An Eigenvalue-Based Algorithm for Graph Bisection

Paper by Ravi B. Boppana

Joline Cheng and Ray Luo

University of Waterloo

March 2025

Introduction

- **bisection**: a partition of the vertices into two even-sized components
- **bisection size**: number of edges between the two components

Graph Bisection Problem (GBP)

For any graph, what is its minimum bisection size? (**bisection width**)

- GBP is **NP-Complete**
  - Approximations (ex. within constant factor) or bounds
  - Average case, instead of worst case
- Our paper: finds a lower bound on the bisection width with high probability that it is the true bisection width.

Applications

- Very-large-scale integration (VLSI)
  - How to best fit millions of transistors on one chip?
- Scientific Computing
- Sparse matrix computation
- Physics
  - Spin Glass Problem (Statistical Physics/Magnetism)
- Many "divide and conquer" problems

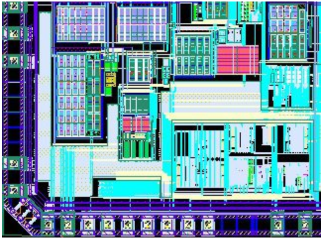


Figure: VLSI Chip

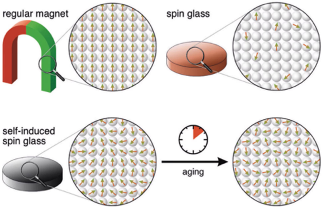


Figure: Spin Glass

Outline

Set Up Function

Let $V(G) = [n]$. Given bisection (V_1, V_2), define its **associated vector** to be the vector where the i-th coordinate is 1 if $i \in V_1$ and -1 if $i \in V_2$.

For input vectors $d, x \in \mathbb{R}^n$, define

$$f(G, d, x) = \sum_{ij \in E} \frac{1 - x_i x_j}{2} + \sum_{i \in V} d_i(x_i^2 - 1) \tag{1}$$

Lemma

For any bisection (V_1, V_2) with associated vector x and any vector $d \in \mathbb{R}^n$, the bisection size is $f(G, d, x)$.

Proof Idea: By inspection; $\frac{1 - x_i x_j}{2}$ is 1 if $i \sim j$ and i, j are in different components (and 0 otherwise). The second sum is 0, since $x_i^2 = 1$.

Intuition

First, let us develop some intuition for our algorithm:

- Define a concave function $g(G, d) = \min_{x \in S \ \|x\|=\sqrt{n}} f(G, d, x)$
- Set a maximum $h(G) = \max_{d \in \mathbb{R}^n} g(G, d)$
- Show $h(G)$ can be calculated to arbitrary precision
- Prove every graph $G = (V, E)$ has bisection width at least $h(G)$

Definitions

Let $S := \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 0\}$ and $\|x\| = (\sum_{i=1}^n x_i^2)^{1/2}$. Then, set:

$$g(G,d) = \min_{x \in S \ \|x\| = \sqrt{n}} f(G,d,x) \tag{2}$$

Let A be the adjacency matrix of G.
For a vector d, let $D = diag(d)$ be the matrix whose diagonal entries are entries of d:

$$\begin{pmatrix} d_1 & 0 \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix}$$

Let $B = A + D$, and let $sum(B)$ be sum of all n^2 entries of B.

Discrete -> Continuous

Let $B_S : S \to S$, mapping x to the projection of Bx onto S (the smallest point on S from Bx).
We can derive the following equation from the definition of $f(G, d, x)$:

$$g(G, d) = \frac{sum(B) - n\lambda(B_S)}{4} \tag{3}$$

Where $\lambda(B_S)$ is the largest eigenvalue of B_S.
Notice that $g(G, d)$ is a concave function of d. Let us define its maximum:

$$h(G) = \max_{d \in \mathbb{R}^n} g(G, d) \tag{4}$$

Grotschel, Lovasz and Schrijver proved, under very general conditions, the maximum of a concave function can be calculated to arbitrary precision in polynomial time. So, it remains to show that $h(G)$ is a lower bound for the bisection width.

Lower Bound

Theorem

Every graph $G = (V, E)$ has bisection width at least $h(G)$,

Proof: Choose an arbitrary vector d. Suppose that (V_1, V_2) is the minimum-size bisection with the associated vector x, which we know from the previous lemma has bisection width $= f(G, d, x)$. Since $x \in S$,

$$f(G, d, x) \geq \min_{x \in S} f(G, d, x) = g(G, d)$$

Since d was chosen arbitrarily, we have:

$$f(G, d, x) \geq h(G)$$

So $G = (V, E)$ has bisection width at least $h(G)$.

Upper Bound

There is also an algorithm for the upper bound of the bisection width using a similar method.

- Given a graph G with adjacency matrix A, find the vector d that minimizes $g(G, d)$ using Grotschel,Lovasz and Schrijver's method.
- Let y be the eigenvector for the largest eigenvalue of $(A + D)_S$, where $D = diag(d)$.
- Output the bisection with the $n/2$ largest components of y on one side, and the $n/2$ smallest components on the other side.

In the next section, we will prove that that our algorithm works well on average.

Statement of Main Theorem

Let $\mathcal{G}(n, m, b) =$ graphs with n vertices, m edges, and bisection width b.

Theorem

Let $G \in \mathcal{G}(n, m, b)$ be chosen randomly, with uniform distribution, and that

$$b \leq \frac{1}{2}m - \frac{5}{2}\sqrt{mn\log n}$$

Then with probability $1 - O(\frac{1}{n})$, we have

$$b = h(G) = \text{our lower bound}$$

Part 1: Reduction

Let (V_1, V_2) be a minimum bisection. Denote $N_1(i), N_2(i)$ to be neighbours of i in V_1, V_2 resp. Define $d \in \mathbb{R}^n$ to be

$$d_i = \begin{cases} |N_2(i)| - |N_1(i)| & i \in V_1 \\ |N_1(i)| - |N_2(i)| & i \in V_2 \end{cases}$$

This is the d we want to use for our D in our algorithm (where $B = A + D$).
We can get the series of implications:

$$\lambda(B_S) = 0 \iff g(G, d) = b \implies h(G) = b$$

So we just need to prove $\lambda(B_S) =$ largest eigenvalue of $B_S = 0$.

The associated vector y is an eigenvector of B_S with eigenvalue 0!
Need to show: all other eigenvalues ≤ 0 with high probability

Part 2: Prove $\lambda(B_S) = 0$ (with High Probability)

Set $p = \frac{m-b}{\frac{n}{2}\left(\frac{n}{2}-1\right)}$, $q = \frac{b}{\left(\frac{n}{2}\right)^2}$ and $M \in \mathbb{R}_{n\times n}$ be the matrix where $M_{ij} = q$ if i, j are in opposite components and p otherwise.

Boppana's Proof Sketch

1. Show $E(B) = M - \frac{1}{2}(p-q)n\mathbb{I}$
2. Note that B_S and $(B-M)_S$ are equivalent on vectors orthogonal to y
3. Suffices to show eigenvectors of $(B-M)_S$ are nonpositive*
4. Suffices to show all eigenvalues of $B - E(B)$ have an upper bound of $\frac{1}{2}(p-q)n$*

(*) = with high probability

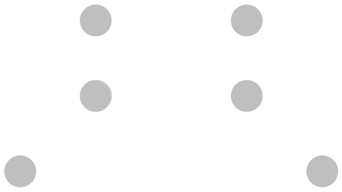
Let μ(matrix) represent any eigenvalue of the input matrix.

$$\mu(B - E(B)) \leq^* \frac{1}{2}(p-q)n \implies \mu((B-M)_S) \leq^* 0 \implies \mu(B_S) \leq^* 0$$

Model of random graphs

To show $\mu(B - E(B)) \leq \frac{1}{2}(p-q)n$, we need to replace the model $\mathcal{G}(n, m, b)$, from which we get our random graph G, with a closely related model. The new model forms a random graph as follows:

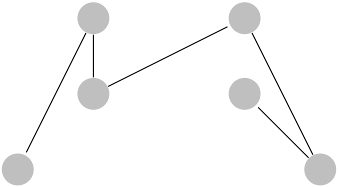
1. Partition the vertices V into 2 equal parts (V_1, V_2)



Model of random graphs

To show $\mu(B - E(B)) \leq \frac{1}{2}(p-q)n$, we need to replace the model $\mathcal{G}(n, m, b)$, from which we get our random graph G, with a closely related model. The new model forms a random graph as follows:

1. Partition the vertices V into 2 equal parts (V_1, V_2)
2. Each edge within either of the 2 parts is chosen independently with probability p
3. Each edge between the 2 parts is chosen independently with a smaller probability q



Model of random graphs

To show $\mu(B - E(B)) \leq \frac{1}{2}(p-q)n$, we need to replace the model $\mathcal{G}(n, m, b)$, from which we get our random graph G, with a closely related model. The new model forms a random graph as follows:

1. Partition the vertices V into 2 equal parts (V_1, V_2)
2. Each edge within either of the 2 parts is chosen independently with probability p
3. Each edge between the 2 parts is chosen independently with a smaller probability q

Such a random graph is likely have about m edges and bisection width b, so it will be similar to a graph from $\mathcal{G}(n, m, b)$. The transfer can be more carefully justified, but assume G was generated this way for now.

Part 3: Prove $\mu(B - E(B)) \leq \frac{1}{2}(p-q)n \ \forall \mu$

We can estimate the eigenvalues of $B - E(B)$ by decomposing it into $A + D$, then using the inequality:

$$\lambda(B - E(B)) \leq \lambda(A - E(A)) + \lambda(D - E(D))$$

The eigenvalues of $D - E(D)$ are its diagonal entries, which can be shown to be small with high probability using Chernoff's bound on the tail of a binomial distribution. We get:

$$\lambda(D - E(D)) \leq 5\sqrt{pn\log n}$$

with high probability.
The eigenvalues of $A - E(A)$ can be estimated with the following theorem, an extension of a result due to Furedi and Komlos.

Part 3: Prove $\mu(B - E(B)) \leq \frac{1}{2}(p-q)n \; \forall\mu$

Theorem

Let Z be a random $n \times n$ symmetric matrix whose entries are independent up to symmetry, have mean zero, have standard deviation at most σ, and are concentrated on $[-a, a]$. Suppose that $\sigma\sqrt{n} \geq 10a\sqrt{\log n}$. Then with probability at least $1 - O(1/n)$, all the eigenvalues of Z have absolute value less than $3\sigma\sqrt{n}$.

Applying this theorem to $Z = A - E(A)$, it follows that $\lambda(A - E(A)) \leq 3\sqrt{pn}$ with high probability. So with high probability,

$$\lambda(B - E(B)) \leq 6\sqrt{pn \log n}$$

By the hypothesis on the size of b, $6\sqrt{pn\log n} < \frac{1}{2}(p-q)n$. So, we get $\mu(B - E(B)) < \frac{1}{2}(p-q)n \; \forall\mu$, which is sufficient to prove the theorem.

Regular Graphs

Let $\mathcal{G}(n, r, b)$ = graphs with n vertices, bisection width b, and are r-regular.

Theorem

Let $G \in \mathcal{G}(n, r, b)$ where $r \geq 3$ and $b = o(n^{1-\frac{1}{\lfloor \frac{r+1}{2} \rfloor}})$. Then with probability $1 - o(1)$, we have $b = h(G)$.

Better than on $\mathcal{G}(n, m, b)$, which was probability $1 - O(\frac{1}{n})$

- Previously: showed that $\mu(B - E(B)) \leq \frac{1}{2}(p - q)n$ (for all graphs)
- New: We can further show
  - $\mu(D - E(D)) \leq \frac{1}{2}(p - q)n - 1$
  - $\mu(A - E(A)) \leq 3d^{\frac{3}{4}}$

$\implies \mu(B - E(B)) \leq \frac{1}{2}(p - q)n - \frac{1}{2}$

Previous Works

- Donath and Hoffman first developed a lower bound for the bisection width problem, based on computing the first 2 eigenvalues of a matrix associated with the graph. The bound in this paper is always greater or equal than the previous bound, and has an average-case analysis.
- Bui et al. developed a heuristic for the bisection of the class of all r-regular graph. It can be shown that this algorithm achieves the same performance over regular graphs.
- Dyer and Frieze presented an algorithm that finds, with high probability, provably optimal bisections over the class of dense graphs.

Other Results

Another result, presented in *Eigenvalues in Combinatorial Optimization* by **Bojan Mohar** and **Svatopluk Poljak**, found another formulation for $g(G, d)$. By noticing that, for some $d \in S = \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i = 0\}$,

$$f(G, d, x) = \frac{1}{4} \sum_{i,j \in V} (x_i - x_j)^2 + \sum_{i=1}^{n} d_i = \frac{1}{4} x^T (L(G) + D) x$$

Where $L(G)$ is the Laplacian matrix of G.

$$g(G, d) = \min_{y \in S, \|y\|=1} \frac{n}{4} y^T (L(G) + D) y \tag{5}$$

Where $D = diag(d)$ is defined in the same way as before. This leads to a easier probabilistic analysis, by using Rayleigh quotient and Fan's Theorem.