1^P, n_2^P, and n_t^P denote the number of voters of party P in district 1, district 2, and a strip t, respectively. Recall that the total number of voters is N.

Since party P loses in district 2, which has $N/2$ voters, we have $n_2^P < N/4$. Hence, there exists a strip t in district 2 such that $n_t^P \leq n_2^P/(n/2) < (N/4)/(n/2) = N/(2n)$.

On the other hand, we have $n_1^P = \alpha^P N - n_2^P > \alpha^P N - N/4$. Even after discounting the top row which has N/n voters, there must exist a strip t' in district 1 such that

$$n_{t'}^P \geq \frac{\alpha^P \cdot N - N/4 - N/n}{n/2}. \tag{1}$$

Let us consider the (valid) districting obtained by exchanging strips t and t' between the two districts. We observe that party P still wins district 1 because by losing strip t', it loses at most N/n of its own voters, and

$$n_1^P - \frac{N}{n} > \alpha^P \cdot N - \frac{N}{4} - \frac{N}{n} \geq \frac{N}{4},$$

where the last inequality follows because $\alpha^P \geq 1/2 + 1/n$. On the other hand, district 2 now has strictly more voters of party P because it loses at most $n_t^P < N/(2n)$ such voters, but gains at least $n_{t'}^P$ such voters. From Equation (1) and the fact that $\alpha^P \geq 1/2 + 1/n$, it readily follows that $n_{t'}^P \geq N/(2n)$. We have created a partition with a higher minimal voting share for party P, contradicting our construction. $\square$

We remark that the districts formed in the proof of Theorem 2, while contiguous, can be far from being compact. In Section 7, we discuss how adding a formal requirement of compactness may influence our results.

While the party with a majority vote share can easily gerrymander large graphs when K is fixed, it is much more difficult to do so when K is large as well. At the extreme, when $K = n^2$, it is easy to show that party P wins $\max(0, 2\alpha^P - 1)$ fraction of the districts in the worst case. This fraction is zero for $\alpha^P \leq 1/2$, and linearly increases to 1 as α^P goes to 1. This is in sharp contrast to Theorem 2, where the fraction jumps from 0 to 1 when going from $\alpha^P = 1/2$ to $\alpha^P = 1/2 + 1/n$.

While our results are for the extreme cases (Theorem 1 holds as n goes to infinity and Theorem 2 holds for $K = 2$), the worst-case viewpoint lends us to a key insight: the gerrymandering power of a party significantly depends on the relationship between n and K. While large graphs are easy to gerrymander, a large number of districts make it hard to gerrymander.

5   Simulating Optimal Gerrymandering

We now conduct an empirical study of the gerrymandering power of political parties. Instead of the worst case partisanship distribution we considered in the previous section, we adopt a more realistic model based on the urban-rural divide referenced in the introduction. We also use grid graphs with a less extreme ratio of the graph size to the number of districts (in fact, we use numbers that are similar to some states within the American Congressional system).

5.1 An Urban-Rural Model

To model an urban-rural divide on a graph G, we use two parameters. The fraction of the urban party U's voters in G, $\alpha^U \in [0, 1]$, and the strength of an urban-rural divide $\phi \in \mathbb{R}_{\geq 0}$. Given G, ϕ and α^U, we use the following process:

1. Set all voters in G to be for the rural party R.

2. Pick a set of urban centres $C \subset V$ randomly. For $v \in V$, let $d(v)$ be the minimum distance of v to any $c \in C$.

3. Pick a node v (with at least one R voter left) with probability proportional to $\frac{1}{1+(d(v))^\phi}$.

4. Convert one of its R voters into a U voter.

5. Repeat steps 3 and 4 until the fraction of U voters in G is at least α^U.

See Figure 1 for sample heat maps generated by this process. Note that in step 3, we pick a node with a probability that decays polynomially in $d(v)$; we also conducted experiments with exponentially decaying probabilities and did not notice a qualitative difference in our results.
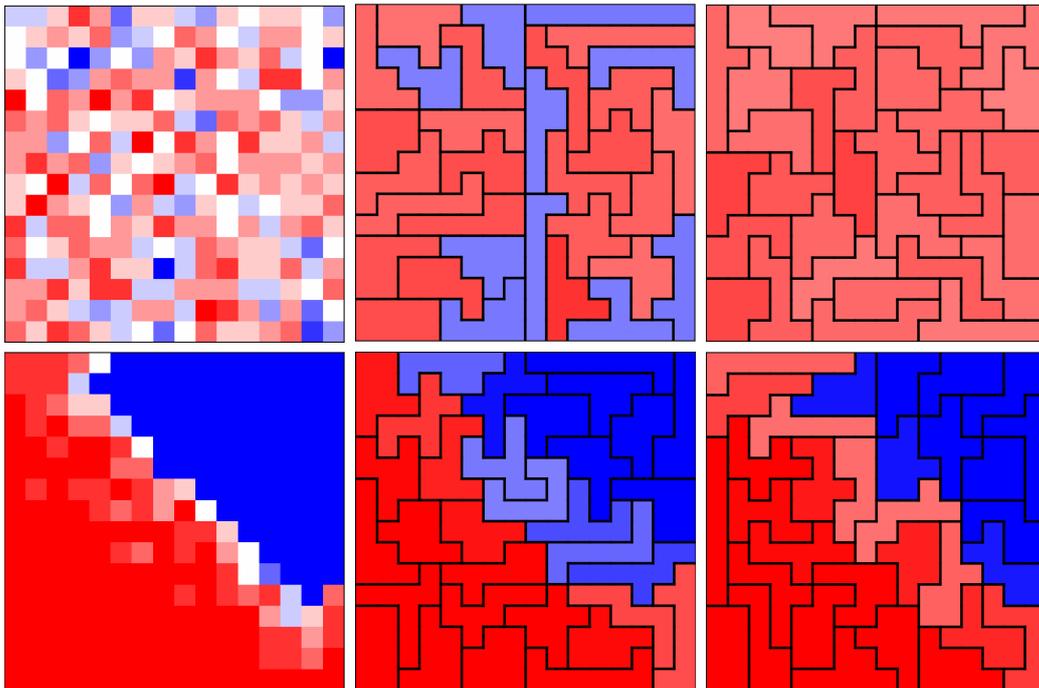

Figure 2: An example of the gerrymandered solutions B^+ found when $\alpha^U = 0.4$. Top row $\phi = 0$ bottom row $\phi = 10$. First column is partisan distribution of voters (similar to Figure 1). Second column is the gerrymandered solution for U, third column is the gerrymandered solution for R. Colour intensity within a district increases with the strength of victory.

5.2 An Algorithm to Gerrymander

Our starting point for an algorithm for optimal gerrymandering is to formulate a Mixed Integer Linear Program (MILP), which uses network flow constraints to ensure connectedness of the districts.

We omit the details due to lack of space. Unfortunately, this program does not scale well, and takes hours on grids with a hundred nodes. Let us call the MILP approach algorithm A. We devise a bottom-up algorithm B, which uses A as a subroutine to optimally solve small sub-problems with at most β nodes. To divide G into K components in favour of party P, algorithm B works as follows:

1. Find an arbitrary division of G into K connected components $(G_1 \cdots G_K)$ of equal or near-equal size.

2. Randomly pick two adjacent components G_i and G_j.

3. Merge them into a new component $G_M = G_i \cup G_j$. If $|V(G_M)| \leq \beta$, use algorithm A to optimally gerrymander G_M into K' districts, where K' is the number of districts in G_i and G_j. Otherwise, let the districting of G_M be dictated by the districtings of G_i and G_j.

4. Repeat steps 2 and 3 until there is one component left.

Finally, we chain algorithm B with itself by feeding the districting found in one execution of B to step 1 in the next execution of B, and repeating until there are no improvements. We call this algorithm B^+. While the algorithm is not guaranteed to find an optimal gerrymandering, we see (see Section 5.4) that it finds highly gerrymandered districting on large instances; in contrast, algorithm A simply fails to work for large instances.

In order to find a districting for step 1 in the first execution of algorithm B, we simply use our MILP but without an objective function, which is reasonably fast. Once we find one valid districting, we can find more for different executions of B^+ using an iterative process I, where we take a pair of adjacent districts G_i and G_j, find one node from each district such that exchanging them gives another valid districting (if possible), and repeat this for a number of steps.

5.3 Simulation Setup

For all of our experiments we use a 16×16 grid graph (i.e., 256 nodes) with 10 voters per node, and divide it into 32 equal sized districts. This problem size is about the same as Vermont's state senate (270 precincts and 30 districts). For the urban party vote share, we use $\alpha^U \in \{0.40, 0.45, 0.48, 0.5, 0.52, 0.55, 0.6\}$, and for the strength of the urban-rural divide, we use $\phi \in \{0, 1, \ldots, 10\}$. Using our urban-rural model, we generate 20 graphs G for each combination of α^U and ϕ, each with a randomly chosen urban centre (more centres would be too crammed with $n = 16$). For each G, we run B^+ 20 times to find the best gerrymandering for each party[2]. To generate the 20 sufficiently different starting points, we use process I with 100,000 swaps. We use IBM CPLEX for solving the MILP in algorithm A, and use $\beta = 16$, i.e., we solve instances with at most 16 nodes optimally using algorithm A. Overall, when provided with a starting point, algorithm $B+$ was able to solve any of our instances within 2 minutes. Finding a starting point did take significantly longer, but since all of our experiments were on the same graph structure, this point could be reused for generating starting points for all our problem instances.

5.4 Some Basic Results

In Section 4, we show that for $K = 2$, a party needs at least 50% vote share to guarantee winning at least one district in the worst case. In our simulations with a moderate urban-rural divide ($\phi = 5$), we observe that just 26% vote share allows a party to win one district with n as low as 8.

For most combinations of α^U and ϕ, our approach was able to secure more districts for the gerrymandering party than its proportional vote share, resulting in a positive gerrymandering power.

---

[2]The gerrymandering outcomes were fairly consistent for fixed values of ϕ and α^U. The standard deviation over the 20 maps was always under 1.4; in comparison the number of districts was 32.
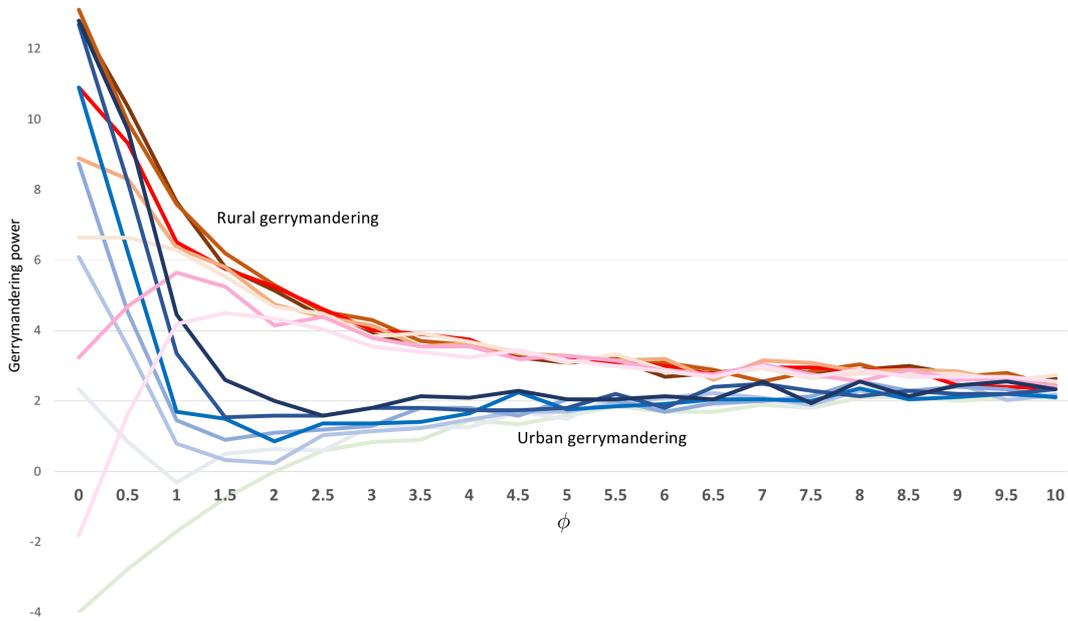
Figure 3: The average gerrymandering power of the parties versus ϕ. The urban/rural party is in blue/red, and a darker colour represents a higher vote share of the gerrymandering party.

The one exception (which we will elaborate on later) was the case of highly unbalanced vote shares with completely homogenous precincts. See Figure 2 and Figure 5 for examples of gerrymandered solutions our algorithm found. Furthermore, we also tested our algorithm on several examples (on the grid), where we knew the optimal gerrymandering outcomes for each party. Our algorithm was always able to win at least 60% of the maximum possible districts, and often it was able to find the optimal solution. We note that often these examples had a unique optimal outcome, so discovering it was difficult. Furthermore we tested the algorithm both with examples similar in size to those in our simulations and ones which were much larger (grids where $n = 64$).

6   Simulation Results

We now describe the results of our simulations, and explain several important trends based on three key figures. Figure 3 shows the gerrymandering power of the two parties for different vote shares as a function of the urban-rural divide.

6.1   Highly Unbalanced Elections

In unbalanced elections with vote share difference of at least 20% (see Figure 4 and Figure 2), we see an expected trend. At $\phi = 0$, when the voters are spread uniformly at random, the party with a majority vote share holds a majority in most precincts despite the randomness in our generation process. This makes it trivial for the majority party to gerrymander to win almost all districts, but difficult for the minority party to gerrymander well. In fact, the minority party has a negative gerrymandering power, i.e., it wins less fraction of districts than its vote share despite gerrymandering.

However, as the voters of the minority party concentrate, this disparity reduces. The majority party sees a reduction in its gerrymandering power as it can no longer avoid forming districts where the minority party wins due to its concentrated voters. Similarly, the minority party finds it easy

to gerrymander to win a larger number of districts. At the extreme, with $\phi = 10$, it is able to win almost half the districts despite being at a 20% vote share disadvantage.
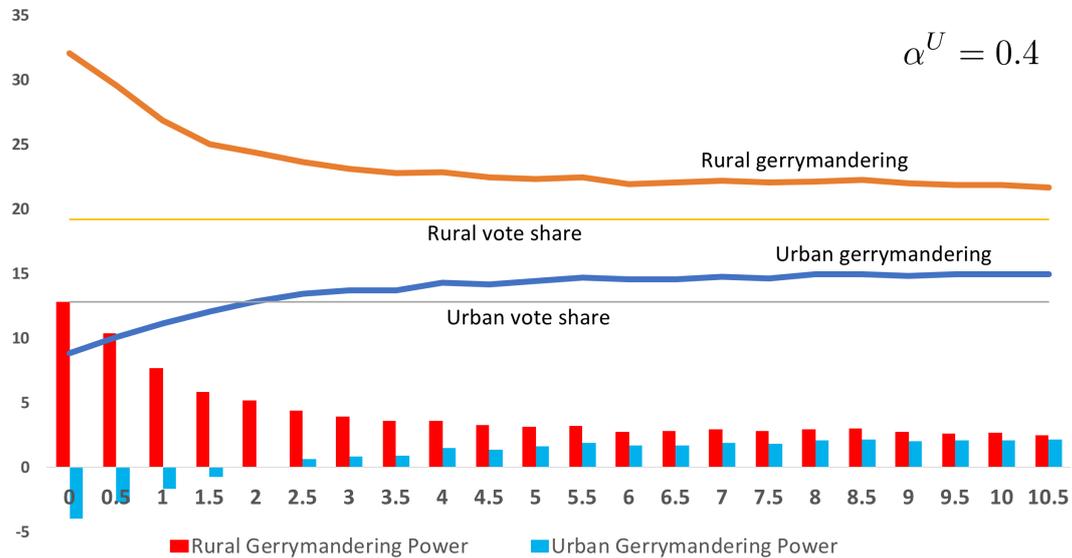

Figure 4: The number of districts won (top) and gerrymandering power (bottom) compared to ϕ for $\alpha^U = 0.4$.

6.2 Close Elections

Arguably, the more interesting elections in practice are the close elections with vote share difference of less than 20%. The trend is very different in these elections. For instance, consider Figure 5 and Figure 6 with $\alpha^U = 0.5$.

At $\phi = 0$, the voters are spread uniformly at random, which makes it easy for the gerrymandering party to put precincts with a slight majority together with precincts with a slight minority to form many districts with a slight majority, leading to a high gerrymandering power. Further, this holds for each party due to symmetry.

As the divide strengthens, the rural party witnesses a diminishing gerrymandering power as in the case of unbalanced elections. However, an interesting pattern emerges in the gerrymandering power of the urban party. As ϕ increases, we see that the gerrymandering power decreases suddenly till $\phi = 2$, then increases slowly, and finally levels out, forming a *trough*.

We do not believe this trough to be an artifact of our algorithm B^+. On a smaller number of instances, we ran the iterative algorithm I for several hours to come up with hundreds of thousands of districting plans, and chose the most gerrymandered of them. A similar pattern emerges, though this approach returns less gerrymandered solutions than B^+, making the pattern a bit less emphasized.

We hypothesize that with a moderate ϕ, there are still many urban voters in deeply rural regions, which constitutes a lot of wasted votes for the urban party as it is unable to put them together with other urban voters and form a district it can win. However, as the concentration further increases, these voters are brought closer to the urban centre, allowing the party to utilize their votes to win a few additional districts.

Due to a similar reason, when the rural party has a minority vote share (say $\alpha^R = 0.45$), we see an inverse pattern with its gerrymandering power initially increasing, and then slightly decreasing, thus forming a *peak*. Again, this is because with a moderate ϕ, the urban party has a lot of wasted votes within rural regions, which helps the rural party gerrymander well.
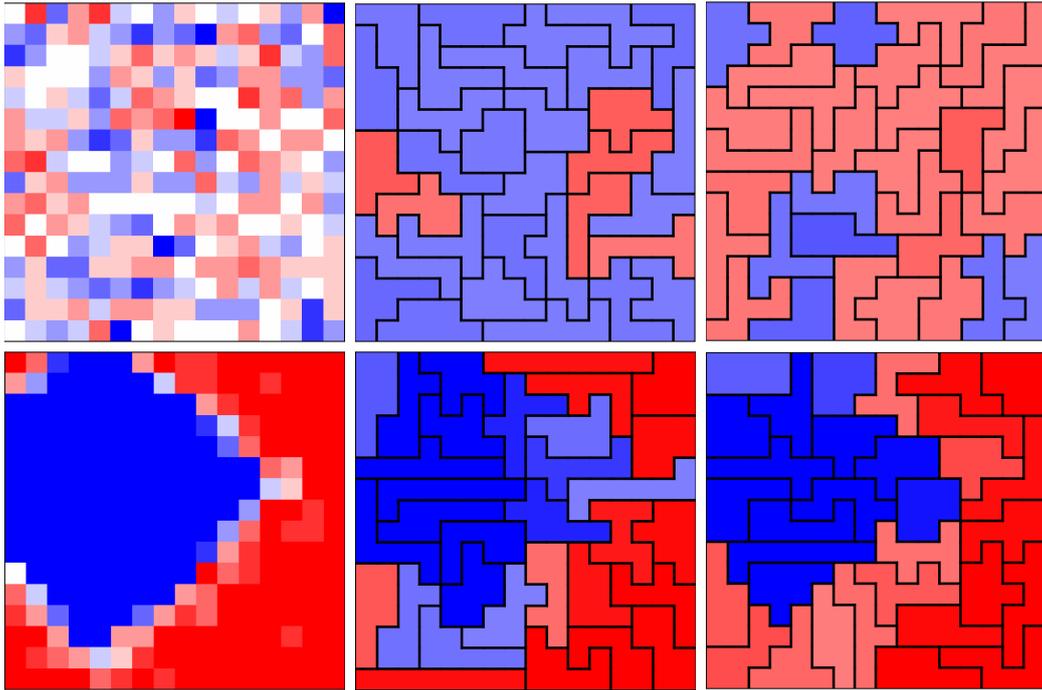
Figure 5: An example of the gerrymandered solutions B^+ found when $\alpha^U = 0.5$. Top row $\phi = 0$ bottom row $\phi = 10$. First column is partisan distribution of voters (similar to Figure 1). Second column is the gerrymandered solution for U, third column is the gerrymandered solution for R. Colour intensity within a district increases with the strength of victory.

6.3 Concentration Leads to Fairer Districting

Interestingly, Figure 3 shows that across all vote shares, the gerrymandering power of both parties converges to $1/16$ (i.e., 2 more districts compared to the vote share with $K = 32$) as ϕ goes to 10. In fact, the convergence seems to begin at a fairly low concentration level (around $\phi = 2.5$). That is, at an extreme level of concentration, both parties are able to gerrymander and win about two more districts than their proportional share (dictated by their share of the votes).

Intuitively, at extreme concentration levels, there is a densely packed region of urban voters near the urban centre, a densely packed region of rural voters surrounding it, and a much sharper boundary in between (Figure 1). Irrespective of which party gerrymanders, districts near the urban centre are won by the urban party, and districts densely packed with rural voters are won by the rural party. This ensures each party approximately its proportional share of the districts. The only control that the gerrymandering party has is near the boundary, where it can merge its own voters with voters of the opponent, creating districts with a slight majority. This is reflected in the urban-gerrymandered and rural-gerrymandered districting shown in the final columns of Figure 2 and Figure 5. Since both parties control the same boundary region when gerrymandering, their gerrymandering power becomes identical with such extreme concentration. Further, since the number of vertices on the boundary is a fraction of the total number of vertices, this gerrymandering power is relatively small.

6.4 A Rural Advantage

Finally, we observe that the gerrymandering power is not symmetric between the urban and rural parties. The rural party almost always has a higher gerrymandering power than the urban party, even
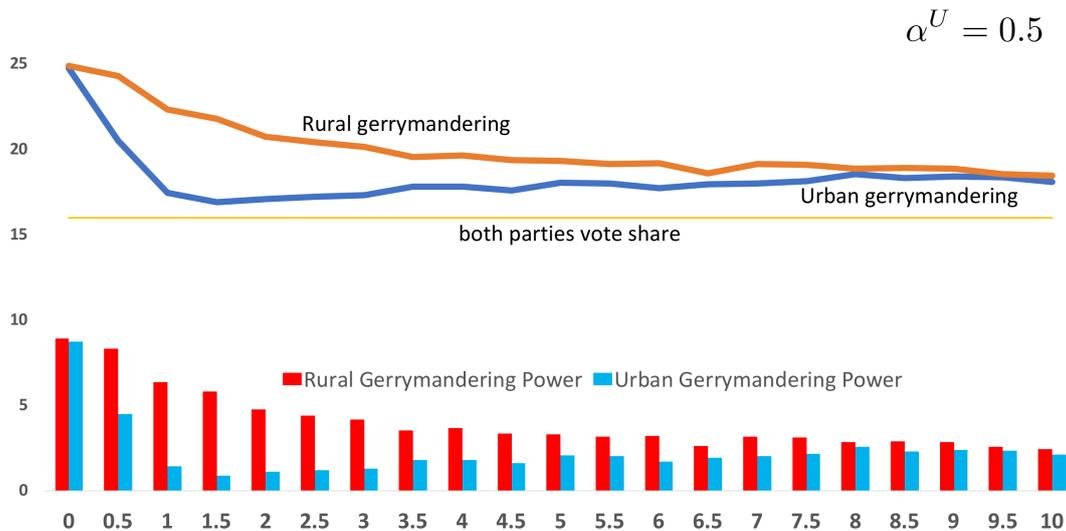
Figure 6: The number of districts won (top) and gerrymandering power (bottom) compared to ϕ for $\alpha^U = 0.5$.

in the case of proportional vote split (Figure 6). This asymmetry is not surprising. As Figure 1 shows, the distribution of voters is also not symmetric; the urban party's voters congregate together in a tight area, while the rural party's voters surround them on all sides.

7   Discussion

Our definition of the gerrymandering power of a political party, the theoretical model for a worst-case analysis, and our empirical observations raise many interesting open questions.

On the theoretical level, our results consider extreme cases of an infinite graph or just two districts. Solving the case of a finite graph with more than two districts is an immediate open question. We also note that while we use convex districts in Theorem 1, our proof of Theorem 2 uses districts that are far from convex. An interesting direction is to incorporate a formal requirement of district compactness into our framework.

Extending our techniques to large real-world planar graphs is clearly the most interesting future direction. For instance, in Theorem 2, how does the vote share needed to win both districts change in non-grid graphs? Insights from our simulations lead us to believe that in close elections on real graphs, the gerrymandering power of both parties will eventually decrease with voter concentration, simply because extreme concentration limits the gerrymandering possibilities to a sharper boundary between voters of the two parties. Note that while our model does not explicitly postulate suburban areas, our party affiliation dispersion model does provide an implicit sense of the more mixed affiliation nature of suburban areas.

We believe that our model is just a starting point to developing a more precise understanding of the gerrymandering power so as to address gerrymandering in the real world.

8   Acknowledgments

References

[1] Micah Altman. The computational complexity of automated redistricting: Is automation the answer. *Rutgers Computer & Technology Law Journal*, 23:81–142, 1997.

[2] N. Apollonio, R.I. Becker, I. Lari, F. Ricca, and B. Simeone. Bicolored graph partitioning, or: gerrymandering at its worst. *Discrete Applied Mathematics*, 157:3601–3614, 2009.

[3] Yoram Bachrach, Omer Lev, Yoad Lewenberg, and Yair Zick. Misrepresentation in district voting. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 81–87, New York City, New York, July 2016.

[4] Bill Bishop. *The Big Sort: Why the Clustering of Like-Minded America is Tearing Us Apart.* Mariner Books, 2009.

[5] D. Butler. The electoral process the redrawing of parliamentary boundaries in britain. *British Elections and Parties Yearbook*, 2(1), 1992.

[6] Jowei Chen and David Cottrell. Evaluating partisan gains from congressional gerrymandering: Using computer simulations to estimate the effect of gerrymandering in the u.s. house. *Electoral Studies*, 44:329–340, 2016.

[7] Jowei Chen and Jonathan Rodden. Tobler's law, urbanization, and electoral bias: Why compact, contiguous districts are bad for the democrats, 2009.

[8] Jowei Chen and Jonathan Rodden. Unintentional gerrymandering: Political geography and electoral bias in legislatures. *Quarterly Journal of Political Science*, 8:239–269, 2013.

[9] M.E. Dyer and A.M. Frieze. On the complexity of partitioning graphs into connected subgraphs. *Discrete Applied Mathematics*, 10:139–153, 1985.

[10] E. J. Engstrom. Stacking the states, stacking the house: The partisan consequences of congressional redistricting in 19th century. *APSR*, 100:419–427, 2006.

[11] Harry Enten. Ending gerrymandering won't fix what ails america. *FiveThirtyEight*, 26 January 2018.

[12] Robert S. Erikson. Malapportionment, gerrymandering, and party fortunes in congressional elections. *The American Political Science Review*, 66(4):1234–1245, 1972. ISSN 00030554, 15375943. URL http://www.jstor.org/stable/1957176.

[13] Dan S. Felsenthal and Nicholas R. Miller. What to do about election inversions under proportional representation? *Representation*, 51(2):173–186, 2015.

[14] Benjamin Fifield, Michael Higgins, Kosuke Imai, and Alexander Tarr. A new automated redistricting simulator using markov chain monte carlo, January 2018.

[15] John N. Friedman and Richard T. Holden. The rising incumbent reelection rate: What's gerrymandering got to do with it? *The Journal of Politics*, 71:593–611, April 2009. ISSN 1468-2508. doi: 10.1017/S0022381609090483.

[16] B. Grofman, W. Koetzle, and T. Brunell. An integrated perspective on the three potential sources of partisan bias: Malapportionment, turnout differences, & the geographic distribution of party vote shares. *Electoral Studies*, 16(4):457–470, 1997.

[17] Bernard Grofman and Gary King. The future of partisan symmetry as a judicial test for partisan gerrymandering after lulac v. perry. *Election Law Journal*, 6(1):2–35, 2007.

[18] S. Issacharoff. Gerrymandering & political cartels. *Harvard Law Review*, 116(2):593–648, 2002.

[19] Donald Kagan. The origin and purposes of ostracism. *Hesperia: The Journal of the American School of Classical Studies at Athens*, 30(4):393–401, October-December 1961.

[20] R N Karasev. Equipartition of several measures. arXiv:1011.4762, 2010.

[21] Brian Klaas. Gerrymandering is the biggest obstacle to genuine democracy in the united states. so why is no one protesting? *Washington Post*, 10 February 2017.

[22] Yoad Lewenberg, Omer Lev, and Jeffrey S. Rosenschein. Divide and conquer: Using geographic manipulation to win district-based elections. In *Proceedings of the 16th International Coference on Autonomous Agents and Multiagent Systems (AAMAS)*, pages 624–632, São-Paulo, Brazil, May 2017.

[23] D. Lublin. *The Paradox of Representation: Racial Gerrymandering and Minority Interests in Congress*. Princeton University Press, 1997.

[24] New York Times Editorial Board. A shot at fixing american politics. *New York Times*, page A22, 30 September 2017.

[25] H. Nurmi. *Voting Paradoxes and How to Deal with Them*. Springer-Verlag, 1999.

[26] Wesley Pegden, Ariel D. Procaccia, and Dingli Yu. A partisan districting protocol with provably nonpartisan outcomes. ArXiv:1710.08781, October 2017.

[27] Clemens Puppe and Attila Tasnádi. A computational approach to unbiased districting. *Mathematical and Computer Modelling*, 48(9–10):1455–1460, 2008. ISSN 0895-7177. doi: http://dx.doi.org/10.1016/j.mcm.2008.05.024. URL http://www.sciencedirect.com/science/article/pii/S0895717708001908. Mathematical Modeling of Voting Systems and Elections: Theory and Applications.

[28] Raza Rumi. 2016 election explainer in 4 words: The urban-rural divide. *HuffPost*, 6 December 2017.

[29] P.H. Schuck. The thickest thicket: Partisan gerrymandering and judicial regulation of politics. *Columbia Law Review*, 87(7):1325–1384, 1987.

[30] Pablo Soberón. Balanced convex partitions of measures in rd. *Mathematika*, 58(1):71–76, 2012.

[31] E.S. Staveley. *Greek and Roman voting and elections*. Thames & Hudson, 1972.

[32] Andranik Tangian. Computational application of the mathematical theory of democracy to arrow's impossibility theorem (how dictatorial are arrow's dictators?). *Social Choice and Welfare*, 35(1):129–161, June 2010.

[33] René van Bevern, Robert Bredereck, Jiehua Chen, Vincent Froese, Rolf Niedermeier, and Gerhard J. Woeginger. Network-based vertex dissolution. *SIAM Journal on Discrete Mathematics*, 29(2):888–914, 2015.

[34] Sam Wang. The great gerrymander of 2012. *New York Times*, page SR1, February 3 2013.

[35] Samuel S.-H. Wang. Three tests for practical evaluation of partisan gerrymandering. *Stanford Law Review*, 68:1263–1321, June 2016.

[36] David Wasserman. Hating gerrymandering is easy. fixing it is harder. *FiveThirtyEight*, 25 January 2018.

[37] Jed Yang. Some np-complete edge packing and partitioning problems in planar graphs. ArXiv:1409.2426, September 2014.

Allan Borodin, Nisarg Shah and Tyrone Strangway
University of Toronto
Toronto, Canada
Email: {bor,nisarg,tyrone}@cs.toronto.edu

Omer Lev
Ben-Gurion University
Beersheba, Israel
Email: omerlev@bgu.ac.il