Finding Small Subgraphs of Given Minimum Degree

Ignasi Sau-Valls Joint work with O. Amini, S. Perénnes, D. Peleg, and S. Saurabh

> Mascotte Project - INRIA/CNRS-I3S/UNSA - FRANCE Applied Mathematics IV Department of UPC - SPAIN

Outline of the talk

- Statement of the problem
- (Motivations)
- Preliminaries
- Hardness result
- (Approximation algorithm for minor free graphs)
- Conclusions and open problems

• MINIMUM SUBGRAPH OF MINIMUM DEGREE $\geq d$ (MSMD_d):

- For d = 2 it is the GIRTH problem (find the length of a shortest cycle), which is in P.
- We will see that for *d* ≥ 3, MSMD_d does not accept any constant-factor approximation (in particular, it is NP-complete).

• MINIMUM SUBGRAPH OF MINIMUM DEGREE $\geq d$ (MSMD_d):

- For d = 2 it is the GIRTH problem (find the length of a shortest cycle), which is in P.
- We will see that for *d* ≥ 3, MSMD_d does not accept any constant-factor approximation (in particular, it is NP-complete).

• MINIMUM SUBGRAPH OF MINIMUM DEGREE $\geq d$ (MSMD_d):

- For *d* = 2 it is the GIRTH problem (find the length of a shortest cycle), which is in P.
- We will see that for *d* ≥ 3, MSMD_d does not accept any constant-factor approximation (in particular, it is NP-complete).

• MINIMUM SUBGRAPH OF MINIMUM DEGREE $\geq d$ (MSMD_d):

- For d = 2 it is the GIRTH problem (find the length of a shortest cycle), which is in P.
- We will see that for *d* ≥ 3, MSMD_d does not accept any constant-factor approximation (in particular, it is NP-complete).

• Density $\rho(G)$ of a graph G = (V, E):

$$\rho(G) := \frac{|E(G)|}{|V(G)|}$$

More generally, for $S \subset V(G)$:

$$\rho(\boldsymbol{S}) := \rho(\boldsymbol{G}[\boldsymbol{S}])$$

• DENSE *k*-SUBGRAPH problem:

DENSE *k*-SUBGRAPH (D*k*S): **Input**: a graph G = (V, E) and a positive integer *k*. **Output**: a subset $S \subseteq V$ with |S| = k, maximizing $\rho(S)$.

(U. Feige, D. Peleg and G. Kortsarz, Algorithmica'01) (S. Khot, FOCS'04)

• Density $\rho(G)$ of a graph G = (V, E):

$$\rho(G) := rac{|E(G)|}{|V(G)|}$$

More generally, for $S \subset V(G)$:

$$\rho(S) := \rho(G[S])$$

• DENSE *k*-SUBGRAPH problem:

DENSE *k*-SUBGRAPH (D*k*S): **Input**: a graph G = (V, E) and a positive integer *k*. **Output**: a subset $S \subseteq V$ with |S| = k, maximizing $\rho(S)$.

(U. Feige, D. Peleg and G. Kortsarz, Algorithmica'01) (S. Khot, FOCS'04)

- Suppose that we want to find an induced subgraph *G*[*S*] of size at most *k* and density at least *ρ*.
 We can suppose that *S* is minimal, i.e. there is no subset of *S* with density greater than *ρ*(*S*).
 - 1) All the vertices of G[S] have degree at least $\rho/2$. If there exists a vertex v with degree strictly smaller than $\rho/2$, then the removal of v results in a smaller subgraph with higher density.
 - 2) If we have a subgraph G[S] with minimum degree at least ρ , then S is a subset with density at least $\rho/2$.
- So, modulo a constant factor, looking for a densest subgraph of G with size at most k is as hard as looking for the greatest ρ such that there exists a subgraph of G with size at most k and minimum degree at least ρ.

- Suppose that we want to find an induced subgraph *G*[*S*] of size at most *k* and density at least *ρ*.
 We can suppose that *S* is minimal, i.e. there is no subset of *S* with density greater than *ρ*(*S*).
 - All the vertices of G[S] have degree at least ρ/2. If there exists a vertex v with degree strictly smaller than ρ/2, then the removal of v results in a smaller subgraph with higher density.
 - 2) If we have a subgraph G[S] with minimum degree at least ρ , then S is a subset with density at least $\rho/2$.
- So, modulo a constant factor, looking for a densest subgraph of G with size at most k is as hard as looking for the greatest ρ such that there exists a subgraph of G with size at most k and minimum degree at least ρ.

- Suppose that we want to find an induced subgraph *G*[*S*] of size at most *k* and density at least *ρ*.
 We can suppose that *S* is minimal, i.e. there is no subset of *S* with density greater than *ρ*(*S*).
 - All the vertices of G[S] have degree at least ρ/2. If there exists a vertex v with degree strictly smaller than ρ/2, then the removal of v results in a smaller subgraph with higher density.
 - 2) If we have a subgraph G[S] with minimum degree at least ρ , then S is a subset with density at least $\rho/2$.
- So, modulo a constant factor, looking for a densest subgraph of G with size at most k is as hard as looking for the greatest ρ such that there exists a subgraph of G with size at most k and minimum degree at least ρ.

- Suppose that we want to find an induced subgraph *G*[*S*] of size at most *k* and density at least *ρ*.
 We can suppose that *S* is minimal, i.e. there is no subset of *S* with density greater than *ρ*(*S*).
 - All the vertices of G[S] have degree at least ρ/2. If there exists a vertex v with degree strictly smaller than ρ/2, then the removal of v results in a smaller subgraph with higher density.
 - 2) If we have a subgraph G[S] with minimum degree at least ρ , then S is a subset with density at least $\rho/2$.
- So, modulo a constant factor, looking for a densest subgraph of G with size at most k is as hard as looking for the greatest ρ such that there exists a subgraph of G with size at most k and minimum degree at least ρ.

Motivation 2: relation with TRAFFIC GROOMING

- Roughly, (one model of) TRAFFIC GROOMING consists in finding subgraphs with **high density** and *bounded* number of edges.
- I.e., we want to find subgraphs with **high average degree** (and *bounded* number of edges).
- Density and average degree of a graph differ by a factor 2.
- So, if we can find small subgraphs with prescribed minimum degree, we can also find small subgraphs with *good* density.

Motivation 2: relation with TRAFFIC GROOMING

- Roughly, (one model of) TRAFFIC GROOMING consists in finding subgraphs with **high density** and *bounded* number of edges.
- I.e., we want to find subgraphs with **high average degree** (and *bounded* number of edges).
- Density and average degree of a graph differ by a factor 2.
- So, if we can find small subgraphs with prescribed minimum degree, we can also find small subgraphs with *good* density.

Motivation 2: relation with TRAFFIC GROOMING

- Roughly, (one model of) TRAFFIC GROOMING consists in finding subgraphs with **high density** and *bounded* number of edges.
- I.e., we want to find subgraphs with **high average degree** (and *bounded* number of edges).
- Density and average degree of a graph differ by a factor 2.
- So, if we can find small subgraphs with prescribed minimum degree, we can also find small subgraphs with *good* density.

Hardness result

Preliminaries: hardness of approximation

Class APX (Approximable):

an NP-complete optimization problem is in APX if it can be approximated within a constant factor.

Example: VERTEX COVER

• Class PTAS (Polynomial-Time Approximation Scheme):

an NP-complete optimization problem is in PTAS if it can be approximated within a constant factor $1 + \varepsilon$, for all $\varepsilon > 0$ (the best one can hope for an NP-complete problem).

Example: MAXIMUM KNAPSACK

Preliminaries: hardness of approximation

Class APX (Approximable):

an NP-complete optimization problem is in APX if it can be approximated within a constant factor.

Example: VERTEX COVER

• Class PTAS (Polynomial-Time Approximation Scheme):

an NP-complete optimization problem is in PTAS if it can be approximated within a constant factor $1 + \varepsilon$, for all $\varepsilon > 0$ (the best one can hope for an NP-complete problem).

Example: MAXIMUM KNAPSACK

Idea of the proof for d = 3

(1) First we will see that $MSMD_3 \notin PTAS$.

(2) Then we will see that $MSMD_3 \notin APX$.

(1) $MSMD_3$ is not in PTAS

• Reduction from VERTEX COVER:

Instance H of VERTEX COVER \rightarrow Instance G of MSMD₃

• We will see that

PTAS for $G \Rightarrow$ PTAS for H

And so,

 \nexists PTAS for MSMD₃

(1) $MSMD_3$ is not in PTAS

• Reduction from VERTEX COVER:

```
Instance H of VERTEX COVER \rightarrow Instance G of MSMD<sub>3</sub>
```

We will see that

PTAS for $G \Rightarrow$ PTAS for H

And so,

 \nexists PTAS for MSMD₃

(1) $MSMD_3$ is not in PTAS

• Reduction from VERTEX COVER:

```
Instance H of VERTEX COVER \rightarrow Instance G of MSMD<sub>3</sub>
```

We will see that

PTAS for $G \Rightarrow$ PTAS for H

And so,

 \nexists PTAS for MSMD₃

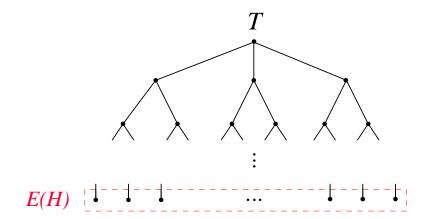
- (1) $MSMD_3$ is not in PTAS
 - Reduction from VERTEX COVER:
 - Instance H of VERTEX COVER \rightarrow Instance G of MSMD₃
 - We will see that

PTAS for $G \Rightarrow$ PTAS for H

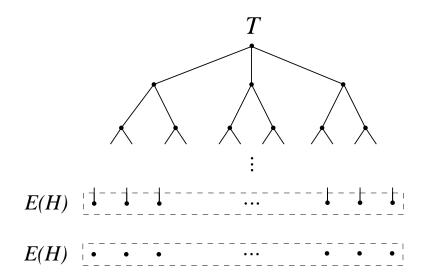
And so,

 \nexists PTAS for MSMD₃

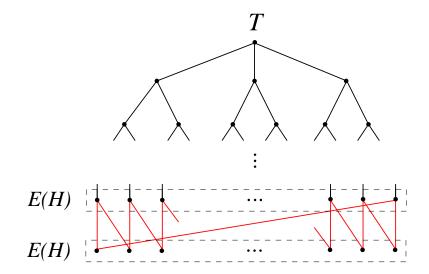
We build a complete ternary tree with $|E(H)| = 3 \cdot 2^m$ leaves:



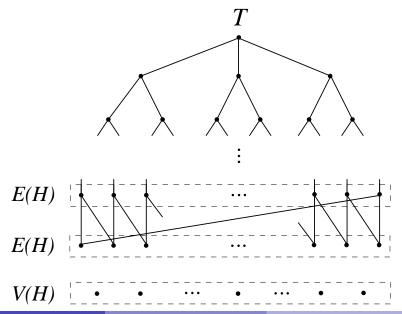
We add a copy of the set of leaves E(H):



We join both sets with a Hamiltonian cycle (for technical reasons):



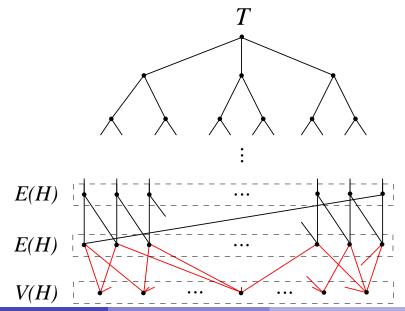
We add all the vertices of *H*:



Ignasi Sau-Valls (COST 293)

Degree-Constrained Subgraph Problems

We add the incidence relations between E(H) and $V(H) \rightarrow G$:



Ignasi Sau-Valls (COST 293)

- If we touch a vertex of G \ V(H), we have to touch all the vertices of G \ V(H)
- Thus, MSMD₃ in G is equivalent to minimize the number of selected vertices in V(H)
 - \rightarrow this is **exactly** VERTEX COVER in *H* !!
- Thus,

 $OPT_{MSMD_3}(G) = OPT_{VC}(H) + |V(G \setminus V(H))| =$ $= OPT_{VC}(H) + 9 \cdot 2^m$

• This clearly proves that

- If we touch a vertex of G \ V(H), we have to touch all the vertices of G \ V(H)
- Thus, MSMD₃ in G is equivalent to minimize the number of selected vertices in V(H)

 \rightarrow this is **exactly** VERTEX COVER in *H* !!

Thus,

 $OPT_{MSMD_3}(G) = OPT_{VC}(H) + |V(G \setminus V(H))| =$ $= OPT_{VC}(H) + 9 \cdot 2^m$

• This clearly proves that

- If we touch a vertex of G \ V(H), we have to touch all the vertices of G \ V(H)
- Thus, MSMD₃ in G is equivalent to minimize the number of selected vertices in V(H)
 - \rightarrow this is **exactly** VERTEX COVER in *H* !!

Thus,

$$OPT_{MSMD_3}(G) = OPT_{VC}(H) + |V(G \setminus V(H))| =$$
$$= OPT_{VC}(H) + 9 \cdot 2^m$$

• This clearly proves that

- If we touch a vertex of G \ V(H), we have to touch all the vertices of G \ V(H)
- Thus, MSMD₃ in G is equivalent to minimize the number of selected vertices in V(H)
 - \rightarrow this is **exactly** VERTEX COVER in *H* !!
- Thus,

$$OPT_{MSMD_3}(G) = OPT_{VC}(H) + |V(G \setminus V(H))| =$$
$$= OPT_{VC}(H) + 9 \cdot 2^m$$

• This clearly proves that

Let α > 1 be the factor of inapproximability of MSMD₃

• We use a technique called error amplification:

- ► We build a sequence of families of graphs G_k, such that MSMD₃ is hard to approximate in G_k within a factor α^k
- This proves that the problem is not in APX (for any constant C, ∃ k > 0 such that α^k > C)

• Let $G_1 = G$.

- Let $\alpha > 1$ be the factor of inapproximability of MSMD₃
- We use a technique called error amplification:
 - We build a sequence of families of graphs G_k, such that MSMD₃ is hard to approximate in G_k within a factor α^k
 - This proves that the problem is not in APX (for any constant C, ∃ k > 0 such that α^k > C)

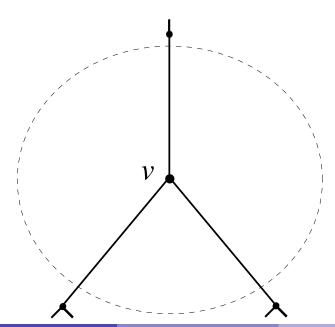
• Let $G_1 = G$.

- Let α > 1 be the factor of inapproximability of MSMD₃
- We use a technique called error amplification:
 - We build a sequence of families of graphs G_k, such that MSMD₃ is hard to approximate in G_k within a factor α^k
 - This proves that the problem is not in APX (for any constant C, ∃ k > 0 such that α^k > C)

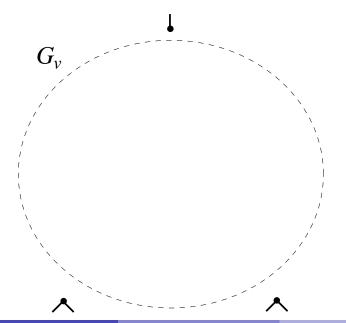
• Let $G_1 = G$.

- Let α > 1 be the factor of inapproximability of MSMD₃
- We use a technique called error amplification:
 - We build a sequence of families of graphs G_k, such that MSMD₃ is hard to approximate in G_k within a factor α^k
 - This proves that the problem is not in APX (for any constant C, ∃ k > 0 such that α^k > C)
- Let $G_1 = G$.

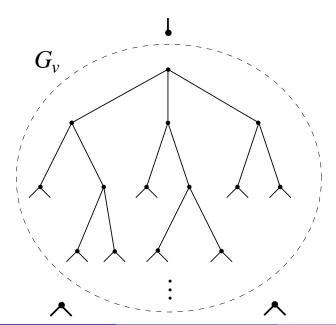
For any vertex v (note its degree by d_v):



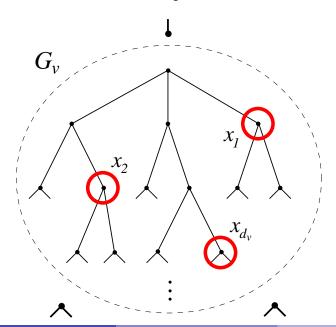
We will replace the vertex v with a graph G_v , built as follows:



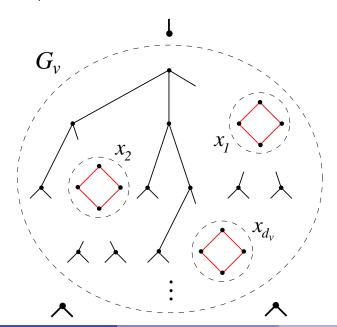
We begin by placing a copy of G (described before):



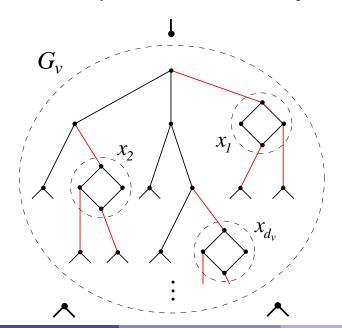
We select d_v vertices of degree 3 in $T \subset G$:



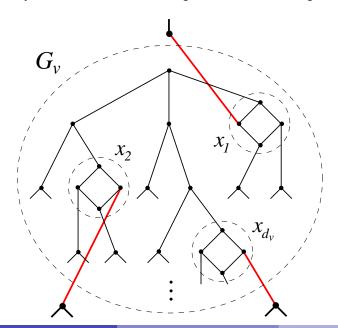
We replace each of these vertices x_i with a C_4 :



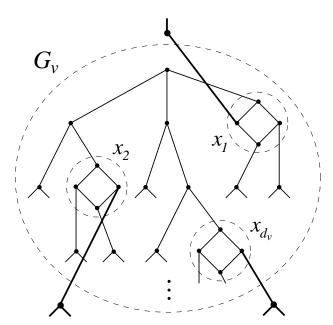
In each C_4 , we join 3 of the vertices to the neighbors of x_i :



We join the d_v vertices of degree 2 to the d_v neighbors of v:



This construction for all $v \in G$ defines G_2 :



Ignasi Sau-Valls (COST 293)

(2) $MSMD_3$ is not in APX

- Once a vertex in one G_ν is chosen → MSMD₃ in G_ν
 (which is hard up to a constant α)
- But minimize the number of *v*'s for which we touch $G_v \rightarrow MSMD_3$ in *G* (which is also hard up to a constant α)

- Thus, in G_2 the problem is hard to approximate up to a factor $\alpha \cdot \alpha = \alpha^2$
- Inductively we prove that in G_k the problem is hard to approximate up to a factor α^k

(2) $MSMD_3$ is not in APX

- Once a vertex in one G_ν is chosen → MSMD₃ in G_ν (which is hard up to a constant α)
- But minimize the number of *v*'s for which we touch $G_v \rightarrow MSMD_3$ in *G* (which is also hard up to a constant α)

- Thus, in G_2 the problem is hard to approximate up to a factor $\alpha \cdot \alpha = \alpha^2$
- Inductively we prove that in G_k the problem is hard to approximate up to a factor α^k

(2) $MSMD_3$ is not in APX

- Once a vertex in one G_ν is chosen → MSMD₃ in G_ν
 (which is hard up to a constant α)
- But minimize the number of *v*'s for which we touch $G_v \rightarrow MSMD_3$ in *G* (which is also hard up to a constant α)

- Thus, in G_2 the problem is hard to approximate up to a factor $\alpha \cdot \alpha = \alpha^2$
- Inductively we prove that in G_k the problem is hard to approximate up to a factor α^k

The problem is in P for graphs of *small* treewidth

Lemma

Let G be a graph on n vertices with treewidth at most t, and let d be a positive integer. Then in time $O((d + 1)^t (t + 1)^{d^2} n)$ we can either • find a smallest subgraph of minimum degree at least d in G, or

• conclude that no such subgraph exists.

Corollary

Let G be an n-vertex graph with treewidth $O(\log n)$, and let d be a positive integer. Then in polynomial time one can either • find a smallest subgraph of minimum degree at least d in G, or • conclude that no such subgraph exists.

The problem is in P for graphs of *small* treewidth

Lemma

Let G be a graph on n vertices with treewidth at most t, and let d be a positive integer. Then in time $O((d + 1)^t (t + 1)^{d^2} n)$ we can either • find a smallest subgraph of minimum degree at least d in G, or

conclude that no such subgraph exists.

Corollary

Let G be an n-vertex graph with treewidth $O(\log n)$, and let d be a positive integer. Then in polynomial time one can either

- find a smallest subgraph of minimum degree at least d in G, or
- conclude that no such subgraph exists.

Nice partition of *M*-minor-free graphs

Theorem

For a fixed graph *M*, there is a constant c_M such that for any integer $k \ge 1$ and for every *M*-minor-free graph *G*, the vertices of *G* can be partitioned into k + 1 sets such that any *k* of the sets induce a graph of treewidth at most $c_M k$.

Furthermore, such a partition can be found in polynomial time.

(E. Demaine, M.T. Hajiaghayi and K.C. Kawarabayashi, FOCS'05)

- Relying on the previous Theorem, partition V(G) in polynomial time into log n + 1 sets V₀,..., V_{log n} such that any log n of the sets induce a graph of treewidth at most c_M log n, where c_M is a constant depending only on the excluded graph M.
- (2) Run the dynamic programming algorithm of the Lemma on all the subgraphs $G_i = G[V \setminus V_i]$ of log *n* sets, $i = 0, ..., \log n$.
- (3) This procedure finds all the solutions of size at most log *n*.
- (4) If no solution is found, output the whole graph G.

This algorithm provides an $O(n/\log n)$ -approximation for MSMD_d in minor-free graphs, for all $d \ge 3$.

- (1) Relying on the previous Theorem, partition V(G) in polynomial time into $\log n + 1$ sets $V_0, \ldots, V_{\log n}$ such that any $\log n$ of the sets induce a graph of treewidth at most $c_M \log n$, where c_M is a constant depending only on the excluded graph M.
- (2) Run the dynamic programming algorithm of the Lemma on all the subgraphs $G_i = G[V \setminus V_i]$ of log *n* sets, $i = 0, ..., \log n$.
- (3) This procedure finds all the solutions of size at most log *n*.
- (4) If no solution is found, output the whole graph G.
 - This algorithm provides an $O(n/\log n)$ -approximation for MSMD_d in minor-free graphs, for all $d \ge 3$.

- Relying on the previous Theorem, partition V(G) in polynomial time into log n + 1 sets V₀,..., V_{log n} such that any log n of the sets induce a graph of treewidth at most c_M log n, where c_M is a constant depending only on the excluded graph M.
- (2) Run the dynamic programming algorithm of the Lemma on all the subgraphs $G_i = G[V \setminus V_i]$ of log *n* sets, $i = 0, ..., \log n$.
- (3) This procedure finds all the solutions of size at most log *n*.
- (4) If no solution is found, output the whole graph G.

This algorithm provides an $\mathcal{O}(n/\log n)$ -approximation for MSMD_d in minor-free graphs, for all $d \ge 3$.

- (1) Relying on the previous Theorem, partition V(G) in polynomial time into $\log n + 1$ sets $V_0, \ldots, V_{\log n}$ such that any $\log n$ of the sets induce a graph of treewidth at most $c_M \log n$, where c_M is a constant depending only on the excluded graph M.
- (2) Run the dynamic programming algorithm of the Lemma on all the subgraphs $G_i = G[V \setminus V_i]$ of log *n* sets, $i = 0, ..., \log n$.
- (3) This procedure finds all the solutions of size at most log *n*.
- (4) If no solution is found, output the whole graph G.

This algorithm provides an $O(n/\log n)$ -approximation for $MSMD_d$ in minor-free graphs, for all $d \ge 3$.

- (1) Relying on the previous Theorem, partition V(G) in polynomial time into $\log n + 1$ sets $V_0, \ldots, V_{\log n}$ such that any $\log n$ of the sets induce a graph of treewidth at most $c_M \log n$, where c_M is a constant depending only on the excluded graph M.
- (2) Run the dynamic programming algorithm of the Lemma on all the subgraphs $G_i = G[V \setminus V_i]$ of log *n* sets, $i = 0, ..., \log n$.
- (3) This procedure finds all the solutions of size at most log *n*.
- (4) If no solution is found, output the whole graph *G*.

This algorithm provides an $O(n/\log n)$ -approximation for $MSMD_d$ in minor-free graphs, for all $d \ge 3$.

- Relying on the previous Theorem, partition V(G) in polynomial time into log n + 1 sets V₀,..., V_{log n} such that any log n of the sets induce a graph of treewidth at most c_M log n, where c_M is a constant depending only on the excluded graph M.
- (2) Run the dynamic programming algorithm of the Lemma on all the subgraphs $G_i = G[V \setminus V_i]$ of log *n* sets, $i = 0, ..., \log n$.
- (3) This procedure finds all the solutions of size at most log *n*.
- (4) If no solution is found, output the whole graph *G*.

This algorithm provides an $\mathcal{O}(n/\log n)$ -approximation for MSMD_d in minor-free graphs, for all $d \geq 3$.

- We have proved that $MSMD_d$, $d \ge 3$, is not in APX.
- We have an $O(n/\log n)$ -approximation for minor free graphs.
- Parameterized version of MSMD_d:

Input: an undirected graph G = (V, E), an integer $d \ge 3$, and a parameter k.

Question: does there exist $S \subseteq V$, with $|S| \leq k$, such that $\delta(G[S]) \geq d$?

• $MSMD_d$, $d \ge 3$, is W[1]-hard.

(and thus the problem is not likely to be FPT in general graphs)

• FPT algorithms for minor free graphs.

(for instance: planar graphs, graphs of bounded local treewidth, graphs of bounded genus,...)

- We have proved that $MSMD_d$, $d \ge 3$, is not in APX.
- We have an $O(n/\log n)$ -approximation for minor free graphs.

• Parameterized version of MSMD_d:

Input: an undirected graph G = (V, E), an integer $d \ge 3$, and a parameter k.

Question: does there exist $S \subseteq V$, with $|S| \leq k$, such that $\delta(G[S]) \geq d$?

• $MSMD_d$, $d \ge 3$, is W[1]-hard.

(and thus the problem is not likely to be FPT in general graphs)

• FPT algorithms for minor free graphs.

(for instance: planar graphs, graphs of bounded local treewidth, graphs of bounded genus,...)

- We have proved that $MSMD_d$, $d \ge 3$, is not in APX.
- We have an $O(n/\log n)$ -approximation for minor free graphs.

• Parameterized version of MSMD_d:

Input: an undirected graph G = (V, E), an integer $d \ge 3$, and a parameter k.

Question: does there exist $S \subseteq V$, with $|S| \leq k$, such that $\delta(G[S]) \geq d$?

• $MSMD_d$, $d \ge 3$, is W[1]-hard.

(and thus the problem is not likely to be FPT in general graphs)

• FPT algorithms for minor free graphs.

(for instance: planar graphs, graphs of bounded local treewidth, graphs of bounded genus,...)

- We have proved that $MSMD_d$, $d \ge 3$, is not in APX.
- We have an $O(n/\log n)$ -approximation for minor free graphs.

• Parameterized version of MSMD_d:

Input: an undirected graph G = (V, E), an integer $d \ge 3$, and a parameter k.

Question: does there exist $S \subseteq V$, with $|S| \leq k$, such that $\delta(G[S]) \geq d$?

• $MSMD_d$, $d \ge 3$, is W[1]-hard.

(and thus the problem is not likely to be FPT in general graphs)

• FPT algorithms for minor free graphs.

(for instance: planar graphs, graphs of bounded local treewidth, graphs of bounded genus,...)

- We have proved that $MSMD_d$, $d \ge 3$, is not in APX.
- We have an $O(n/\log n)$ -approximation for minor free graphs.

• Parameterized version of MSMD_d:

Input: an undirected graph G = (V, E), an integer $d \ge 3$, and a parameter k.

Question: does there exist $S \subseteq V$, with $|S| \leq k$, such that $\delta(G[S]) \geq d$?

• $MSMD_d$, $d \ge 3$, is W[1]-hard.

(and thus the problem is not likely to be FPT in general graphs)

• FPT algorithms for minor free graphs.

(for instance: planar graphs, graphs of bounded local treewidth, graphs of bounded genus,...)

- We have proved that $MSMD_d$, $d \ge 3$, is not in APX.
- We have an $O(n/\log n)$ -approximation for minor free graphs.

• Parameterized version of MSMD_d:

Input: an undirected graph G = (V, E), an integer $d \ge 3$, and a parameter k.

Question: does there exist $S \subseteq V$, with $|S| \leq k$, such that $\delta(G[S]) \geq d$?

• $MSMD_d$, $d \ge 3$, is W[1]-hard.

(and thus the problem is not likely to be FPT in general graphs)

• FPT algorithms for minor free graphs.

(for instance: planar graphs, graphs of bounded local treewidth, graphs of bounded genus,...)

Thanks!