# Are Stable Instances Easy?

Yonatan Bilu
Mobileye Inc.

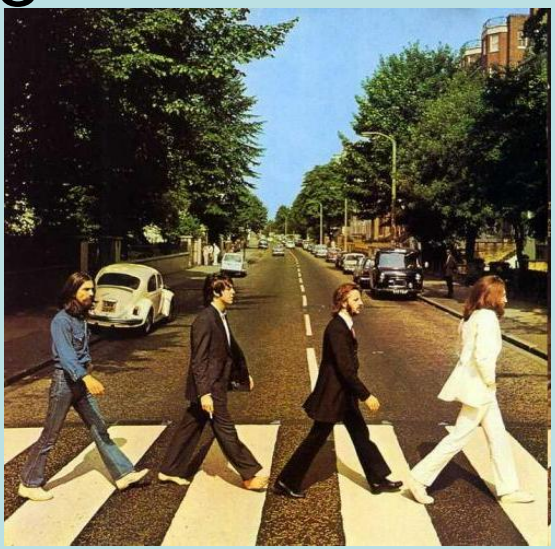
Nati Linial
Hebrew University

ICS 2010

# Clustering Problems

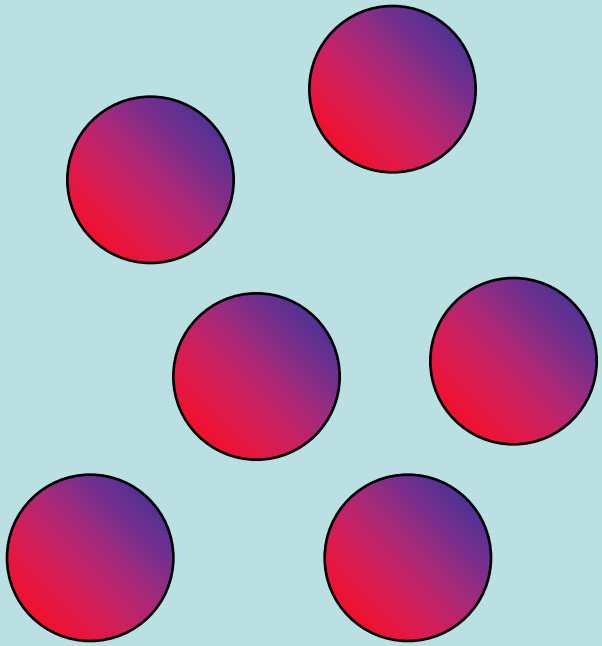
Abundant in science:

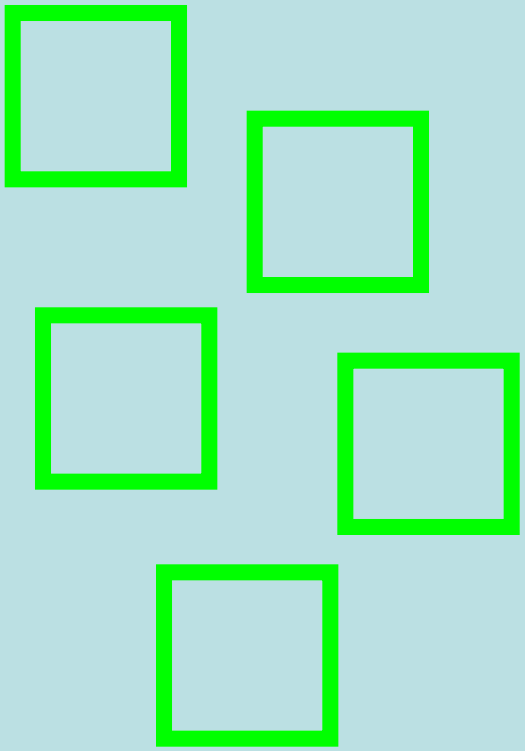
- Clustering proteins by edit distance
- Clustering genes by expression profiles
- Clustering pixels in an image into foreground and background



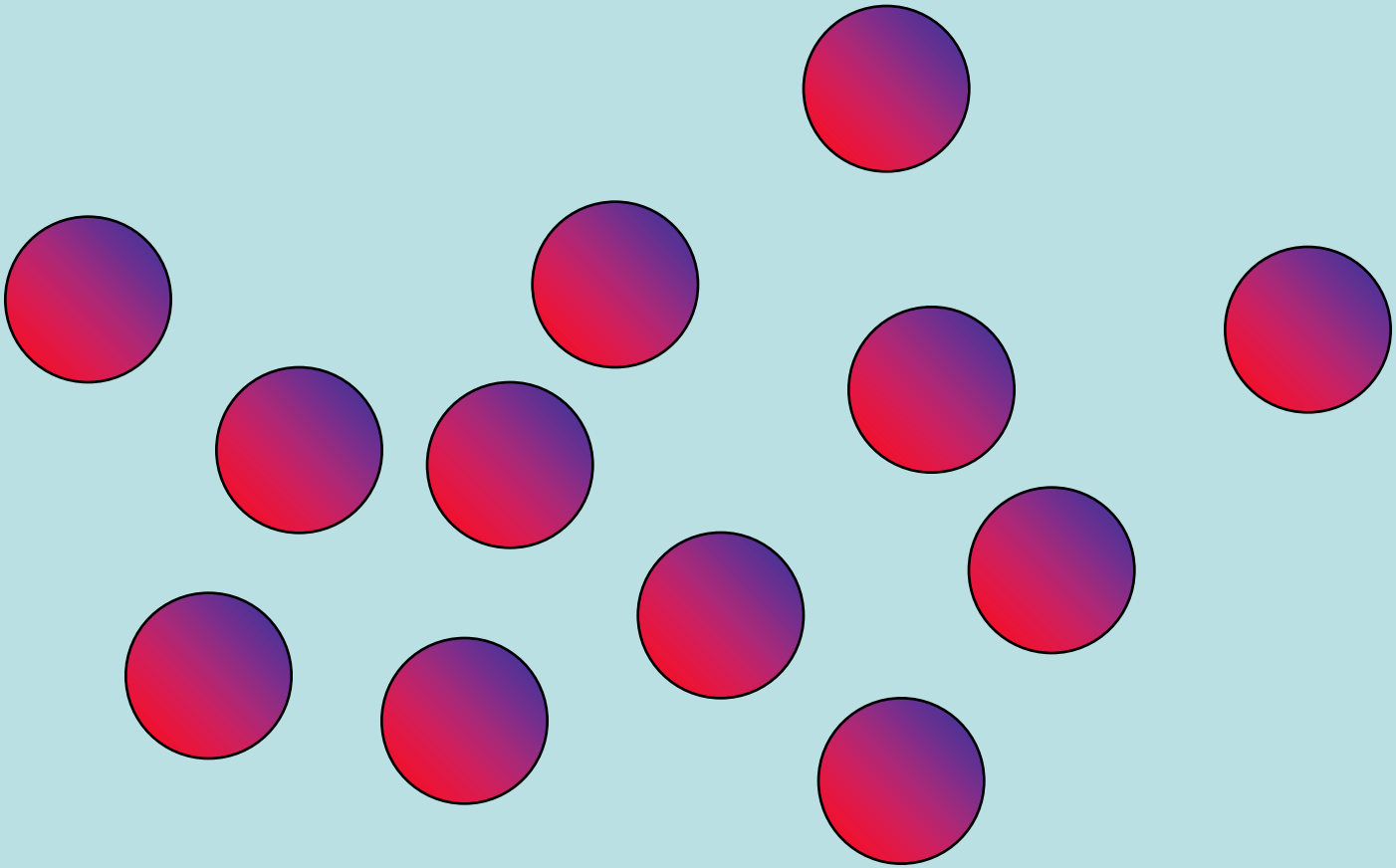
Like all good things: NP-hard

# What is the solution?





# What is the solution?



# Clustering Problems

**Are they really hard when we actually care?**

# Clustering Problems

**Which instances are interesting?**

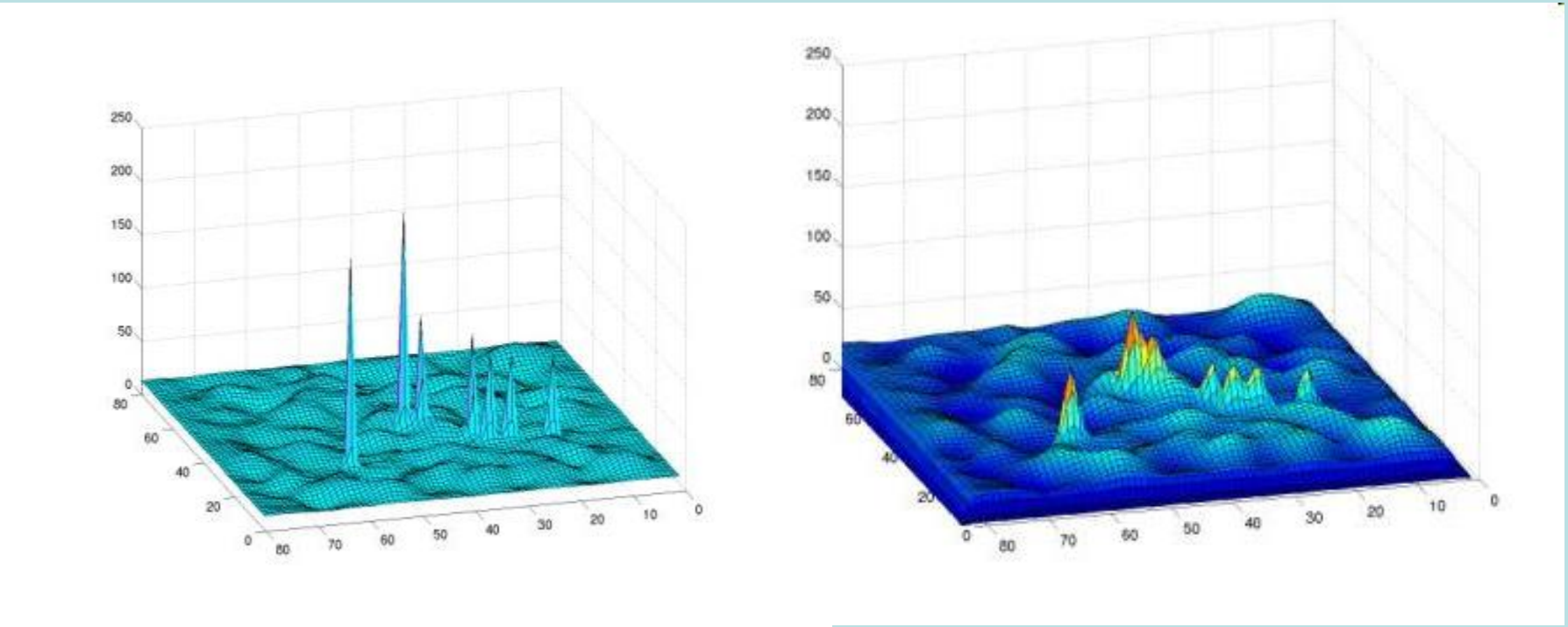
Are interesting instances easy to solve?

Are interesting instances easy to approximate?

Are interesting instances amenable to practical heuristics?

# Related work

**Smoothed Analysis** (Spielman & Teng 2001): Don’t say it’s hard if the hard instances are “singular” points.



# Related work

**THE (c, ε)-PROPERTY** (Balcan, Blum and Gupta 2008):

All c-approximations of the optimal solution are ε-close to it.

Then an O(ε)-close clustering is easily found (*K-means*, *k-medians* and *min-sum**).

# Related work

**Clusterability** (Ostrovsky, Rabani, Schulman and Swamy 2006):

Optimal k-clustering is much better than (k-1)-clustering.

Then there's a PTAS.

**More Clusterability** (Ackerman & Ben David)

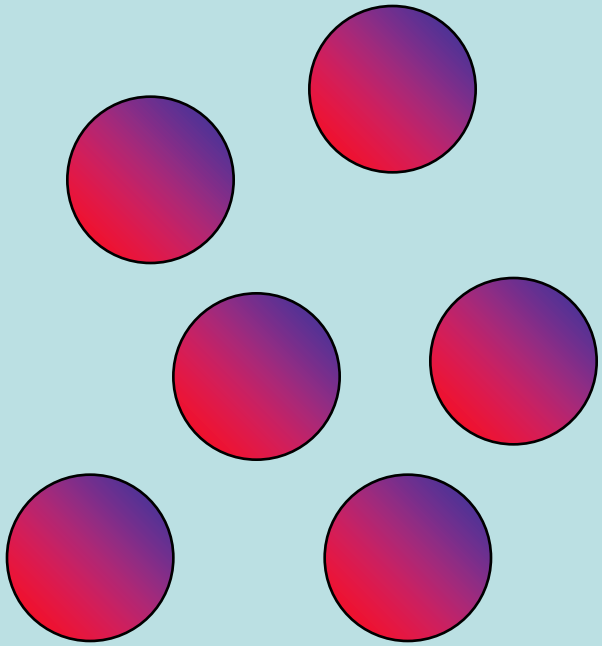
# Stability

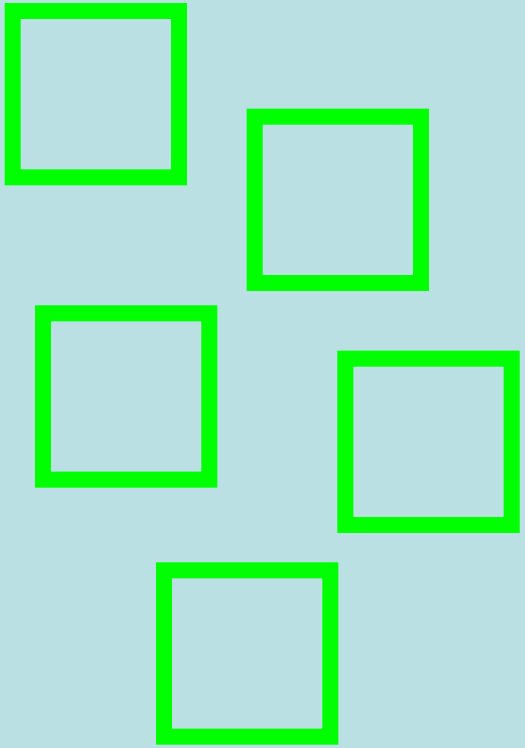
We don’t care about small perturbations (we never have exact measurements anyway).

Small perturbations should not change ***the structure*** of the optimal solution.

# Stability

Small perturbations do not change the structure of the optimal solution





# Stability

**γ-perturbation**: multiply inputs by numbers between 1 and γ.

**γ-stability**: γ-perturbations do not change the structure of the optimal solution.

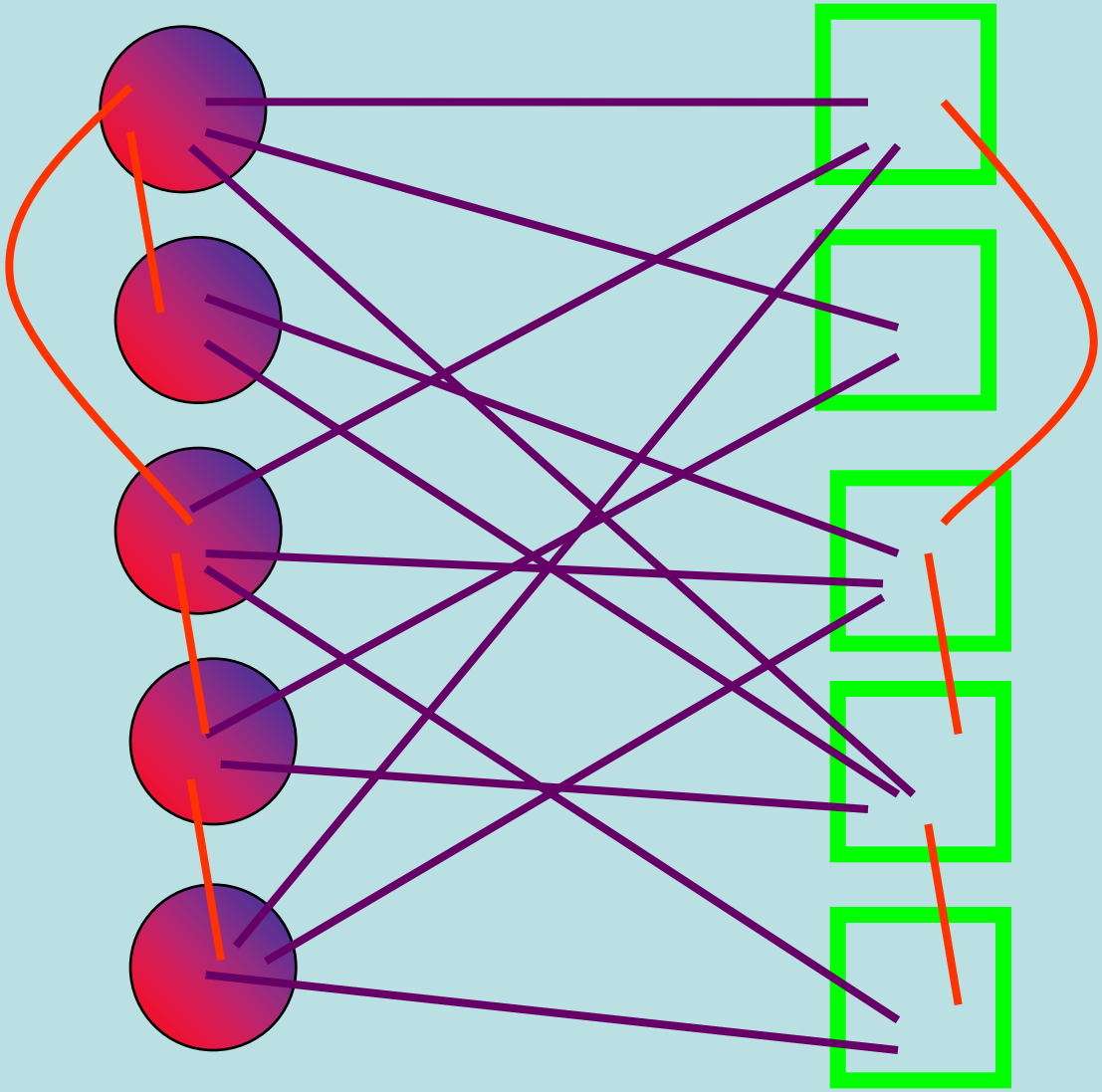
**Are Stable Instances Easy?**

# Max Cut

Maximize weight of crossing edges

Minimize weight of internal edges

Minimize weight of internal edges - crossing edges



# Max Cut

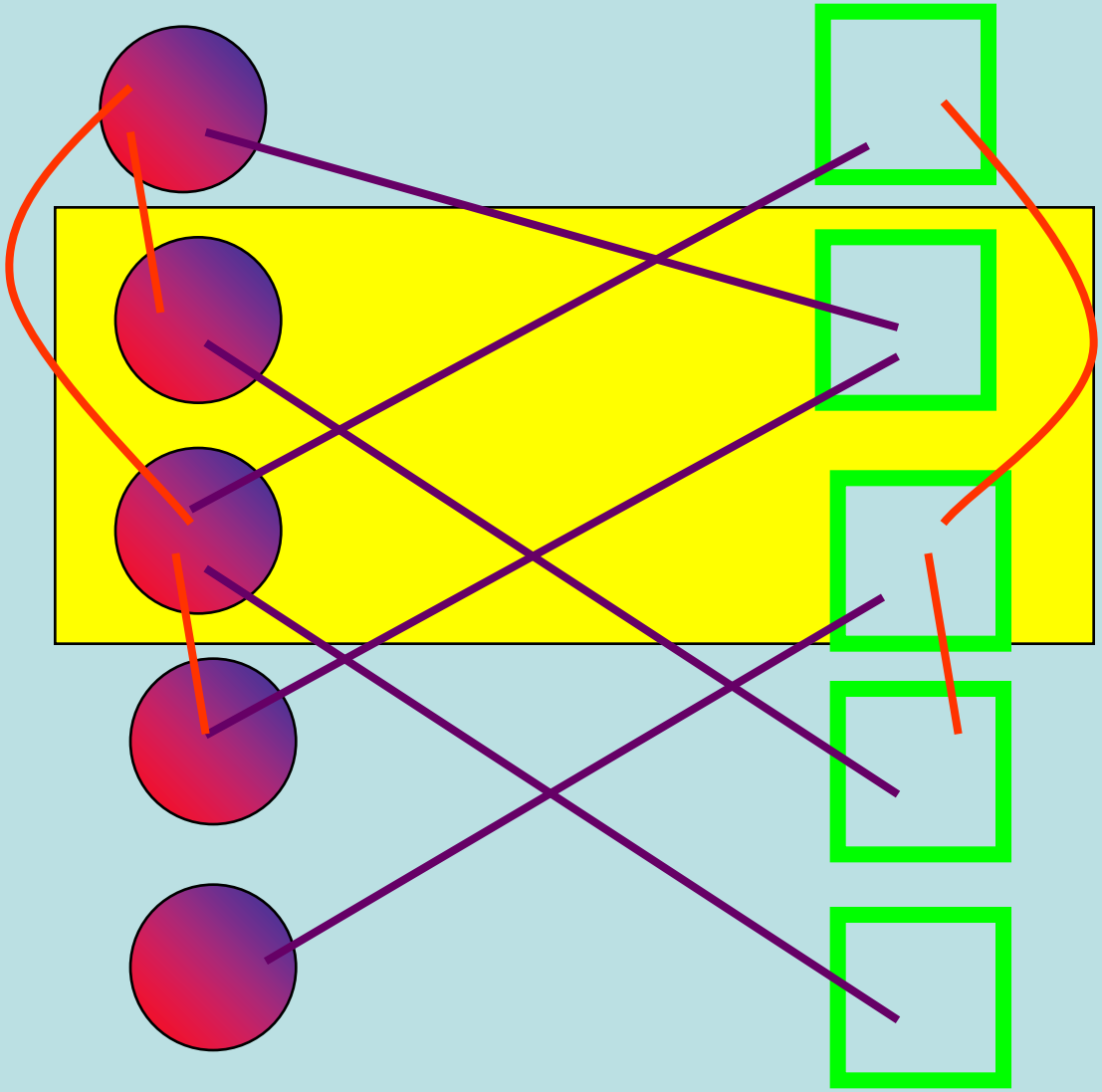
**For every subset -**

**Weight of crossing ≥ Weight of internal**

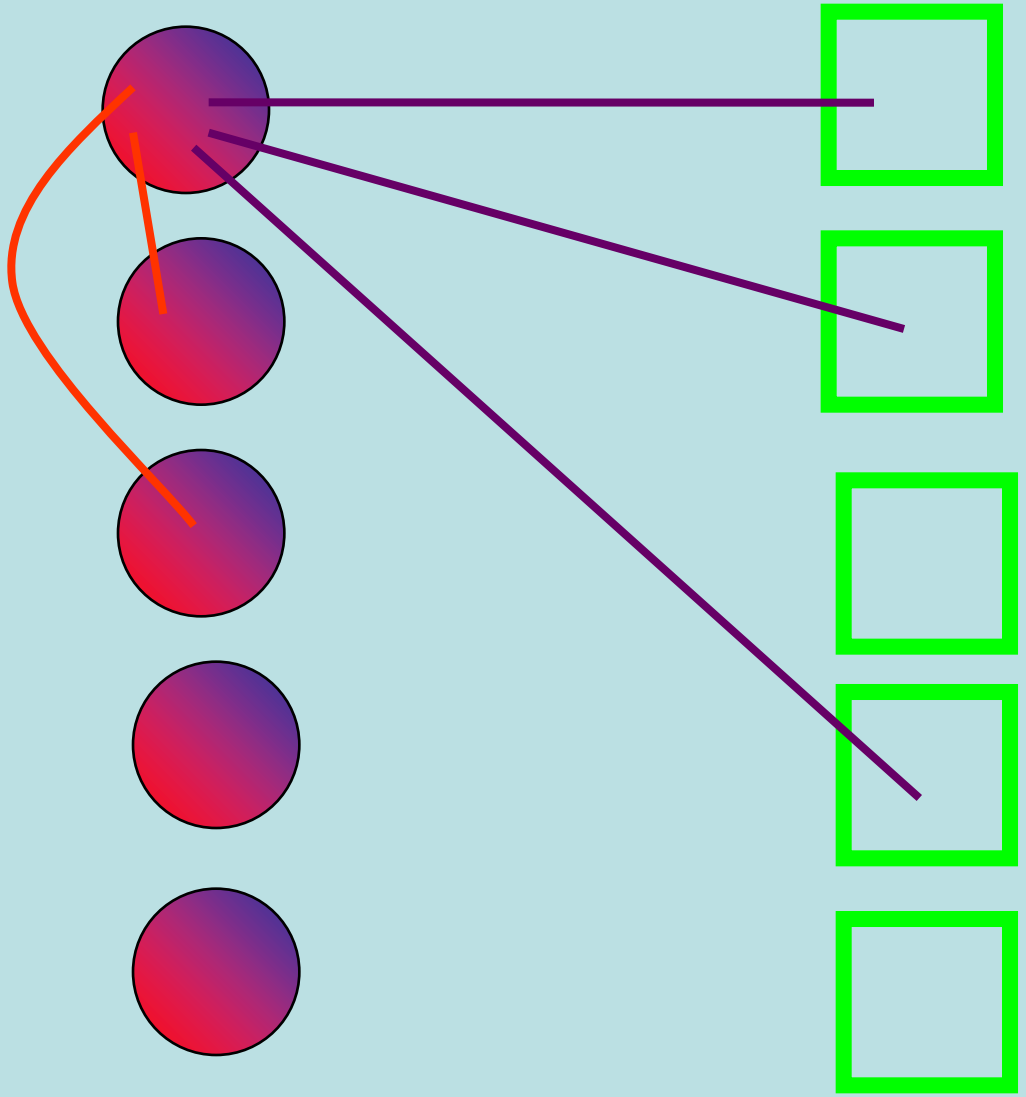
**γ-Stability:**

**For every subset -**

**Weight of crossing ≥ γ x (Weight of internal)**



# Max Cut



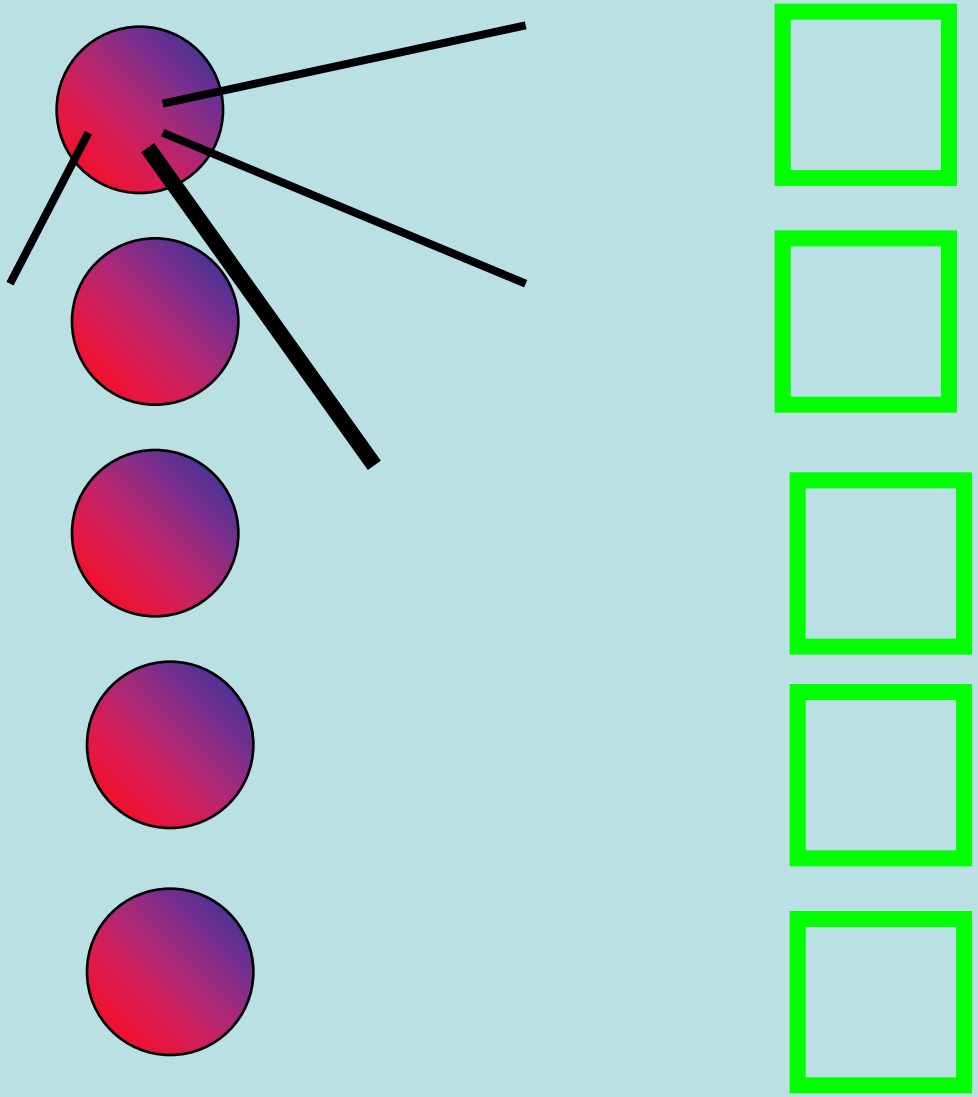
**γ-Local Stability:**

**For every vertex -**

**Weight of crossing ≥**
**γ** x (**Weight of internal**)

# Combinatorial algorithm

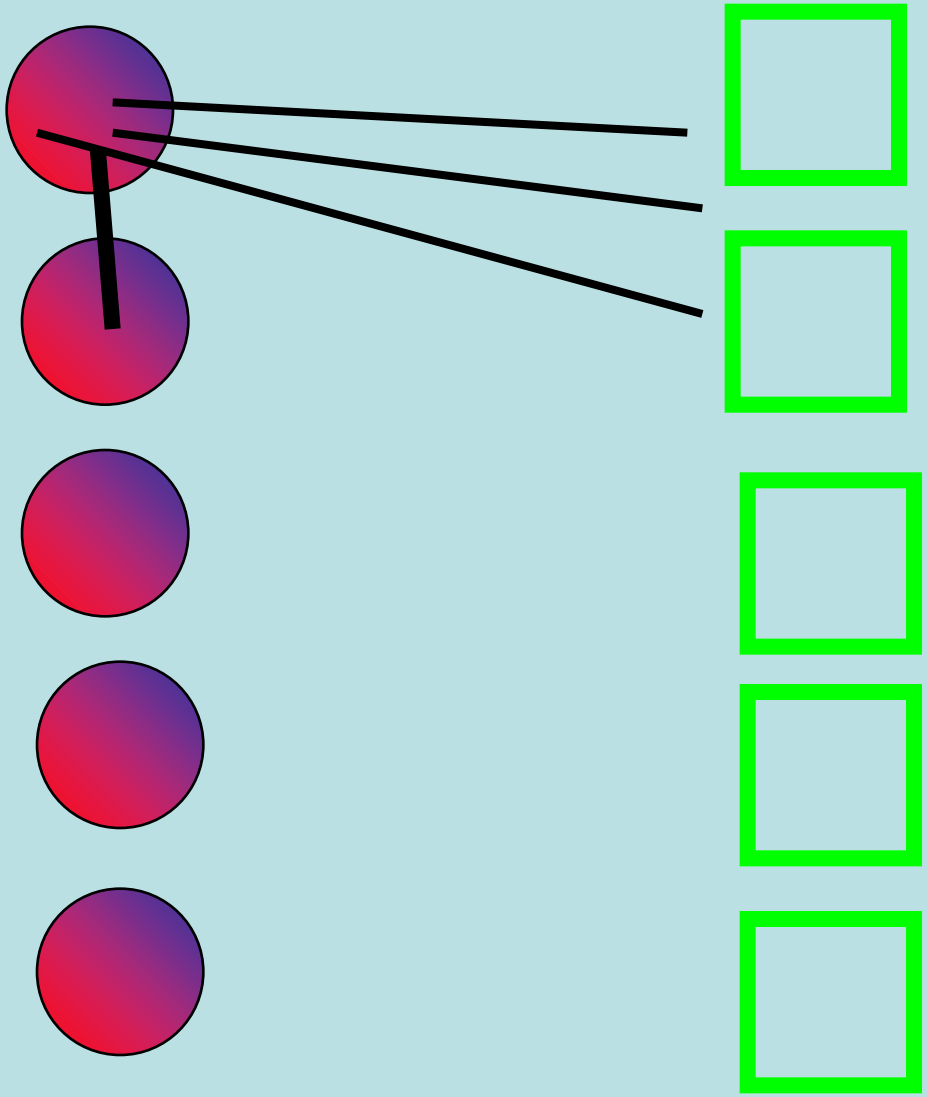
Suppose the graph is 4-stable…



# Combinatorial algorithm

Suppose the graph is 4-stable…

Look at the heaviest edge (at each vertex) – it can't be internal…

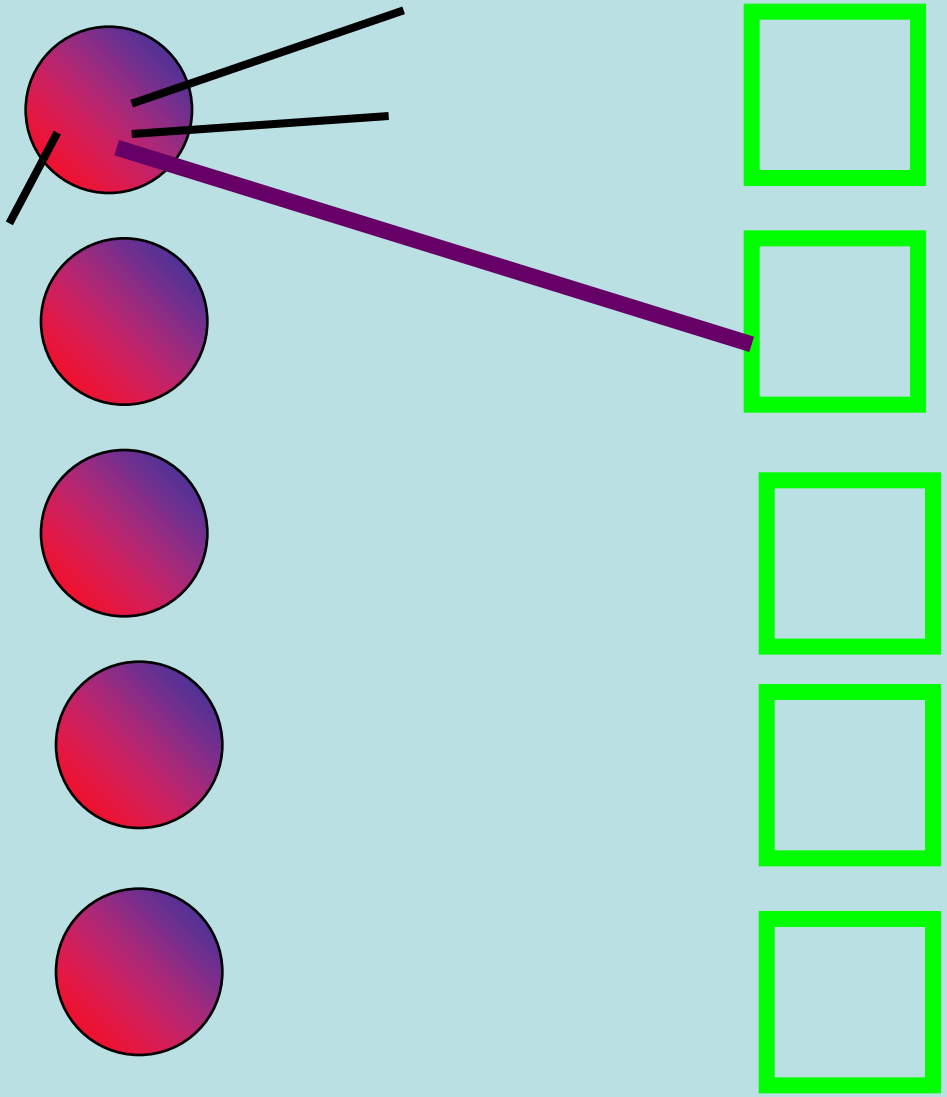


# Combinatorial algorithm

Suppose the graph is 4-stable…

Look at the heaviest edge (of each vertex) – it can't be internal…

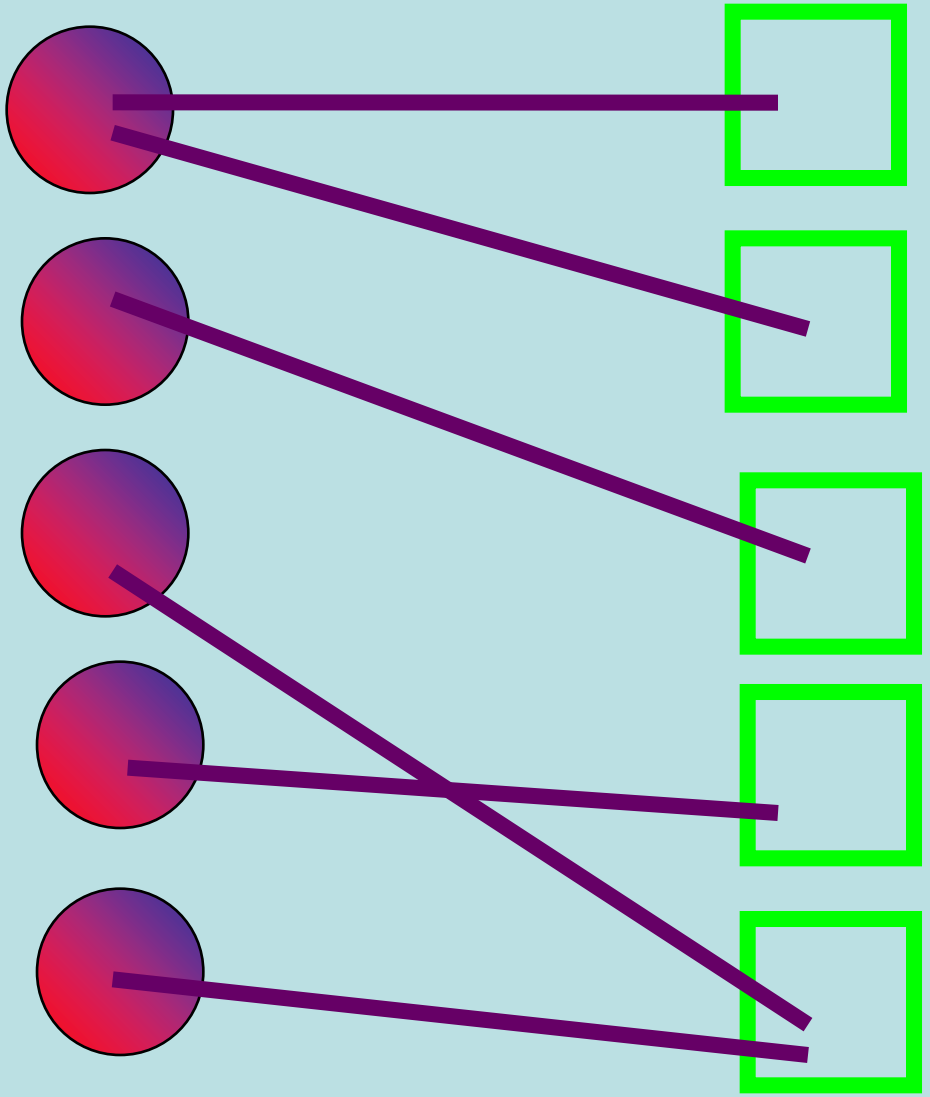
Its vertices must be on different side of the Max Cut.



# Combinatorial algorithm

Continue this way as long as you can rely on stability…

You can go all the way if $\gamma \geq \sqrt{\Delta n}$



# Combinatorial algorithm II

For simple graphs with minimal degree $\delta$,

Max-Cut is easy when $\gamma \geq 4n/\delta$.

# Spectral Algorithm

**Max Cut:** minimize $x^tAx$ with $x_i$ in {-1,1}

**Stability:** the same x will minimize $x^t(A \cdot P)x$

(P symmetric, and entries between 1 and γ)

**Eigenvector:** minimize $u^tAu$ (all of u with norm 1)

$U = |uu^t|$ : What x in $\{-1,1\}^n$ minimizes $x^t(A \cdot U)x$?

- The sign vector of u ($A \cdot U = |u^t|A|u|$)
- The Max Cut (if entries of U between 1 and γ)

# Spectral Algorithm

**Spectral partitioning:** Partition the graph according to the sign of the eigenvector associated with the least eigenvalue.

Works if entries of U between 1 and γ :

$$\gamma \geq (|\max u_i \cdot u_j / \min u_i \cdot u_j|)^2$$

(u eigenvector associated with least eigenvalue)

# Spectral Algorithm

**Second Least Eigenvalue:** If G is a $\gamma$-locally stable, d-regular graph:

$$\gamma > 2d/(\lambda_{n-1}+d) - 1$$

Max-Cut solvable for 3-regular graphs with

$$\lambda_{n-1} > -1$$

Boppana '87

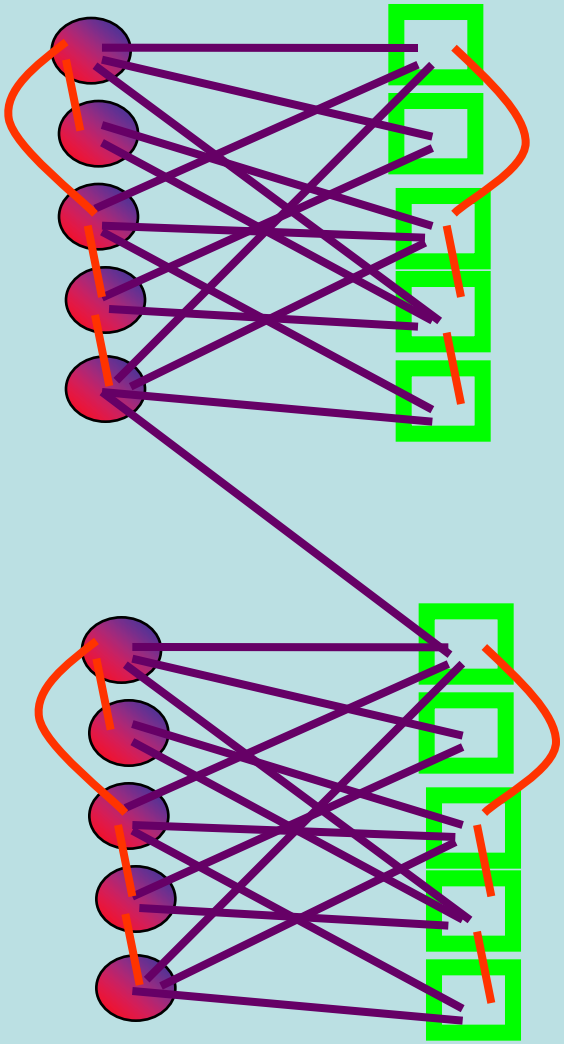
# Spectral Algorithm

**Expanders:** If G is a γ-locally stable, (d,λ)-expander with:

$$\gamma > (5d + \lambda) / (d - \lambda)$$

Spectral Algorithm works unless there are small cuts.

# (γ,ε)-Stability



# (γ,ε)-Stability

**1.** γ-stability (local stability)

**2**. Adding an ε-fraction of edges does not change the structure of the optimal solution, while increasing degrees by at most 2ε**.**

**Easy when: $\gamma > 12 / \varepsilon^2$**

# Conclusion & Open Problems

- **Are stable instances easy?**
  - Dense simple graphs
  - Eigenvalue separation (expanders)
  - ($\gamma$,$\varepsilon$)-Stability
- **Are stable instances hard?**
- **When does spectral partitioning work?**
- **Approximations?**
- **Alternative definitions?**