The diameter problem on graphs: a geometric point of view

G. Ducoffe

University of Bucharest & I.C.I. Bucharest, Romania

December 9th, 2021

Joint work with Pierre Bergé, Feodor Dragan, Michel Habib, Heather Guarnera and Laurent Viennot.

My first meeting with Victor

- 13èmes Journées Combinatoires et Algorithmes du Littoral Méditerranéen (JCALM **2013**).

SIAM J. Discrete Math.
Vol. 16, No. 2, pp. 323–334

1-HYPERBOLIC GRAPHS*

HANS-JÜRGEN BANDELT† AND VICTOR CHEPOI‡

Abstract. The shortest-path metric d of a graph $G = (V, E)$ is called *δ-hyperbolic* if for any four vertices $u, v, w, x \in X$ the two larger of the three sums $d(u, v) + d(w, x), d(u, w) + d(v, x), d(u, x) + d(v, w)$ differ by at most δ. In this paper, we characterize the graphs with 1-hyperbolic metrics in terms of a convexity condition and forbidden isometric subgraphs.

- A question about the recognition of graphs with small hyperbolicity.

SIAM J. Discrete Math.
Vol. 28, No. 3, pp. 1601–1617

RECOGNITION OF C_4-FREE AND 1/2-HYPERBOLIC GRAPHS*

DAVID COUDERT† AND GUILLAUME DUCOFFE‡

Abstract. The shortest-path metric d of a connected graph G is 1/2-*hyperbolic* if and only if it satisfies $\mathrm{d}(u,v) + \mathrm{d}(x,y) \leq \max\{\mathrm{d}(u,x) + \mathrm{d}(v,y), \mathrm{d}(u,y) + \mathrm{d}(v,x)\} + 1$, for every 4-tuple u, x, v, y of G. We show that the problem of deciding whether an unweighted graph is 1/2-hyperbolic is *subcubic equivalent* to the problem of determining whether there is a chordless cycle of length 4 in a graph. An improved algorithm is also given for both problems, taking advantage of fast rectangular matrix multiplication. In the worst case it runs in $O(n^{3.26})$-time.

Problems considered

Problem (ECCENTRICITIES)

Input: *A connected unweighted graph* $G = (V, E)$.

Output: $e(v) =$ *max. # of edges on a shortest path between* v *and any other vertex,* $v \in V$.

Harary's **Graph centrality**: $1/e(v)$.

Special cases:

- **Diameter**: $diam(G) = \max_{v \in V} e(v)$.
- Center: $C(G) = \{v \in V \mid e(v) = argmin_{u \in V} e(u)\}$.

• First step toward a better understanding of the shortest-path structure of some graph classes

(applications to spanners, distance-labeling schemes, etc. . .)

Computing all eccentricities

- In $\mathcal{O}(nm)$ using n BFS **[folklore]**

Seems to require $\Theta(n)$ BFS to be solved on general graphs

- In $\tilde{\mathcal{O}}(n^{2.373})$ time using fast matrix multiplication **[Seidel, STOC'92]**
    - $\omega = 2 \implies$ almost linear time for **dense** graphs
    - "Combinatorial algorithms" seem to require $\Omega(n^{3-o(1)})$.

In all cases it is a reduction to APSP: $\Omega(n^2)$ time, even on sparse graphs.

Can we do better?

Lower bound

- **Orthogonal Vector** (*a.k.a.*, **Disjoint Sets**)

Input: two families A, B of n sets over some universe C, with $|C| = d \ll n$.

Question: $\exists?\ a \in A,\ b \in B$ such that $a \cap b = \emptyset$

Theorem (Chepoi and Dragan, unpublished)

Orthogonal Vector $\propto$ **Diameter** ≤ 2

- **Strong Exponential-Time Hypothesis**: SAT cannot be solved in $\mathcal{O}^*((2-\epsilon)^n)$ time, for any $\epsilon > 0$.

Theorem (Williams, TCS 2005)

SETH $\Longrightarrow$ OV not in $\mathcal{O}(n^{2-\varepsilon})$ time, for any $\varepsilon > 0$.

Breaking the quadratic barrier

- DIAMETER remains SETH-hard in very restricted cases:
  - Bounded-degree graphs with diameter in $\omega(\log n)$;
  - Bipartite graphs with one side of size $\omega(\log n)$;
  - Split graphs of clique number $\omega(\log n)$.

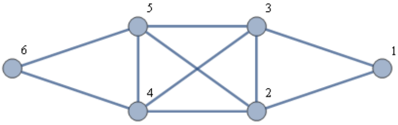
- However, it seems that "all" conditional lower bounds for DIAMETER start with a reduction from OV.

- There are several graph classes where we can solve OV efficiently.

$\Longrightarrow$ *Does it imply fast diameter computation?*

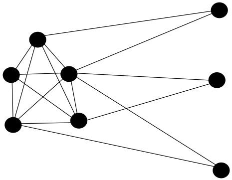
The case of **chordal graphs**

- No induced cycle $C_k, \ k > 3$.



Applications in Data Bases and in Biology (leaf-power), etc.

- Special case: **Split Graphs** (vertex-partition Clique + Stable Set)



Computing $diam(G)$ in $\mathcal{O}(n^{2-\epsilon})$ is already SETH-hard!

From chordal graphs to split graphs

[D. & Dragan, Networks 2021]

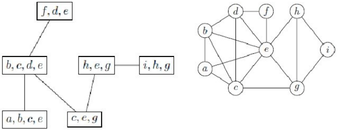
Theorem (Dragan & D., Networks 2021)

For a subclass $\mathcal{C}$ of chordal graphs, let $\mathcal{S}$ contain all split graphs that are induced subgraphs of a graph in $\mathcal{C}$.

If for all $G \in \mathcal{S}$, for every bipartition A, B of its stable set, we can solve OV in $\mathcal{O}(\ell^b)$ time, then there is a randomized $\mathcal{O}(m^b \log^2 n)$-time algorithm for computing w.h.p. the diameter of chordal graphs in $\mathcal{C}$.

Intuition: Centroid decomposition on a clique-tree.

Split graphs are constructed from centroid-cliques + a subset of neighbours ("gates" of the other vertices).



Applications

- For every n-vertex m-edge chordal graph, we can compute an additive $+1$-approximation of all eccentricities in total $\mathcal{O}(m \log n)$ time.

- Almost linear-time (randomized) algorithm for computing the diameter of chordal graphs of bounded asteroidal number

[D., JGT 2021+]

- Study of a new parameter on split graphs: the "clique-interval number"

[D., Habib, Viennot; DMTCS 2021]

$\Longrightarrow$ at most a constant on bounded-treewidth split graphs, comparability split graphs, etc.

$\Longrightarrow$ Almost linear-time algorithms for strongly chordal graphs, chordal comparability graphs, bounded-treewidth chordal graphs, etc.

A partial characterization

Recall that: **hereditary = closed by induced subgraphs**

Theorem (D., MFCS'21)

Under SETH, for any hereditary subclass $\mathcal{C}$ of chordal graphs, the following statements are equivalent:

1. *There is a truly subquadratic algorithm for computing the Wiener index within $\mathcal{C}$ $(= \sum_{u,v} d(u,v))$.*
2. *There is a truly subquadratic algorithm for computing the diameter within $\mathcal{C}$.*
3. *There is a truly subquadratic algorithm for deciding if a graph in $\mathcal{C}$ has diameter ≤ 2.*
4. *$\mathcal{C}$ does not contain all the split graphs.*
5. $\mathcal{C}$ **has bounded VC-dimension.**

Consequences

The following subclasses of chordal graphs all admit truly subquadratic algorithms for the Wiener index and the diameter problem:

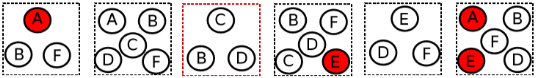
- chordal bull-free graphs
- chordal claw-free graphs
- block graphs
- interval graphs
- strongly chordal graphs
- directed path graphs
- undirected path graphs
- chordal dominating pair graphs
- hereditary Helly graphs
- k-separator chordal graphs
- chordal graphs of bounded interval number
- chordal graphs of bounded asteroidal number
- chordal graphs of bounded leafage

For **non hereditary** chordal subclasses the picture is much less clear: linear-time for chordal Helly graphs **[D., Dragan; Networks 2021]**

VC-dimension

Let $\mathcal{H} = (X, \mathcal{R})$ be a hypergraph/range space/set family (possibly infinite)

- $S \subseteq X$ is **shattered** if $\{S \cap R \mid R \in \mathcal{R}\} = 2^S$
- The **VC-dimension** of $\mathcal{H}$ is the largest cardinality of a shattered subset **[Vapnik and Chervonenkis, 1971]**.



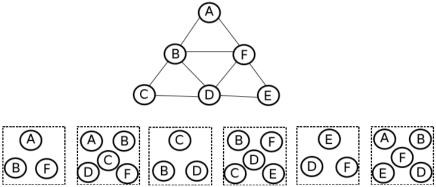
$S = \{A, E\}$ is shattered, VC-dim $= 2$

- Several applications in Learning Theory and Computational Geometry.

VC-dimension **of graphs**

The **neighbourhood hypergraph** of $G = (V, E)$ is defined as $\mathcal{N}(G) = (V, \{N_G[v] \mid v \in V\})$.

$$\text{VC-dim}(G) =^{def} \text{VC-dim}(\mathcal{N}(G))$$



A key lemma:

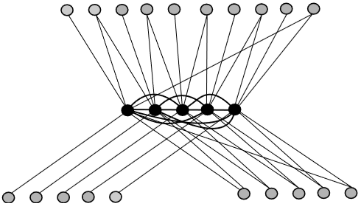
Lemma

If H is a split graph, then every H-free chordal graph has VC-dimension at most $|V(H)| - 1$.

Improves on Bousquet et al. **[SIDMA 2015]**.

Warm-up: Split graphs with small clique

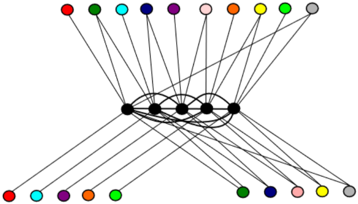
Consider a split graph $G = (K \cup S, E)$ of clique-number $|K| = \log^{\mathcal{O}(1)}(n)$.



Warm-up: Split graphs with small clique

Consider a split graph $G = (K \cup S, E)$ of clique-number $|K| = \log^{\mathcal{O}(1)}(n)$.

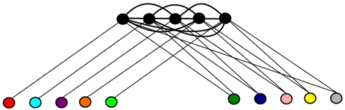
- Two vertices $s, t \in S$ are *twins* if $N(u) = N(v)$.



Warm-up: Split graphs with small clique

Consider a split graph $G = (K \cup S, E)$ of clique-number $|K| = \log^{\mathcal{O}(1)}(n)$.

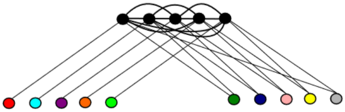
- Two vertices $s, t \in S$ are *twins* if $N(u) = N(v)$.
- We can keep one vertex per twin class **[Coudert et al., SODA'18]**.



Warm-up: Split graphs with small clique

Consider a split graph $G = (K \cup S, E)$ of clique-number $|K| = \log^{\mathcal{O}(1)}(n)$.

- Two vertices $s, t \in S$ are *twins* if $N(u) = N(v)$.
- We can keep one vertex per twin class **[Coudert et al., SODA'18]**.



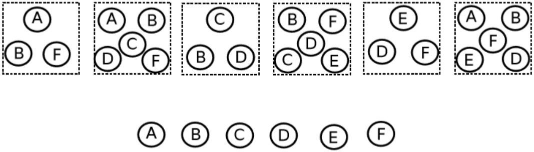
(Sauer-Shelah-Perles) VC-dim$(\mathcal{H}) = \mathrm{d} \Longrightarrow \#\{Y \cap R \mid R \in \mathcal{R}\} = \mathcal{O}(|Y|^d)$.

$\longrightarrow$ There are only $\mathcal{O}(|K|^d) = \log^{\mathcal{O}(d)}(n)$ twin classes!

Stabbing Number

For a hypergraph $\mathcal{H} = (X, \mathcal{R})$, a spanning path = a total order over X.

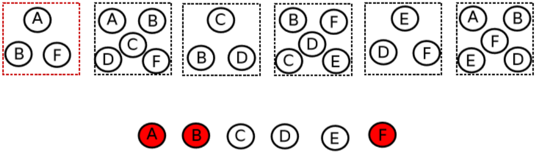
- stabbing number $\sim$ max. # of intervals to represent a hyperedge



Stabbing Number

For a hypergraph $\mathcal{H} = (X, \mathcal{R})$, a spanning path = a total order over X.

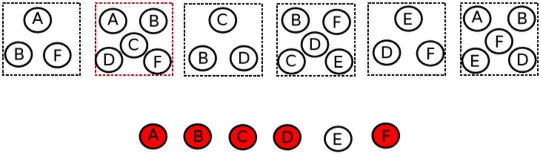
- stabbing number $\sim$ max. # of intervals to represent a hyperedge



Stabbing Number

For a hypergraph $\mathcal{H} = (X, \mathcal{R})$, a spanning path = a total order over X.

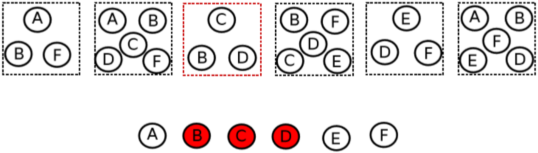
- stabbing number ∼ max. # of intervals to represent a hyperedge



Stabbing Number

For a hypergraph $\mathcal{H} = (X, \mathcal{R})$, a spanning path = a total order over X.

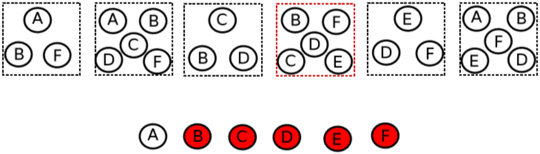
- stabbing number $\sim$ max. # of intervals to represent a hyperedge



Stabbing Number

For a hypergraph $\mathcal{H} = (X, \mathcal{R})$, a spanning path = a total order over X.

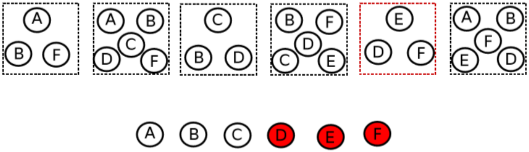
- stabbing number $\sim$ max. # of intervals to represent a hyperedge



Stabbing Number

For a hypergraph $\mathcal{H} = (X, \mathcal{R})$, a spanning path = a total order over X.

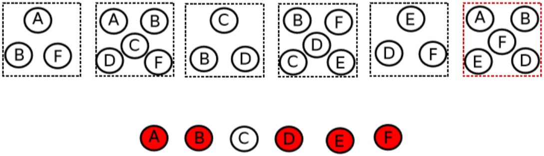
- stabbing number $\sim$ max. $\#$ of intervals to represent a hyperedge



Stabbing Number

For a hypergraph $\mathcal{H} = (X, \mathcal{R})$, a spanning path = a total order over X.

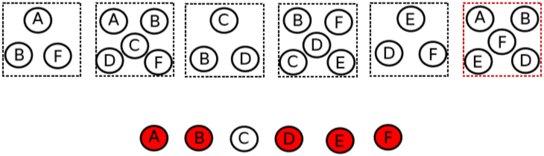
- stabbing number $\sim$ max. # of intervals to represent a hyperedge



Stabbing Number

For a hypergraph $\mathcal{H} = (X, \mathcal{R})$, a spanning path = a total order over X.

- stabbing number $\sim$ max. # of intervals to represent a hyperedge

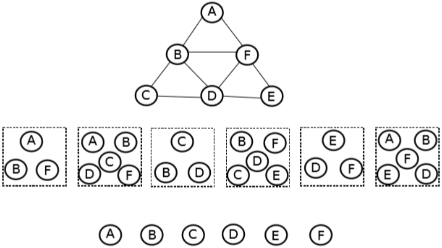


Every hypergraph of VC-dim $\leq d$ has a spanning path of stabbing number $\tilde{\mathcal{O}}(n^{1-1/2^d})$ **[Chazelle and Welzl, DCG'89]**

Application to DIAMETER on Split graphs

Goal: decide whether $diam(G) \leq 2$

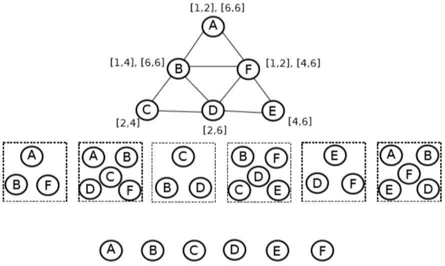
- Compute a spanning path of stabbing number t for $\mathcal{N}(G)$.



Application to DIAMETER on Split graphs

<u>Goal</u>: decide whether $diam(G) \leq 2$

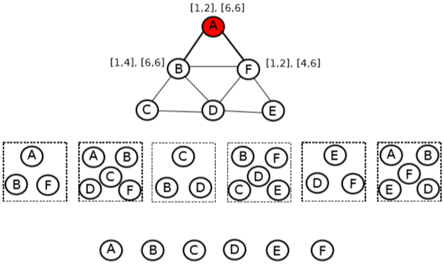
- Compute a spanning path of stabbing number t for $\mathcal{N}(G)$.
- Every vertex v collects the ends $I(v)$ of the $\mathcal{O}(t)$ intervals for $N[v]$.



Application to DIAMETER on Split graphs

Goal: decide whether $diam(G) \leq 2$

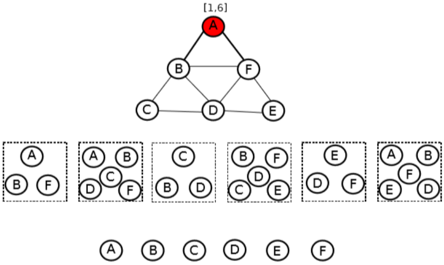
- Compute a spanning path of stabbing number t for $\mathcal{N}(G)$.
- Every vertex v collects the ends $I(v)$ of the $\mathcal{O}(t)$ intervals for $N[v]$.
- Every vertex v collects the sets $I(u),\ \forall u \in N[v]$.



Application to DIAMETER on Split graphs

Goal: decide whether $diam(G) \leq 2$

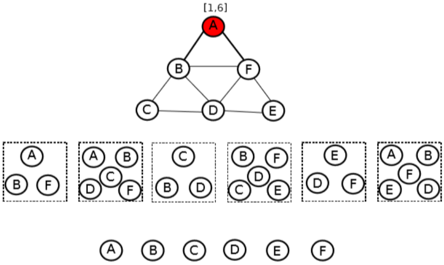
- Compute a spanning path of stabbing number t for $\mathcal{N}(G)$.
- Every vertex v collects the ends $I(v)$ of the $\mathcal{O}(t)$ intervals for $N[v]$.
- Every vertex v collects the sets $I(u),\ \forall u \in N[v]$.
- Check whether $\forall v \in V,\ \bigcup_{u \in N[v]} I(u) = V$.



Application to DIAMETER on Split graphs

Goal: decide whether $diam(G) \leq 2$

- Compute a spanning path of stabbing number t for $\mathcal{N}(G)$.
- Every vertex v collects the ends $I(v)$ of the $\mathcal{O}(t)$ intervals for $N[v]$.
- Every vertex v collects the sets $I(u),\ \forall u \in N[v]$.
- Check whether $\forall v \in V,\ \bigcup_{u \in N[v]} I(u) = V$.



Complexity: Spanning Path Computation + $\sum_v (deg(v)+1) \cdot \mathcal{O}(t)$

$$= \text{Spanning Path Computation} + \mathcal{O}(tm).$$

VC-dimension & Diameter: Main Result

[D., Habib, Viennot; SODA'20]

The proof of **[Chazelle and Welzl, DCG'89]** for computing a "good" spanning path is constructive... but it leads to an $\tilde{\mathcal{O}}(n^3 m)$-time algorithm.

Fortunately, a "good enough" spanning path can be computed in sub-quadratic time:

Theorem

For any $d > 0$, there is some $\varepsilon_d \in (0;1)$ s.t. in $\tilde{\mathcal{O}}(m + n^{2-\varepsilon_d})$ time, for every n-vertex hypergraph $\mathcal{H} = (X, \mathcal{R})$ of VC-dimension $\leq d$ and size $m = \sum_{e \in \mathcal{R}} |e|$, we can compute a spanning path of stabbing number $\tilde{\mathcal{O}}(n^{1-\varepsilon_d})$. Moreover, $\varepsilon_d = \frac{1}{2^{d+1}[c(d+1)-1]+1}$ for some constant $c > 2$.

$\Longrightarrow$ Diameter **in $\tilde{\mathcal{O}}(mn^{1-\varepsilon_d})$ for split graphs of VC-dim $\leq d$**

Beyond Chordal graphs: graphs of bounded VC-dimension

Our previous proof only relies on bounded VC-dim. + that split graphs have diameter ≤ 3.

Therefore:

Theorem

For every $d > 0$, there exists a constant $\varepsilon_d \in (0;1)$ such that in deterministic time $\tilde{\mathcal{O}}(mn^{1-\varepsilon_d})$ we can decide whether a graph of **VC-dimension** *at most d has diameter two.*

Cannot be extended to sub-quadratic diameter computation because of bounded-degree graphs.

Generalization: **distance VC-dimension**

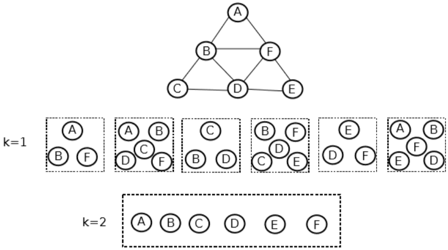
Recall: $N^k[v] = \{u \in V \mid dist_G(v,u) \leq k\}$.

- The **k-neighbourhood hypergraph** of G is: $\mathcal{N}_k(G) = (V, \{N^k[v] \mid v \in V\})$.

k-dist-VC-dim$(G) =^{def}$ VC-dim$(\mathcal{N}_k(G))$

- The **ball hypergraph** of G is: $\mathcal{B}(G) = (V, \{N^k[v] \mid v \in V, k \geq 0\})$.

dist-VC-dim$(G) =^{def}$ VC-dim$(\mathcal{B}(G))$



Generalization: **distance VC-dimension**

Related work

Notion first studied by **[Chepoi et al., DCG'07]**

- Constant distance VC-dimension for:

$\longrightarrow$ *proper minor-closed* graph classes **[Chepoi et al., DCG'07]**.

$\longrightarrow$ *bounded clique-width* graph classes **[Bousquet and Thomassé, DM'15]**.

- A monotone graph class $\mathcal{G}$ is *nowhere dense* iff $\forall k,\ \sup_{G \in \mathcal{G}} k\text{-dist-VC-dim}(G) < +\infty$ **[Nešetřil and Ossona de Mendez, RMS'16]**

distance VC-dimension & Diameter: Our Results

[D., Habib, Viennot; SODA'20]

Theorem

There exists a Monte Carlo algorithm such that, for every positive integers d and k, we can decide whether a graph of **distance VC-dimension** *at most d has diameter at most k. The running time is in $\tilde{\mathcal{O}}(k \cdot mn^{1-\varepsilon_d})$, where $\varepsilon_d \in (0;1)$ only depends on d.*

Can be improved for planar graphs, all proper-minor closed graph classes and beyond:

Theorem

Let $\mathcal{C}$ be a monotone (subgraph-closed) graph class with **bounded dist. VC-dim + polynomial expansion***.*

All eccentricities of a graph in $\mathcal{C}$ can be computed in time $\tilde{\mathcal{O}}(n^{2-\varepsilon_{\mathcal{C}}})$, with a deterministic algorithm, for some $\varepsilon_{\mathcal{C}} \in (0;1)$.

Challenges and techniques

Main Issue: we cannot compute $\mathcal{N}_k(G)$ for any $k \geq 2$.

Theorem (ε-net)

If VC-dim($\mathcal{H}$) $\leq d$, then any random subset of size $\approx \frac{d}{\varepsilon} \log n$ intersects all hyperedges of cardinality $\geq \varepsilon \cdot n$.

Algorithm:

- Use a random subset S as above in order to partition the vertices into equivalence classes: $u \sim v \Longleftrightarrow^{def} N^k[u] \cap S = N^k[v] \cap S$.
- **(Sauer's Lemma)** There are only $\tilde{\mathcal{O}}(\varepsilon^{-d})$ equivalence classes $V_1, V_2, \ldots, V_q$.
- **(ε-net)** $u \sim v \Longrightarrow |N^k[u] \Delta N^k[v]| = \mathcal{O}(\varepsilon n)$. *We keep one representative $v_i \in V_i$ per equivalence class.* Let $\mathcal{H}_k =^{def} (V, \{N^k[v_i] \mid 1 \leq i \leq q\})$.
- Deduce a spanning path for $\mathcal{N}_k(G)$ from $\mathcal{H}_k$ **and** $\mathcal{N}_{k-1}(G)$.

A nice conjecture

Conjecture: The diameter can be computed in $\tilde{\mathcal{O}}(mn^{1-\varepsilon_d})$ time in the class of graphs of dist. VC-dimension at most d, for some absolute constant ε_d.

$\longrightarrow$ True for **constant-diameter** graphs

$\longrightarrow$ True for chordal graphs **[D. & Dragan, Networks 2020]**

$\longrightarrow$ True for classes of polynomial expansion.

$\longrightarrow$ True for bounded clique-width graphs **[D., IPEC'21]**

Further results on the conjecture

Isometric embedding in a system/product of trees

Theorem

Let $G = (V, E)$ and $T_1, T_2, \ldots, T_k$ be a collection of k trees, where $N := \sum_{i=1}^{k} |V(T_i)|$. Given an isometric embedding of G in $\square_{i=1}^{k} T_i$ (Cartesian product), we can compute all eccentricities in G in $\mathcal{O}(2^{\mathcal{O}(k \log k)}(N+n)^{1+\varepsilon})$ time.

$\longrightarrow$ Special case of a product of graphs of bounded (distance) VC-dimension.

$\longrightarrow$ Also applies to quasi-isometric embeddings and to other type of tree products (*e.g.*, strong product).

Applications to several subclasses of **partial cubes** and ℓ_1-graphs (slightly extending prior results from **[Chepoi et al., SODA'02]** for benzenoid and triangular systems).

The odd case of retracts

$$H \text{ is a retract of } G \iff \text{there is a homomorphism } f : G \longmapsto H \text{ s.t. } f_{|H} = id_H$$

Implies that H is an isometric subgraph of G.

Several important classes in MGT are retracts of a product of trees

- **Median graphs**: retracts of hypercubes

- **Helly graphs**: retracts of strong products of paths

- **Absolute bipartite retracts** (including chordal bipartite graphs): retracts of relational products of paths

The odd case of retracts

$$H \text{ is a retract of } G \iff \text{there is a homomorphism } f : G \longmapsto H \text{ s.t. } f_{|H} = id_H$$

Implies that H is an isometric subgraph of G.

Several important classes in MGT are retracts of a product of trees

- **Median graphs**: retracts of hypercubes

  ECCENTRICITIES in $\tilde{\mathcal{O}}(n^{1.65})$ time **[Bergé & D. & Habib, *arXiv*]**

- **Helly graphs**: retracts of strong products of paths

  ECCENTRICITIES in $\mathcal{O}(m\sqrt{n})$ time **[Dragan & D. & Guarnera, WADS'21]**

- **Absolute bipartite retracts** (including chordal bipartite graphs): retracts of relational products of paths

  ECCENTRICITIES in $\tilde{\mathcal{O}}(m\sqrt{n})$ time **[D., WG'21]**

Short conclusion

- Some interesting connections exist between faster diameter computation and important geometric properties

(VC-dimension, tree embeddings, **Helly-type properties**)

- A large part of prior works in this area can be revisited and extended by using this unifying MGT-inspired framework.

La mulți ani, Victor!