The intrinsic dimensionality of graphs

Robert Krauthgamer (UC Berkeley, IBM Almaden)

Joint work with James R. Lee (UC Berkeley)

Geometry of graphs

- Geometric representations $ graph properties.
  - Various relevant representations, e.g., low-distortion embeddings.
- Dimension of a graph G
  - Is the minimum d such that G has a d-dimensional representation (of certain type).
- Thesis: Low-dimension $ large diameter.

A dimensionality notion

- $\mathbf{Z}^d{}_1$:= d-dimensional integer lattice with diagonals .
  - Namely, the infinite graph whose vertex set is $\mathbf{Z}^d$ and which has an edge (u,v) whenever $\|u-v\|_1 = 1$.

- Let G=(V,E) be a (finite) graph and let n=|V|.
- dim(G) := smallest d such that G occurs as a (not necessarily induced) subgraph of $\mathbf{Z}^d{}_1$.
  - Always defined: dim(G) ≤ log n.

- Why l_1 ?
  - Balls of radius r contain ~$(2r)^d$ vertices.
  - In l_1 grid (no diagonals) degree is 2d << 2^d vertices.

A simple lower bound

- The growth rate of G is

$$r_G := \min\{r : |B(v,r)| \le r^r \text{ for all } v \in V, r > 1\}$$

where B(v,r) is a ball (in G) of radius r around v.

- That is, $r_G = \max\{\frac{\log|B(v,r)|}{\log r} : v \in V, r > 1\}$

- For example, $r(\mathbf{Z}^d{}_1) = \Theta(d)$.

- Corollary: $\dim(G) = \Omega(r_G)$. (Volume argument.)

The dimensionality conjecture

Conjecture (Levin & [Linial,London,Rabinovich]):

For every graph G, dim(G) = $O(r_G)$.

- I.e., structure has no role (only volume)
- Weaker form: Any bound in terms of r_G.

Previously known:

- Holds for cubes and complete binary trees [LLR].
- What about trees (even the weaker form)? [Linial]

Another Variant

- Equivalent formulation dim(G):

Smallest d such that $\exists$embedding $f:V\to\mathbf{Z}^d$ with

1. Contractive: $\|f(u)-f(v)\|_1 \leq 1$ for all (u,v)2 E.
2. Injective: $f(u) \neq f(v)$ for all $u \neq v$.

- A Euclidean analogue $\dim_2(G)$, due to Linial:

Smallest d such that $\exists$embedding $g:V\to\mathbf{R}^d$ with

1. $\|g(u)-g(v)\|_2 \leq 2$ for all (u,v)2 E.
2. $\|g(u)-g(v)\|_2 \geq 1$ for all $u \neq v$.

Easy to see that $\dim_2(G) = \Omega(r_G)$.

Our results

We resolve Levin s conjecture.

- Upper bounds:
  2 • dim(G) $\leq$ O(r_G) for trees (and chordal graphs etc.).
  - dim(G) $\leq$ O(r_G) for graphs excluding a fixed minor.
  3 • dim(G) $\leq$ O($r_G \log r_G$) for general graphs.
- Lower bound:
  1 • $\exists$graphs G where dim(G) $\geq \Omega(r_G \log r_G)$.
- Extensions:
  - Our results extend to $\dim_2$(G) and to metric spaces.

dim(G) vs. $\dim_2(G)$

Direct relationship: $\dim(G) \le O(\dim_2(G) \log \dim_2(G))$.

- **But:** dim(G) is equivalent to

1 . Contraction: For all (u,v)2E and j, $|f_j(u)-f_j(v)| \le 1$

2 . Injection: For all $u \ne v$ $\exists j$ s.t. $f_j(u) \ne f_j(v)$

- If (2) holds actually for $\ge \alpha$-fraction of j s then

1 . $\|f(u)-f(v)\|_2^2 \le d$ for all (u,v)2E

2 . $\|f(u)-f(v)\|_2^2 \ge \alpha d$ for all $u \ne v$

- Giving also the $\dim_2(G)$ result (up to scaling)

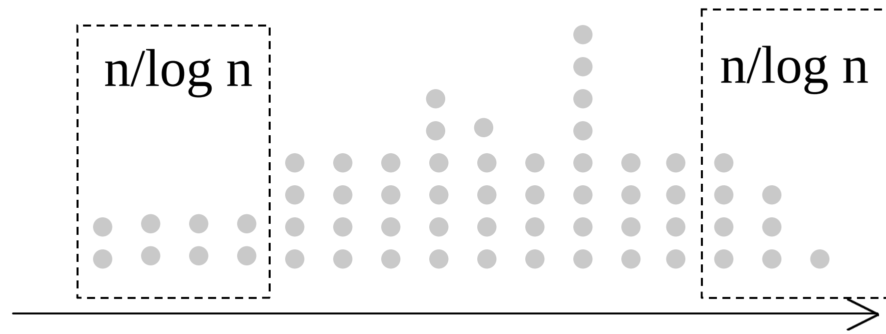
- This is the case in our constructions.

Related work

- Numerous geometric representations (too many to list)
- Similar-flavor embeddings:
  - Nonexpansive hashing [Linial, Sasson], [Indyk, Motwani, Raghavan, Vempala]
  - Network emulation by another network
  - Bilipschitz embedding of snowflake metrics [Assouad]
- A dual question to ours is Bandwidth:
  - Injective embedding to **Z** with minimum edge-stretch
- Growth-restricted metrics occur in practice:
  - Data points on a manifold
  - Networks: peer-to-peer, internet (?)

1. Lower bound

- Consider k-regular expander (e.g., random graph).
  - $r_G = O(\log n / \log \log_k n)$.
  - Take $k = O(\log n)$ so $r_G = O(\log n / \log \log n)$.
- Assume 9 f:V ! $\mathbf{Z}^d$ with $d << r_G \log r_G = O(\log n)$.
- For every j:
  - Project on *j*th coordinate.
  - Remove the leftmost and rightmost n/log n vertices.



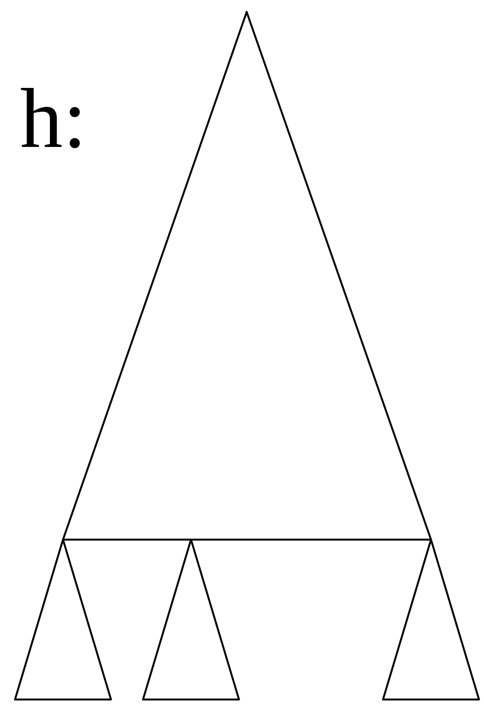
  - Remaining vertices are in $O(\log_k \log n)=O(1)$ interval.
- Remaining $> n/2$ vertices are mapped by f to $[O(1)]^d$.
- Since f is injective, $d \geq \Omega(\log n)$.
  - In general $d \geq \Omega(\log n / \log \log_k \log n)$.

Upper bounds

- The general scheme:
  - Construct coordinates iteratively (one at a time).
  - Each coordinate is contractive by definition.
  - Every u,v2V are separated in at least one coordinate.
- How many coordinates are sufficient?

Where is the problem?

- Exploiting the growth-restriction
- E.g., a complete binary tree of height h:
  - $r = \Theta(\mathrm{h}/\log \mathrm{h})$.
  - About 2^h pairs of sibling leaves
  - Seems to require ~h dimensions
  - [Linial, London, Rabinovich]:
    - Correlate all pairs at the same depth

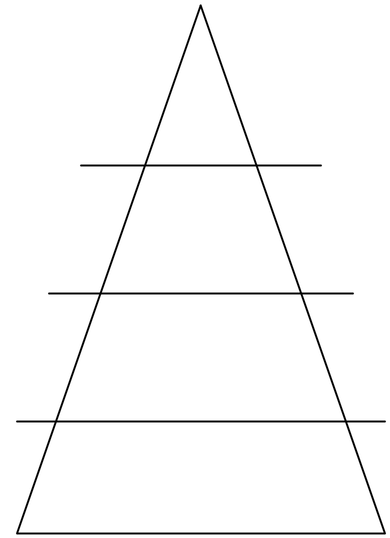


- In general:
  - We wish to handle each locality separately,
  - But we only have an upper bound on their total size!

2. Trees naive approach

- Construct a coordinate f_j:
  - Fix a root and map it to 0.
  - Assign each edge 1/+1 at random.
  - Map vertices to their weighted distance from root.
- Consider a specific scale r
  - Let u,v2 V with d(u,v) = r.
  - $\Pr[f_j(u)=f_j(v)] = O(r^{1/2})$. (Random walk)
  - $O(r)$ repetitions (i.e., coordinates) reduces the collision probability to $O(r^{8r})$.
  - We cannot take union bound over all u,v with d(u,v) = r because number of pairs u,v is huge! (depends on n)

Trees local approach



- Cut the tree every 2r levels.
- For every connected component (cluster) C:
  - C has radius ≤ 2r.
  - C has ≤ $(2r)^r$ vertices.
  - C has ≤ $(2r)^{2r} \le r^{4r}$ pairs of vertices.
  - So with high probability, every $u,v \in C$ with $d(u,v)=r$ is separated (in the $O(r)$ coordinates) at least once.
  - The random-walk & union bound argument can be applied.
- Glue these mappings:
  - Assign the inter-cluster edges arbitrarily (say 0).

Trees   staggering

- What about u,v that belong to different clusters?
  - Construct another layer of cuts that is shifted , i.e., cut every 2r levels but start from level r.
  - Every pair u,v with d(u,v) = r belongs to the same cluster in at least one layer.
- We have 2 layers, each using O(r) coordinates.
  - Their concatenation separates all u,v with d(u,v)=r.
  - In fact, they handle all distances 2 [$r^{1/2}$,r].
- For all scales, we need O(r loglog n) coordinates!
  - Can we handle different scales simultaneously?

Trees   reusing coordinates
(or not using all your ammo at once)

- Observation: When a scale r is handled, edges between different clusters (of radius 2r) remain open .
- Consider every other scale $r^{1/4}$, $\cancel{r^{1/2}}$, r, $\cancel{r^{2}}$, r^4, $\cancel{r^{8}}$, r^{16},
  - Increases the dimension by a factor of 2.
- Let s handle scale r (i.e., distances 2 [$r^{1/2}$,r]) assuming scale $r^{1/4}$ was already handled:
  - There is an open edge every $2r^{1/4}$ edges
  - Every path of length $\geq r^{1/2}$ contains $\geq r^{1/4}/2$ open edges
  - Random assignment of the open edges will do!
- Technicality: Induction interferes with staggering.

3. Embedding via decomposition

Overview:

- Fixing r, decompose G into clusters of diameter $\leq r^4$.
  - Let m be the number of layers used
  - Every u,v with d(u,v) = r are, in some layer, in the same cluster
- Embed each cluster C such that:
  - Every u,v2 C with d(u,v) = r are separated
  - The boundary of C is mapped to 0 (to allow glueing).
  - Let q be the number of coordinates used.
- Total number of coordinates O(mq loglog n).
- Nest different r s to use only O(mq)=O(r^3) dimensions.

Naive embedding of a cluster C

- C has diameter $\leq r^4$ and thus $\leq r^{4r}$ vertices.
- Inner decomposition:
  - Decompose C into subclusters of diameter $\leq r^{1/4}$.
  - Then every u,v2C with d(u,v)2[$r^{1/2}$,r] are in different subclusters.
- Suppose every subcluster was randomly mapped to a value 2 [0,$r^{1/8}$].
  - Then $\Pr[f_j(u)=f_j(v)] \leq 1/r^{1/8}$.
  - Pr[collision in O(r) repetitions] $\leq 1/r^{8r}$.
- Number of vertex pairs in C is $\leq r^{8r}$.
- So by union bound, WHP all these pairs are separated!
- But how to glue the different subclsuters?

Smooth embedding via padding

- We decompose C using q layers s.t.
  - Diameter: Every subcluster has diameter $\leq r^{1/4}$.
  - Subcluster boundary is $\partial(C_i) := \{u \in C_i$ adjacent in G to $v \notin C_i\}$.
    - This includes the boundary of the cluster C.
  - Padding: Every $v \in C$ is $r^{1/8}$-far from the boundary of its subcluster in at least one layer.
- Every layer yields a coordinate:
  - Extend subcluster boundary (inwards) by a value $\in [0, r^{1/8}]$.
  - Map each vertex to its distance to the extended boundary.
  - Thus $\Pr[f_j(u)=f_j(v)] \leq 1/r^{1/8}$ so $O(r)$ repetitions suffice etc.
  - Glueing is possible since all boundaries are mapped to 0.
- Total number of coordinates = $O(m\ q\ \ r$ loglog n).

Further issues

- Reusing coordinates
  - Forced nesting
  - Base case r = *r*
- Decomposition
  - Requirements
  - Implementation
- Techniques:
  - Decomposition a la [Linial-Saks]
  - Lovasz Local Lemma

Reusing coordinates   forced nesting

- Observation: Edges between clusters are still open
  - I.e., boundaries of different clusters can be mapped to different values.
- Let us embed scale r (i.e., cluster of diameter $\leq r^4$)
  - Suppose we already embedded smaller scales, say clusters of diameter $\leq r^{1/20}$.
  - Then every $r^{1/8}$ path contains $\geq r^{1/16}$ edges that are open between clusters of diameter $r^{1/20}$.
  - But these open edges might belong to clusters of even smaller diameter, e.g. $r^{1/640}$.
  - We thus shrink clusters (inwards) to avoid cutting smaller clusters (losing at most $r^{1/20}$ of the $r^{1/8}$ padding).

The Lovasz local lemma

- Let $A_1, \quad , A_n$ be events where each A_i is independent of all but at most d other events A_j.
- Suppose that $\Pr[A_i] \leq p$ for all i, and $ep(d+1) \leq 1$.
- Then $\Pr[\cap_i \quad _i] > 0$.

Forced nesting   base case

Consider $\mathrm{r} \le \boldsymbol{r}$.

- Assume first that $\mathrm{n} \le \mathrm{r}^{O(r)} \le \boldsymbol{r}^{O(r)}$.
  - Map the vertices injectively (or randomly) to $\{0,1\}^{O(r \log r)}$.
- For general n,
  - Map the vertices randomly to $\{0,1\}^{cr \log r}$ for constant c.
  - Let $\mathsf{E}_v = \{f(v)$ is not unique within distance r from $v\}$.
  - Then $\Pr[\mathsf{E}_v] \le \mathrm{r}^{O(r)}/\boldsymbol{r}^{cr} \le 1/\boldsymbol{r}^{c\,r}$.
  - Degree of dependency is $\le \mathrm{r}^{O(r)}$.
  - So by the Local Lemma: $\Pr[\text{none of } \mathsf{E}_v] > 0$.

Decomposition requirements

- Inner decomposition (of a cluster C with |C| ≤ r^{4r}):
  - Diameter of subclusters ≤ $r^{1/4}$.
  - Every u,v2C with d(u,v) ≤ r belong to the same cluster in at least one layer with (say) u being at least $r^{1/8}$-far from boundary.
- Outer decomposition (of V with |V|=n):
  - Diameter of clusters ≤ r^4.
  - Every u,v2V with d(u,v) ≤ r belong to the same cluster in at least one layer with (say) u being at least $r^{1/8}$-far from the boundary.
- Nearly the same requirements!

Graphs excluding $K_{s,s}$-minor

- We use the [Klein, Plotkin, Rao] decomposition.
  - Take a BFS from an arbitrary vertex.
  - Cut
    1. Every $r^2/100$ levels.
    2. Every $r^2/100$ levels starting at level $r^2/200$.
  - Repeat recursively until depth s to get 2^s layers.
- Our requirements holds:
  - Diameter of every cluster $\le r^2$ by [KPR].
  - Every $u,v \in V$ with $d(u,v) \le r$ belong to the same cluster in at least one layer with (say) u being at least $r^{1/8}$-far from the boundary.
- Thus, $\dim(G) \le O(4^s\, r)$.

Inner decomposition

- Let $\mathrm{r} > r$ and assume first that $\mathrm{n} \le \mathrm{r}^{O(r)}$.
- We use a decomposition of [Linial,Saks],[Bartal].
  - Fix an ordering of the vertices.
  - Every vertex defines a ball whose radius is randomly chosen from exponential distribution with mean r^2.
  - WHP all radii are bounded by $O(\mathrm{r}^2 \log \mathrm{n})$.
  - Define clusters:
    - A vertex belongs to the first ball that contains it.
    - Each cluster has weak diameter $O(\mathrm{r}^2 \log \mathrm{n}) = O(\mathrm{r}^4)$.
    - Each cluster has size $\le \mathrm{r}^{O(r)}$.
  - Pr[a ball of radius r around u is cut] $\le O(1/\mathrm{r})$.
  - Pr[cutting such ball in $O(r)$ repetitions] $\le 1/\mathrm{r}^{cr}$.
- Hence $\mathrm{m} = \mathrm{q} = O(r)$.

Outer decomposition

- For general n we use the local lemma.
  - Let E_v = {ball of radius r around u is cut in all $O(r)$ layers of the decomposition}
  - Then $\Pr[E_v] \le 1/r^{cr}$.
  - Degree of dependency graph is $\le r^{O(r)}$.
  - So by the Local Lemma Pr[none of E_v] > 0.

- We obtain $m = O(r)$.
- Total number of coordinates = $O(r^3)$.

A tight upper bound

- We can achieve O(r log r) by
  - Applying all these arguments on the same O(r) coordinates,
  - And arguing using Chernoff bounds that the desired event occurs in constant fraction of the coordinates.
  - We need O(r log r) coordinates to handle the base case r $= r$.

Conclusion

- Is it true that $\dim(G) = \Theta(\dim_2(G))$?
- Applications of these notions?
  - Our embeddings actually map distance $r > r$ to distance $\in [r^{1/2}, r]$.
- A dual question is Bandwidth:
  - Density lower bound: $D_G := \max \{|B(v,r)|/2r : 8\ v, r\}$.
  - What is the gap between them?
    - Can be $\Omega(\log n)$ (e.g., expander).
    - At most polylog(n) [Feige].
  - What is the tradeoff between dimension and stretch?

Thank you!