On the Generalized Mean Densest Subgraph Problem: Complexity and Algorithms

Karthekeyan Chandrasekaran
University of Illinois, Urbana-Champaign, USA

Chandra Chekuri
University of Illinois, Urbana-Champaign, USA

Manuel R. Torres
University of Illinois, Urbana-Champaign, USA

Weihao Zhu
University of Illinois, Urbana-Champaign, USA

Abstract

Dense subgraph discovery is an important problem in graph mining and network analysis with several applications. Two canonical polynomial-time solvable problems here are to find a *maxcore* (subgraph of maximum min degree) and to find a *densest subgraph* (subgraph of maximum average degree). Both of these problems can be solved in polynomial time. Veldt, Benson, and Kleinberg [47] introduced the generalized p-mean densest subgraph problem which captures the maxcore problem when $p = -\infty$ and the densest subgraph problem when $p = 1$. They observed that for $p \geq 1$, the objective function is supermodular and hence the problem can be solved in polynomial time. In this work, we focus on the p-mean densest subgraph problem for $p \in (-\infty, 1)$. We prove that for every $p \in (-\infty, 1)$, the problem is NP-hard, thus resolving an open question from [47]. We also show that for every $p \in (0, 1)$, the weighted version of the problem is APX-hard. On the algorithmic front, we describe two simple $\frac{1}{2}$-approximation algorithms for every $p \in (-\infty, 1)$. We complement the approximation algorithms by exhibiting non-trivial instances on which the algorithms simultaneously achieve an approximation factor of at most $\frac{1}{2}$.

2012 ACM Subject Classification Theory of computation → Approximation algorithms analysis

Keywords and phrases Densest subgraph problem, Hardness of approximation, Approximation algorithms

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2024.9

Category APPROX

Funding *Karthekeyan Chandrasekaran*: partially supported by NSF grant CCF-1907937 AND CCF-2402667.
Chandra Chekuri: partially supported by NSF grants CCF-1907937 and CCF-2402667.
Manuel R. Torres: supported in part by fellowships from NSF and the Sloan Foundation, and NSF grant CCF-1910149.
Weihao Zhu: supported by a graduate fellowship from the CS department.

Acknowledgements We thank Sanjeev Khanna and Euiwoong Lee for pointers to [23] and [26] on the hardness of Exact ℓ-Cover. We thank Farouk Harb for helpful discussions. This work was done when Manuel R. Torres was a student at University of Illinois, Urbana-Champaign.

1 Introduction

Dense subgraph discovery is an essential tool in graph mining and network analysis. The aim here is to find clusters in a graph which are denser than the entire graph. There are a number applications of dense subgraph discovery in biological settings [29, 20, 35, 4, 42], web mining [24, 14], social network analysis [34], real-time story identification [2], and finance

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2024).
Editors: Amit Kumar and Noga Ron-Zewi; Article No. 9; pp. 9:1–9:21
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

and fraud detection [15, 48, 31]. Motivated by the needs of applications and theoretical considerations, various density measures have been used and studied in the literature (see [18, 25, 37, 45, 36] for some surveys). Each density definition leads to a corresponding combinatorial optimization problem: given a graph G, find a subgraph of maximum density. Two of the most popular density measures in the literature are (i) the minimum degree of the subgraph and (ii) the average degree of the subgraph. These measures lead to the maxcore problem and densest subgraph problem respectively. They are both polynomial-time solvable and have been extensively studied. We briefly describe them before discussing a common generalization that will be the focus of this work.

In the maxcore problem (MAXCORE), the goal is to find a subgraph with maximum minimum degree. The optimum value of this problem is known as the *degeneracy* of the graph and the subgraph achieving the optimum is known as a *maxcore* of the graph. A k-*core* of a graph is a maximal connected subgraph in which all vertices have degree at least k. Min-degree is a popular measure of density, commonly finding use in what is known as the k-core decomposition, a nested sequence of subgraphs that captures k-cores for every k. One nice feature of a k-core decomposition is that there is a simple linear-time peeling algorithm to compute it. The peeling algorithm – denoted Greedy – produces an ordering of the vertices by repeatedly removing the vertex with least degree in the current graph. This ordering can in turn be used to solve MAXCORE. We refer the reader to [38] for a survey on k-core decomposition and applications.

In the densest subgraph problem (DSG), the goal is to find a subgraph of maximum average degree. DSG is widely used in graph mining applications. It is a well-studied problem in combinatorial optimization and is polynomial time solvable via a variety of techniques including network flow [40, 21], submodular function minimization (folklore), and linear programming [9]. Even though DSG can be solved exactly, the algorithms are slow and this has spurred the design of fast approximation algorithms [9, 5, 13, 8, 7, 10, 27]. Amongst these approximation algorithms is a peeling algorithm introduced by Asahiro, Iwama, Tamaki, and Tokuyama [3] which was shown to be a $\frac{1}{2}$-approximation by Charikar [9]. We note that the peeling order of this algorithm is the same as the one for computing a maxcore, namely Greedy; a second step of the algorithm returns the best subgraph induced by a suffix of the peeling order (best in terms of average degree). The specific density measure for DSG is used only in the second step. Charikar's analysis has spurred the development and analysis of a variety of peeling algorithms for several variants of DSG in both graphs and hypergraphs [1, 46, 44, 30, 33, 47].

Veldt, Benson, and Kleinberg [47] introduced the generalized mean densest subgraph problem that unifies MAXCORE and DSG. The input here is a real value $p \in \mathbb{R} \cup \{-\infty, \infty\}$ and an undirected graph $G = (V, E)$. For a subset $S \subseteq V$ of vertices, the density of the subgraph $G[S]$ induced by S is defined as:

$$M_p(S) := \left(\frac{1}{|S|} \sum_{v \in S} d_S(v)^p \right)^{1/p},$$

where $d_S(v)$ is the degree of vertex v in the subgraph $G[S]$. We note that $M_{-\infty}(S) = \min_{v \in S} d_S(v)$ is the minimum degree in the induced subgraph $G[S]$, while $M_\infty(S) = \max_{v \in S} d_S(v)$ is the maximum degree. For $p = 0$, the density of the subgraph $G[S]$ is $M_0(S) = (\prod_{v \in S} d_S(v))^{1/|S|} = \exp(\frac{1}{|S|} \sum_{v \in S} \ln d_S(v))$. The goal is to find a subset S of vertices with maximum $M_p(S)$. We refer to this problem as the p-mean densest subgraph problem (p-MEAN DSG). As p varies from $-\infty$ to ∞, $M_p(S)$ prioritizes the smallest degree in S to the largest degree in S and consequently, p-MEAN DSG provides a smooth way to generate subgraphs with different density properties.

p-MEAN DSG generalizes to weighted graphs in a natural manner. For a graph $G = (V, E)$ with positive edge weights $w : E \to \mathbb{R}_+$, we define $d_S(v)$ as the sum of the weight of edges that are incident to vertex v in $G[S]$. For a subset $S \subseteq V$ of vertices, its p-mean density $M_p(S)$ is defined using $d_S(v)$ as it was for the unweighted case. The goal again is to find a subset S of vertices with maximum $M_p(S)$. We refer to this problem as WEIGHTED p-MEAN DSG.

Veldt, Benson, and Kleinberg made several contributions to p-MEAN DSG. They observed that 1-MEAN DSG is equivalent to DSG and that $(-\infty)$-MEAN DSG is equivalent MAXCORE. For $p \geq 1$, they observed that the set function $f_p : 2^V \to \mathbb{R}_{\geq 0}$ defined by $f_p(S) := \sum_{v \in S} d_S(v)^p$ is a supermodular set function[1]. This implies that one can solve p-MEAN DSG in polynomial time for all $p \geq 1$ via a standard reduction to submodular set function minimization, a classical polynomial-time solvable problem in combinatorial optimization [41]. Motivated by the fact that exact algorithms are very slow in practice, they described a greedy peeling algorithm, denoted Greedy-p, that runs in $O(mn)$ time and achieves an approximation factor of $1/(p+1)^{1/p}$ for $p \geq 1$ (here m and n are the number of edges and nodes of the graph). The peeling order of their Greedy-p algorithm is *not* the same as that of Greedy – in particular, the peeling order depends on p. They supplement these theoretical results with empirical evaluation, showing that Greedy-p returns solutions with desirable characteristics for values of p in the range $[1, 2]$. We note that the function $f_p(S)$ is *not* supermodular if $p < 1$, which partially stems from the fact that the univariate function $g(x) := x^p$ is not convex if $p < 1$.

1.1 Our Results

We study the complexity and algorithmic status of p-MEAN DSG for $p \in (-\infty, 1)$ which was mentioned as a compelling direction for future work by Veldt et al. [47]. It is intriguing that p-MEAN DSG is polynomial-time solvable for for $p = -\infty$ and $p \geq 1$ while the status for $p \in (-\infty, 1)$ is non-trivial to understand. Our work fills this gap.

Hardness of p-mean DSG for $p \in (-\infty, 1)$

We prove that p-MEAN DSG is NP-Hard for every $p \in (-\infty, 1)$. We also show that WEIGHTED p-MEAN DSG is APX-hard for every fixed constant $p \in (0, 1)$. The hardness results are the main contribution of this work. They are technically involved for two reasons. First, the objective function is non-linear and does not fall into a clean and known class of functions. Second, the problem is effectively an unconstrained problem. The initial inspiration for our reduction came from a high-level connection to submodularity due to the concavity of the univariate function x^p for $p \in (0, 1)$; constrained versions of submodular optimization are NP-hard. However, the objective function for p-MEAN DSG is not a submodular function and there are no constraints. Nevertheless, we are able to model it via a gadget. Although the reduction is quite simple to describe, the proof of the reduction requires careful parameter setting and a detailed case analysis. The reduction/analysis for the weighted case is somewhat easier, however, we prove NP-Hardness for the unweighted case since it is of particular interest. We prove APX-hardness for the weighted case, and only for $p \in (0, 1)$, to mitigate the calculations. It may be possible to extend our APX-hardness proof to the unweighted case and also to the full range $(-\infty, 1)$.

[1] A real-valued set function $f : 2^V \to \mathbb{R}$ is *supermodular* if $f(A) + f(B) \leq f(A \cup B) + f(A \cap B)$ for all $A, B \subseteq V$. We recall that f is supermodular iff $-f$ is *submodular*.

Approximation algorithms for $p \in (-\infty, 1)$

The NP-Hardness result for p-MEAN DSG motivates the search for approximation algorithms for $p \in (-\infty, 1)$. We note that the peeling algorithm for p-MEAN DSG, namely Greedy-p, given by Veldt, Benson, and Kleinberg [47] is well-defined only for $p > 0$. In the same paper, the authors show empirical results for Greedy-p for $p \in (0, 1)$ even though the corresponding function f_p is not supermodular; however, no approximation guarantee is known for Greedy-p for $p \in (0, 1)$. We describe two different and simple algorithms for p-MEAN DSG– one based on simple greedy peeling and the other based on an exact solution to DSG. We show that both algorithms achieve a $\frac{1}{2}$-approximation for all $p \in (-\infty, 1)$. These are the first algorithms with approximation guarantees for p in the regime $(-\infty, 1)$. We complement the algorithms by exhibiting tight instances on which *both* algorithms exhibit an approximation factor of at most $\frac{1}{2}$, thus ruling out the possibility of improving the ratio by taking the best of the two algorithms.

This paper builds upon and extends an earlier version [11] by two of the authors. The previous version included results on faster algorithms for p-MEAN DSG for $p > 1$ and empirical evaluation of several algorithms. A full version including those results will be made available in the future.

Organization

We present the NP-hardness result in Section 2 and the APX-hardness result in Section 3. We present our algorithmic results in Section 4. All proofs that are omitted from the main body of the paper are given in the appendix.

Notation

Let $G = (V, E)$ be a graph. For a subset $S \subseteq V$ of vertices and a vertex $v \in S$, we recall that $d_S(v)$ is the (weighted) degree of v in the induced subgraph $G[S]$. Let S be a subset of vertices. We define $f_p(S) := \sum_{v \in S} d_S(v)^p$ if $p \in [-\infty, 0) \cup (0, \infty]$ and $f_p(S) := \sum_{v \in S} \ln d_S(v)$ if $p = 0$. We also define $\rho_p(S) := f_p(S)/|S|$ for all p. With these definitions, we have $M_p(S) = \rho_p(S)^{1/p}$ for all p. Thus, finding a set S of vertices with maximum $M_p(S)$ is equivalent to finding a set S of vertices with *maximum* $\rho_p(S)$ if $p \geq 0$, and to finding a set S of vertices with *minimum* $\rho_p(S)$ if $p < 0$.

1.2 Other Related Work

DSG and the subfield of dense subgraph discovery is large. We point the reader to a recent survey [36] and restrict our attention to discussing some closely related work, specifically on sequential models and approximability.

As we remarked, DSG is poly-time solvable by several techniques including via maximum flow. Although maximum flow now admits an almost-linear time algorithm [12], the existing exact algorithms for DSG are slow in practice for large graphs. Thus approximation algorithms have also been considered (especially before the recent developments on network flow). In particular, there has been a line of work that obtained a $(1 - \epsilon)$-approximation in $\tilde{O}(m \cdot \mathsf{poly}(\frac{1}{\epsilon}))$-time [5, 8, 10]; in particular, the algorithm in [10] runs in time $\tilde{O}(\frac{m}{\epsilon})$. These faster approximation algorithms also have some limitations in practice for large graphs. Several simpler iterative algorithms based on continuous optimization methods have been developed – these include algorithms based on the Frank-Wolfe method [13], an algorithm based on iterating Greedy called Greedy++ [7], and the projected gradient descent method [27].

They are simple to implement and have been shown to converge quickly to near-optimal solutions on large graphs, both synthetic and real-world, even though the known worst-case theoretical convergence rates are fairly large.

We discuss some other measures of density considered in the literature. Given a graph $G = (V, E)$ and a finite collection of pattern graphs $\mathcal{F}$, one can consider the problem of finding a set S that maximizes $f(S)/|S|$ where $f : 2^V \to \mathbb{Z}_+$ counts the number of occurrences of the patterns from $\mathcal{F}$ in the induced subgraph $G[S]$. If we consider a single edge as the pattern, then we obtain DSG. [46] considered the special case where $\mathcal{F}$ is a single triangle graph and [44] considered the special case where $\mathcal{F}$ is a clique on k vertices. The densest subgraph problem under the general notion of patterns was considered in [17]. Densest subgraph has also been studied for hypergraphs with density of a set S of vertices being defined as $|E(S)|/|S|$ where $E(S)$ is the set of hyperedges with all vertices in S [30]. We can reduce the densest subgraph problem with pattern based densities to the densest subgraph problem in hypergraphs by introducing a hyperedge for each occurrence of the pattern in the input graph. Charikar's analysis of Greedy can be generalized to show an approximation factor of at least $\frac{1}{r}$ in rank r-hypergraphs. Veldt, Benson, and Kleinberg [47] showed that Greedy is not a good worst-case algorithm for p-mean DSG when $p > 1$, and as we mentioned earlier, they developed a different peeling algorithm. Chekuri, Quanrud and Torres [10] unified several results by considering density measures of the form $f(S)/|S|$ where $f : 2^V \to \mathbb{R}_+$ is an arbitrary non-negative supermodular set function over a vertex set V. They showed that there is a natural peeling algorithm for each f and derived an approximation bound in terms of certain properties of f; Greedy-p from [47] and its approximation bound are derived as special cases. They also generalized Greedy++ to supermodular densities and showed that the resulting algorithm converges to an optimum solution, partially answering a conjecture from [7] (the conjecture has a strong convergence rate). See [27, 28] for additional insights. One can also consider density measures of the form $f(S)/g(S)$ where $g : 2^V \to \mathbb{R}_+$ is another set function such as a concave or convex function of $|S|$. We refer the reader to [36, 10] for results and pointers on this aspect.

Constrained versions of DSG such as the k-densest-subgraph (find a densest subgraph with at most k vertices) are well-studied in theoretical computer science. k-densest-subgraph is of particular importance due to its connection to various other problems, and due to the intriguing difficulty in understanding its approximability. Since there is a large literature on this problem and since constrained versions are not the focus of this paper, we point the reader to some relevant papers on algorithms and hardness [19, 6, 39].

2 NP-hardness

In this section, we prove the following theorem.

▶ **Theorem 1.** *p-MEAN DSG is NP-hard for all $p \in (-\infty, 1)$.*

We reduce from the Exact ℓ-Cover problem.

▶ **Problem 1.** *EXACT ℓ-COVER: the input is a finite ground set $\mathcal{U} = \{e_1, e_2, \dots, e_{\ell n}\}$ of cardinality ℓn for some positive integers ℓ and n, and a family of subsets $\mathcal{S} \subseteq 2^{\mathcal{U}}$, where each $X \in \mathcal{S}$ has cardinality ℓ. The goal is to determine whether there exist n disjoint sets $S_{i_1}, S_{i_2}, \dots, S_{i_n} \in \mathcal{S}$ whose union is $\mathcal{U}$.*

We will say that the input instance $(\mathcal{U}, \mathcal{S})$ of EXACT ℓ-COVER has an exact ℓ-cover if there exist n disjoint sets whose union is $\mathcal{U}$. EXACT 3-COVER is a well-known NP-complete problem [22]. A standard padding approach reduces EXACT 3-COVER to EXACT ℓ-COVER for every $\ell \geq 3$.

▶ **Theorem 2.** *EXACT ℓ-COVER is* NP*-complete for every integer $\ell \geq 3$.*

2.1 Reduction from exact ℓ-cover

We reduce EXACT ℓ-COVER to p-MEAN DSG. Let $\mathcal{U} = \{e_1, e_2, \ldots, e_{\ell n}\}$ and $\mathcal{S} = \{S_1, S_2, \ldots, S_m\}$ with $S_i \subseteq \mathcal{U}$ for every $i \in [m]$ be the input instance of EXACT ℓ-COVER. For a positive integer d (to be chosen later), we construct a graph $G_d = (L \cup A, E)$, where $L := \{v_i : i \in [m]\}$ and $A := \{u_j : j \in [\ell \cdot n]\}$. For every $1 \leq i \leq m$ and $1 \leq j \leq \ell \cdot n$, if the set S_i contains element e_j, then we add an edge between v_i and u_j in graph G_d. Further, we add edges between vertices in A to make $G_d[A]$ to be a connected d-regular graph, where d will be chosen appropriately (we may assume that n is even so that such a connected d-regular graph always exists). See Figure 1. If $p \neq 0$, then we set $\rho^* := \frac{\ell^p + \ell \cdot (d+1)^p}{\ell+1}$ and if $p = 0$, then we set $\rho^* := \frac{\ln \ell + \ell \cdot \ln(d+1)}{\ell+1}$. We will show that there exist positive integers $\ell \geq 3$ and d such that the input instance admits an exact ℓ-cover if and only if $\max\{M_p(X) : X \subseteq V\} \geq (\rho^*)^{1/p}$.

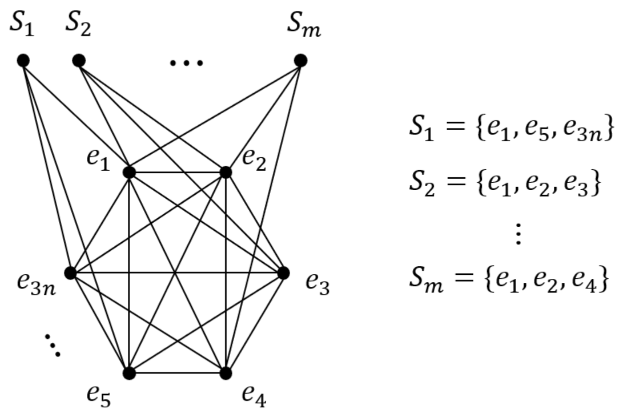

Figure 1 Graph constructed in our reduction from EXACT ℓ-COVER for $\ell = 3$ and $d = 5$. The EXACT 3-COVER instance consists of the ground set $\mathcal{U} := \{e_1, \ldots, e_{3n}\}$ and the family $\mathcal{S} := \{S_1, \ldots, S_m\}$.

Next, we prove the NP-hardness of p-MEAN DSG by casing on the value of p via the above mentioned reduction. We prove NP-hardness for $p \in (0, 1)$ in Section 2.2 (see Theorem 7). The missing proofs are given in the appendix. The proofs for $p \in (-\infty, 0]$ are given in the full version owing to space limitations.

2.2 NP-hardness for $p \in (0, 1)$

We recall that $\rho_p(X) = M_p(X)^p$ for every subset X of vertices and hence, finding a set X of vertices that maximizes $M_p(X)$ reduces to finding a set X of vertices that maximizes $\rho_p(X)$ if $p \in (0, 1)$. We define $\mathsf{OPT}_{G_d} := \max_{X \subseteq V} \rho_p(X)$. We observe that OPT_{G_d} is a maximization problem for $p \geq 0$. Hence, in order to show correctness of our reduction, we will need an upper bound on $\rho_p(X)$ for $p \geq 0$. The following lemma gives an upper bound.

▶ **Lemma 3.** *Let $G_d = (L \cup A, E)$ be the graph constructed in the reduction. Let $S \subseteq L$ and $A' \subseteq A$. Then, for $p > 0$, we have that*

$$\rho_p(S \cup A') \leq \frac{\ell^p \cdot |S| + \sum_{v \in A'} (d + d_{S+v}(v))^p}{|S| + |A'|}.$$

Moreover, the inequality above is strict if $|A'| < |A|$.

Proof. For the case of $p > 0$, we note that

$$\begin{aligned}
\rho_p(S \cup A') &= \frac{f_p(S \cup A')}{|S \cup A'|} \\
&= \frac{\sum_{u \in S} d_{S \cup A'}(u)^p + \sum_{v \in A'} d_{S \cup A'}(v)^p}{|S| + |A'|} \\
&\leq \frac{\sum_{u \in S} d_{S \cup A}(u)^p + \sum_{v \in A'} d_{S \cup A}(v)^p}{|S| + |A'|} \quad \text{(since } p > 0\text{)} \\
&= \frac{\ell^p \cdot |S| + \sum_{v \in A'} d_{S \cup A}(v)^p}{|S| + |A'|} \quad \text{(since } d_{S \cup A}(u) = \ell\text{)} \\
&= \frac{\ell^p \cdot |S| + \sum_{v \in A'} (d + d_{S+v}(v))^p}{|S| + |A'|}. \quad \text{(since } G[A] \text{ is a } d\text{-regular graph)}
\end{aligned}$$

If $|A'| < |A|$, then there exists a vertex $u \in A'$ such that $d_{A'}(u) < d_A(u) = d$ because $G[A]$ is connected, which implies that the inequality above is strict. ◀

We need the following technical lemma about the maximizer of a relevant function.

▶ **Lemma 4.** *Let $c \in \mathbb{R}_{\geq 0}$. Let $f : \mathbb{R}^n_{\geq 0} \to \mathbb{R}$ be defined as $f(x) := \sum_{i=1}^n (c + x_i)^p$. For $0 < p < 1$, consider the following maximization problem parameterized by $s \in \mathbb{N}$:*

$$\textit{maximize} \quad \left\{ f(x) : x \in \mathbb{Z}^n, x \geq 0, \sum_{i=1}^n x_i \leq s \right\}.$$

Every maximizer for the above problem has μ coordinates set to $\lceil s/n \rceil - 1$ and $n - \mu$ coordinates set to $\lceil s/n \rceil$, where $\mu = n \cdot \lceil s/n \rceil - s$. If s is a multiple of n, then the maximizer has all coordinates set to s/n.

Proof. Let $x \in \mathbb{Z}^n$ and $x \geq 0$ be a maximizer. If $\sum_{i=1}^n x_i < s$, then by setting $x'_n = x_n + 1$ and $x'_k = x_k$ for every $k \in [n-1]$, we have $f(x') > f(x)$, a contradiction to optimality of $f(x)$.

Suppose that $\sum_{i=1}^n x_i = s$. We prove that all coordinates are in $\{\lceil s/n \rceil, \lceil s/n \rceil - 1\}$. Assume that x has at least one entry not in $\{\lceil s/n \rceil, \lceil s/n \rceil - 1\}$. Then, there exists some coordinate x_i that is strictly larger than $\lceil s/n \rceil$ or smaller than $\lceil s/n \rceil - 1$. Without loss of generality, we assume that $x_i > \lceil s/n \rceil$. Consequently, there exists some index $j \in [n]$ such that $x_j < \lceil s/n \rceil$. Since $x_j < \lceil s/n \rceil < x_i$ and x_i, x_j are both integers, we have $x_i - x_j \geq 2$. Because $(c+x)^p$ is a concave function of x, $(c+x_i)^p + (c+x_j)^p < (c+x_i-1)^p + (c+x_j+1)^p$. By setting $x'_i = x_i - 1$, $x'_j = x_j + 1$, and $x'_k = x_k$ for every $k \in [n] \backslash \{i, j\}$, we have $f(x') > f(x)$, a contradiction to optimality of $f(x)$.

Hence, for every maximizer $x \in \mathbb{Z}^n$, we have $\sum_{i=1}^n x_i = s$ and $x_i \in \{\lceil s/n \rceil, \lceil s/n \rceil - 1\}$ for every $i \in [n]$. This implies that $(n - \mu)$ coordinates are $\lceil s/n \rceil$ and μ coordinates are $\lceil s/n \rceil - 1$, where $\mu = n \cdot \lceil s/n \rceil - s$. When s is a multiple of n, by Jensen's inequality, we have $x_i = s/n$ for every $1 \leq i \leq n$. ◀

We will use the following lemma about the existence of an integer ℓ that satisfies two inequalities simultaneously for every given p.

▶ **Lemma 5.** *For every $p \in (0, 1)$, there exists an integer $\ell \geq 3$ s.t the following two inequalities hold:*

$$\left(1 - \frac{1}{2\ell}\right)^p < 1 - \frac{1 - 1/2^p}{\ell + 1} \quad \textit{and} \quad \left(1 + \frac{1}{2\ell}\right)^p < 1 + \frac{1 - 1/2^p}{\ell + 1}. \tag{1}$$

We need the following lemma about NO-instances of EXACT ℓ-COVER.

▶ **Lemma 6.** *Let $p \in (0,1)$ and $\ell \geq 3$ be an integer that satisfies the two inequalities in (1). Consider an instance of EXACT ℓ-COVER with ground set $\mathcal{U}$ of size ℓn and family $\mathcal{S} \subseteq 2^{\mathcal{U}}$. Suppose that the instance has no exact ℓ-cover. Then, for every non-negative integers $s, a' \in \mathbb{Z}_{\geq 0}$ with $s \leq |\mathcal{S}|$, $a' \leq \ell n$, and $s + a' \geq 1$ and every non-negative integer vector $X \in \mathbb{Z}_{\geq 0}^{a'}$ with $\sum_{i=1}^{a'} X_i \leq \ell s$, we have that*

$$\frac{\ell^p \cdot s + \sum_{i=1}^{a'}(2\ell - 1 + X_i)^p}{s + a'} \leq \frac{\ell^p + \ell \cdot (2\ell)^p}{\ell + 1}. \tag{2}$$

Moreover, if there exists $i \in [a']$ such that $X_i \neq \ell s/a'$, then the above inequality is strict.

Proof. We case on the value of the ratio s/a'.

Case 1. Suppose $\ell \cdot s = a'$. Then, we have

$$\begin{aligned}
\frac{\ell^p \cdot s + \sum_{i=1}^{a'}(2\ell - 1 + X_i)^p}{s + a'} &\leq \frac{\ell^p \cdot s + a' \cdot (2\ell - 1 + \frac{\ell \cdot s}{a'})^p}{s + a'} \quad \left(\text{by } \sum_{i=1}^{a'} X_i \leq \ell \cdot s \text{ and Lemma 4}\right) \\
&= \frac{\ell^p \cdot a'/\ell + a' \cdot (2\ell)^p}{a'/\ell + a'} \quad (\text{since } \ell \cdot s = a') \\
&= \frac{\ell^p + \ell \cdot (2\ell)^p}{\ell + 1}.
\end{aligned}$$

By Lemma 4, if there exists $i \in [a']$ such that $X_i \neq \ell s/a'$, then the above inequality is strict.

Case 2. Suppose $\ell \cdot s = \beta \cdot a'$ for some $0 \leq \beta < 1$. Then, we have

$$\begin{aligned}
\frac{\ell^p \cdot s + \sum_{i=1}^{a'}(2\ell - 1 + X_i)^p}{s + a'} &\leq \frac{\ell^p \cdot s + \beta a' \cdot (2\ell)^p + (1-\beta)a' \cdot (2\ell - 1)^p}{s + a'} \\
&\qquad \left(\text{by } \sum_{i=1}^{a'} X_i \leq \ell \cdot s \text{ and Lemma 4}\right) \\
&= \frac{\ell^p \cdot \beta + \ell\beta \cdot (2\ell)^p + \ell(1-\beta) \cdot (2\ell - 1)^p}{\beta + \ell}. \quad (\text{since } \ell \cdot s = \beta \cdot a')
\end{aligned}$$

We define $h : [0,1] \to \mathbb{R}$ as

$$h(\beta) := \frac{\ell^p \cdot \beta + \ell\beta \cdot (2\ell)^p + \ell(1-\beta) \cdot (2\ell - 1)^p}{\beta + \ell},$$

which implies that the left hand side expression in the lemma is at most $h(\beta)$. By differentiating the function h with respect to β, we have

$$\begin{aligned}
\frac{d}{d\beta} h(\beta) &= \frac{(\ell^p + \ell \cdot (2\ell)^p - \ell \cdot (2\ell - 1)^p)(\beta + \ell) - (\ell^p \cdot \beta + \ell\beta \cdot (2\ell)^p + \ell(1-\beta) \cdot (2\ell - 1)^p)}{(\beta + \ell)^2} \\
&= \frac{\ell^{p+1} + \ell^2 \cdot (2\ell)^p - \ell^2 \cdot (2\ell - 1)^p - \ell \cdot (2\ell - 1)^p}{(\beta + \ell)^2} \\
&= \frac{\ell^{p+1} + \ell^2 \cdot (2\ell)^p - (\ell^2 + \ell) \cdot (2\ell - 1)^p}{(\beta + \ell)^2} \\
&> \frac{\ell^{p+1} + \ell^2 \cdot (2\ell)^p - (\ell^2 + \ell) \cdot (2\ell)^p \cdot (1 - \frac{1 - 2^{-p}}{\ell + 1})}{(\beta + \ell)^2} \quad (\text{by inequality (1)}) \\
&= \frac{\ell^{p+1} + \ell^2 \cdot (2\ell)^p - \ell \cdot (2\ell)^p \cdot (\ell + 2^{-p})}{(\beta + \ell)^2} = 0.
\end{aligned}$$

Hence, function $h(\beta)$ is strictly increasing for $\beta \in [0,1]$. Thus,

$$\frac{\ell^p \cdot s + \sum_{i=1}^{a'}(2\ell - 1 + X_i)^p}{s + a'} \le h(\beta) < h(1) = \frac{\ell^p + \ell \cdot (2\ell)^p}{\ell + 1}.$$

Case 3. Suppose $\ell \cdot s = \alpha \cdot a'$ for some $\alpha > 1$. Let $\alpha = t + \beta$ where $t \ge 1$ is an integer and $0 < \beta \le 1$. Then, we have

$$\begin{aligned}\frac{\ell^p \cdot s + \sum_{i=1}^{a'}(2\ell - 1 + X_i)^p}{s + a'} &\le \frac{\ell^p \cdot s + \beta a' \cdot (2\ell + t)^p + (1-\beta)|A'| \cdot (2\ell - 1 + t)^p}{|S| + |A'|} \\ &\quad \text{(by } \sum_{i=1}^{a'} X_i \le \ell \cdot s \text{ and Lemma 4)} \\ &= \frac{\ell^p \cdot (t+\beta) + \ell\beta \cdot (2\ell + t)^p + \ell(1-\beta) \cdot (2\ell - 1 + t)^p}{t + \beta + \ell} \\ &\quad \text{(since } \ell \cdot s = (t+\beta) \cdot a'\text{).}\end{aligned}$$

We define $g : [0, +\infty) \times [0,1] \to \mathbb{R}$ as

$$g(t,\beta) := \frac{\ell^p \cdot (t+\beta) + \ell\beta \cdot (2\ell + t)^p + \ell(1-\beta) \cdot (2\ell - 1 + t)^p}{t + \beta + \ell},$$

which implies that the left hand side expression in the lemma is at most $g(t,\beta)$. We note that $g(t,1) = g(t+1,0)$ for every $t \ge 0$.

By differentiating the function g with respect to β, we have

$$\begin{aligned}&\frac{d}{d\beta} g(t,\beta) \\ &= \frac{(\ell^p + \ell \cdot (2\ell+t)^p - \ell \cdot (2\ell - 1 + t)^p)(t+\beta+\ell) - (\ell^p \cdot (t+\beta) + \ell\beta \cdot (2\ell+t)^p + \ell(1-\beta) \cdot (2\ell - 1 + t)^p)}{(t+\beta+\ell)^2} \\ &= \frac{\ell^{p+1} + (\ell^2 + t\cdot\ell)\cdot(2\ell+t)^p - (\ell^2 + (t+1)\cdot\ell)\cdot(2\ell+t-1)^p}{(t+\beta+\ell)^2}.\end{aligned}$$

Now, we note that

$$\begin{aligned}&\frac{d}{dt}\left(\ell^{p+1} + (\ell^2 + t\cdot\ell)\cdot(2\ell+t)^p - (\ell^2 + (t+1)\cdot\ell)\cdot(2\ell+t-1)^p\right) \\ &= p\cdot(\ell^2 + t\cdot\ell)\cdot(2\ell+t)^{p-1} + \ell\cdot(2\ell+t)^p - p\cdot(\ell^2 + (t+1)\cdot\ell)\cdot(2\ell+t-1)^{p-1} - \ell\cdot(2\ell+t-1)^p \\ &\le p\cdot(\ell^2 + t\cdot\ell)\cdot(2\ell+t)^{p-1} + \ell\cdot p\cdot(2\ell+t-1)^{p-1} - p\cdot(\ell^2 + (t+1)\cdot\ell)\cdot(2\ell+t-1)^{p-1} \\ &\qquad \text{(since } (x+1)^p - x^p \le p\cdot x^{p-1} \text{ for every } x > 0\text{)} \\ &= p\cdot(\ell^2 + t\cdot\ell)((2\ell+t)^{p-1} - (2\ell+t-1)^{p-1}) < 0.\end{aligned}$$

Thus,

$$\begin{aligned}\frac{d}{d\beta} g(t,\beta) &= \frac{\ell^{p+1} + (\ell^2 + t\cdot\ell)\cdot(2\ell+t)^p - (\ell^2 + (t+1)\cdot\ell)\cdot(2\ell+t-1)^p}{(t+\beta+\ell)^2} \\ &\le \frac{\ell^{p+1} + (\ell^2 + \ell)\cdot(2\ell+1)^p - (\ell^2 + 2\ell)\cdot(2\ell)^p}{(1+\beta+\ell)^2} \quad \text{(since } t \ge 1\text{)} \\ &= \frac{(2\ell)^p\cdot(2^{-p}\cdot\ell + (\ell^2+\ell)\cdot(1+\frac{1}{2\ell})^p - (\ell^2+2\ell))}{(1+\beta+\ell)^2} \\ &< \frac{(2\ell)^p\cdot(2^{-p}\cdot\ell + (\ell^2+\ell)\cdot(1+\frac{1-2^{-p}}{\ell+1}) - (\ell^2+2\ell))}{(1+\beta+\ell)^2} \quad \text{(by inequality (1))} \\ &= \frac{(2\ell)^p\cdot(2^{-p}\cdot\ell + (\ell^2+\ell) + (1-2^{-p})\cdot\ell - (\ell^2+2\ell))}{(1+\beta+\ell)^2} = 0.\end{aligned}$$

Hence, function $g(t,\beta)$ is strictly decreasing with respect to β for $\beta \in [0,1]$. Thus,

$$\frac{\ell^p \cdot s + \sum_{i=1}^{a'}(2\ell - 1 + X_i)^p}{s + a'} \leq g(t,\beta) < g(t,0).$$

For every positive integer $r > 1$, we have $g(r,0) = g(r-1,1) < g(r-1,0)$. Thus,

$$\frac{\ell^p \cdot s + \sum_{i=1}^{a'}(2\ell - 1 + X_i)^p}{s + a'} < g(t,0) < g(1,0) = \frac{\ell^p + \ell \cdot (2\ell)^p}{\ell + 1}.$$ ◀

Now, we are ready to prove the NP-hardness of p-MEAN DSG for $p \in (0,1)$. We recall that $\mathsf{OPT}_G := \max_{X \subseteq V} \rho_p(X)$.

▶ **Theorem 7.** *For every $p \in (0,1)$, there exist positive integers $\ell \geq 3$ and d such that for an instance of* EXACT *ℓ-*COVER *with ground set $\mathcal{U}$ and family $\mathcal{S} \subseteq 2^{\mathcal{U}}$, there exists an exact ℓ-cover iff $OPT_G \geq \rho^* = \frac{\ell^p + \ell \cdot (d+1)^p}{\ell + 1}$, where $G = G_d$ is the graph constructed in the reduction above.*

Proof. By Corollary 5, we know that there exists a positive integer $\ell \geq 3$ satisfying the two inequalities in (1). Fix such an ℓ. Let $\mathcal{U}$ be the ground set and $\mathcal{S}$ be the collection of subsets of an instance of Exact ℓ-Cover. We set $d = 2\ell - 1$ and consider the graph $G = G_d$ constructed in the reduction in Section 2.1. We show that there exists an exact ℓ-cover iff $\mathsf{OPT}_G \geq \rho^* = \frac{\ell^p + \ell \cdot (d+1)^p}{\ell + 1}$.

Suppose $\mathcal{S}$ contains an exact ℓ-cover $S_{i_1}, \ldots, S_{i_n}$. Let $S = \{v_{i_1}, v_{i_2}, \ldots, v_{i_n}\}$. We note that $|S| = n = \frac{|A|}{\ell}$. Thus, we have

$$\mathsf{OPT}_G \geq \rho_p(S \cup A) = \frac{f_p(S \cup A)}{|S \cup A|} = \frac{\ell^p \cdot |S| + (d+1)^p \cdot |A|}{|S| + |A|} = \frac{\ell^p \cdot |A|/\ell + (d+1)^p \cdot |A|}{|A|/\ell + |A|} = \rho^*.$$

Suppose $\mathcal{S}$ does not contain an exact ℓ-cover. Let $S \subseteq L$ and $A' \subseteq A$. We note that $\ell \cdot |S| = \sum_{v \in A} d_{S+v}(v) \geq \sum_{v \in A'} d_{S+v}(v)$. Thus, we have

$$\rho_p(S \cup A') = \frac{f_p(S \cup A')}{|S \cup A'|} \leq \frac{\ell^p \cdot |S| + \sum_{v \in A'}(d + d_{S+v}(v))^p}{|S| + |A'|} \quad \text{(by Lemma 3)}$$
$$\leq \rho^*. \quad \text{(by Lemma 6)}$$

We note that since S is not an exact ℓ-cover, we have that either $|A'| < |A|$ or there exists $u, v \in A$ with $d_{S+u}(u) \neq d_{S+v}(v)$. This implies that either the first inequality or the second inequality above is strict according to the respective lemmas, that is, $\rho_p(S \cup A') < \rho^*$. ◀

3 APX-hardness for $p \in (0,1)$

In this section, we adapt the NP-hardness proof from Section 2 to show that WEIGHTED p-MEAN DSG is APX-hard for every fixed constant $p \in (0,1)$.

▶ **Theorem 8.** *For every fixed constant $p \in (0,1)$, there exists a constant $\delta_p > 0$ that depends only on p such that it is* NP*-hard to obtain a $(1 - \delta_p)$-approximation for* WEIGHTED *p-*MEAN *DSG.*

In order to prove Theorem 8, we will rely on the APX-hardness of EXACT ℓ-COVER as stated below.

▶ **Theorem 9.** *There exists a constant $\varepsilon \in (0,1)$ such that for every integer $\ell \geq 3$, it is NP-hard to distinguish between the following two cases for a given finite ground set $\mathcal{U}$ of size ℓn and a family $\mathcal{S} \subseteq 2^{\mathcal{U}}$ of subsets each of cardinality ℓ:*

- *YES-instance: There exists a collection of n sets in $\mathcal{S}$ whose union is $\mathcal{U}$.*
- *NO-instance: The union of every collection of $n \cdot (1+\varepsilon)$ sets in $\mathcal{S}$ has size at most $\ell \cdot n \cdot (1-\varepsilon)$.*

Exact ℓ-Cover is a special case of Set Cover. The hardness we seek requires a disjoint set cover in the YES case, and we also need the hardness to hold for every fixed integer $\ell \geq 3$. Related results have been proved in the literature [32, 43, 26, 23], however the precise version we need requires a formal argument. We provide the proof in the full version and also comment on the relation to previous work.

Reduction from Exact ℓ-Cover to weighted p-mean DSG

We reduce from the APX-hard variant of EXACT ℓ-COVER, namely the problem mentioned in Theorem 9. Consider an instance of the problem mentioned in Theorem 9: namely, let $\mathcal{U}$ be a ground set of size ℓn and let $\mathcal{S} \subseteq 2^{\mathcal{U}}$ of subsets each of cardinality ℓ. For a positive integer d (to be chosen later), we construct a graph $G_d = (L \cup A, E)$ as follows: we define $L := \{v_i : i \in [m]\}$ and $A := \{u_j : j \in [\ell \cdot n]\}$. For every $i \in [m]$ and $j \in [\ell \cdot n]$, if set S_i contains element e_j, then we add an edge with unit weight between v_i and u_j in graph G. We add an edge between all pairs of vertices in A with weight $\frac{d}{|A|-1}$, where d will be chosen appropriately (instead of $G[A]$ being a connected d-regular graph as used in the NP-hardness reduction in Section 2). We note that for every vertex $v \in A$, the sum of weight of edges incident to v in the induced subgraph $G[A]$ is d. We define $\mathsf{OPT}_{G_d} := \max_{X \subseteq V} \rho_p(X)$ and set $\rho^* := \frac{\ell^p + \ell \cdot (d+1)^p}{\ell+1}$. We will prove that if the instance is a YES instance, then $\mathsf{OPT}_{G_d} \geq \rho^*$ and if the instance is a NO instance, then $\mathsf{OPT}_{G_d} < (1-\delta_p) \cdot \rho^*$ for some constant $\delta_p > 0$ that depends only on p. We now state the main theorem of the section below.

▶ **Theorem 10.** *For every $p \in (0,1)$, there exist positive integers $\ell \geq 3$ and d such that for an instance $(\mathcal{U}, \mathcal{S})$ of the problem mentioned in Theorem 9, where the ground set $\mathcal{U}$ has size ℓn and every set in $\mathcal{S}$ has size ℓ, the following two hold:*

- *if the instance is a YES-instance, then $\mathsf{OPT}_{G_d} \geq \rho^*$, and*
- *if the instance is a NO-instance, then $\mathsf{OPT}_{G_d} < (1-\delta_p) \cdot \rho^*$ for some constant $\delta_p > 0$ that depends only on p.*

Here, G_d is the graph constructed in the reduction from Exact ℓ-Cover for Weighted Version.

Theorem 8 follows from Theorem 10. We briefly outline our proof of Theorem 10 and refer to the full version for the full proof. It is easy to see that if the instance $(\mathcal{U}, \mathcal{S})$ of the problem mentioned in Theorem 9 is a YES-instance, then $OPT_{G_d} \geq \rho^*$ (similar to the proof of NP-hardness). We focus on showing that if the instance is a NO-instance, then $OPT_{G_d} < (1-\delta)\rho^*$. Let $S \subseteq L$ and $A' \subseteq A$. We need to show that $\rho_p(S \cup A') < (1-\delta)\rho^*$. For this, we recall the proof of NP-hardness in Section 2.2. There, we showed that if the instance does not have an exact ℓ-cover, then $\rho_p(S \cup A') < \rho^*$. For this, we proved that $\rho_p(S \cup A')$ is maximized and is at most ρ^* if $\ell|S|/|A'| = 1$. That proof can be adapted in a straightforward fashion to show that $\rho_p(S \cup A') < (1-\delta)\rho^*$ if $\ell|S|/|A'| \geq 1+\varepsilon$ or if $\ell|S|/|A'| \leq 1-\varepsilon$ for some constants $\delta, \varepsilon > 0$ (even for the graph G_d that appears in the reduction to unweighted p-MEAN DSG) – see cases 1 and 2 in the proof of Theorem 10. Thus, the non-trivial case to handle is if $\ell|S|/|A'| \in (1-\varepsilon, 1+\varepsilon)$. In this situation, we consider two cases: (i) Suppose that $|A'| \leq (1-\varepsilon)|A|$. In this case, we exploit the clique in the weighted graph constructed in the reduction above to conclude that $\rho_p(S \cup A') < (1-\delta)\rho^*$ for some constant $\delta > 0$ (see case 3 in the proof of Theorem 10). (ii) Suppose that $|A'| > (1-\varepsilon)|A|$.

In this case, we rely on the APX-hardness of EXACT ℓ-COVER (i.e., the instance $(\mathcal{U}, \mathcal{S})$ is a NO-instance of the problem mentioned in Theorem 9) to conclude that $\rho_p(S \cup A') < (1-\delta)\rho^*$ for some constant $\delta > 0$ (see case 4 in the proof of Theorem 10). We emphasize that the weighted clique over the set A of vertices in the reduction graph (as opposed to an unweighted d-regular graph over the set A of vertices) is useful in the first case. We also mention that the constant δ_p in Theorem 10 is very small. We give an estimation of $\delta_{1/2}$ in the full version.

4 Approximation Algorithms

We give two new approximation algorithms for p-MEAN DSG. Our algorithms achieve an approximation factor of $\frac{1}{2}$ for all $p \in (-\infty, 1)$. Our algorithms rely on the fact that MAXCORE and DSG can be solved in polynomial time. First, we show that the peeling algorithm used to compute MAXCORE can be adapted to obtain a $\frac{1}{2}$-approximate solutions to p-MEAN DSG for every $p \in (-\infty, 1)$. Secondly, we show that an optimum solution to DSG is a $\frac{1}{2}$-approximate solution to p-MEAN DSG for every $p \in (-\infty, 1)$. We complement these results with a family of graphs for which both algorithms *simultaneously* achieve only a $\frac{1}{2}$-approximation.

Let $G = (V, E)$ be the input graph. We let $S_p^* := \arg\max_{S \subseteq V} M_p(S)$ and let $M_p^* := M_p(S_p^*)$. We need the following fact about the monotonicity of the objective.

▶ **Proposition 11.** *Let $S \subseteq V$. For every $p \le q$, we have $M_p(S) \le M_q(S)$.*

We have the following statement connecting different values of M_p^*.

▶ **Proposition 12.** *For every $p \in [-\infty, 1]$, we have $M_{-\infty}^* \le M_p^* \le M_1^* \le 2M_{-\infty}^*$.*

The first two inequalities follow directly from Proposition 11 and the last inequality follows via a simple known argument connecting degeneracy to the maximum average degree of a subgraph (e.g., see [16]).

$\frac{1}{2}$-approximation via maxcore approach

Our first algorithm leverages the standard greedy peeling algorithm for the maxcore. Our algorithm, denoted Simple-Greedy-p, is given in Figure 2. The algorithm for $p = -\infty$ is the peeling algorithm used to compute maxcore and the algorithm for $p = 1$ is Charikar's greedy peeling algorithm. We recall that Charikar showed that the algorithm achieves a $\frac{1}{2}$-approximation for 1-MEAN DSG.

Simple-Greedy-p($G = (V, E)$)
$S_1 \leftarrow V$
for $i = 1$ **to** $n - 1$ **do**
  $v_i \leftarrow \arg\min_{v \in S_i} d_{S_i}(v)$
  $S_{i+1} \leftarrow S_i - v_i$
return $\arg\max_{S_i} M_p(S_i)$

Figure 2 $\frac{1}{2}$-approximation via greedy peeling for p-MEAN DSG where $p < 1$.

▶ **Theorem 13.** *Let $p \in [-\infty, 1]$ and let S be the output of* Simple-Greedy-$p(G)$. *Then,* $M_p(S) \geq \frac{1}{2} M_p^*$.

Proof. The algorithm for $p = -\infty$ is the peeling algorithm used to compute maxcore. In particular, it is well-known that there exists $i \in [n]$ with $M_{-\infty}(S_i) = M_{-\infty}^*$. By Proposition 11, $M_{-\infty}(S_i) \leq M_p(S_i)$ and by choice of S, we have $M_p(S_i) \leq M_p(S)$. Therefore, $M_{-\infty}^* \leq M_p(S)$. Finally, by Proposition 12, we have $\frac{1}{2} M_p^* \leq M_{-\infty}^*$. Combining these two statements, we get $\frac{1}{2} M_p^* \leq M_p(S)$. ◀

▶ Remark 14. Simple-Greedy-p returns an optimum solution for $p = \infty$. Our results show that for $p \in (-\infty, 1]$, Simple-Greedy-p returns a $\frac{1}{2}$-approximation. However, for $p > 1$, [47] showed that the approximation factor of Simple-Greedy-p can be arbitrarily small.

$\frac{1}{2}$-approximation via 1-mean densest subgraph

Our second algorithm is to simply return a 1-mean densest subgraph. We recall that $S_1^* = \arg\max_{S \subseteq V} M_1(S)$ and it can be computed in polynomial time. We analyze its approximation factor.

▶ **Theorem 15.** *Let $p \in [-\infty, 1]$. Then, $M_p(S_1^*) \geq \frac{1}{2} M_p^*$.*

Proof. We first prove that

$$M_{-\infty}(S_1^*) \geq \frac{1}{2} M_1^*. \tag{3}$$

It suffices to show that $d_{S_1^*}(v) \geq \frac{|E(S_1^*)|}{|S_1^*|}$ for every $v \in S_1^*$. Suppose towards a contradiction that there exists $v \in S_1^*$ such that $d_{S_1^*}(v) < \frac{|E(S_1^*)|}{|S_1^*|}$. Using this and observing $|E(S_1^*)| - |E(S_1^* - v)| = d_{S_1^*}(v)$, after rearranging, we have $\frac{|E(S_1^* - v)|}{|S_1^* - v|} > \frac{|E(S_1^*)|}{|S_1^*|}$. Multiplying through by 2, we obtain $M_1(S_1^* - v) > M_1(S_1^*)$, contradicting the optimality of S_1^*.

Thus, we have

$$M_p(S_1^*) \geq M_{-\infty}(S_1^*) \geq \frac{1}{2} M_1^* \geq \frac{1}{2} M_p^*$$

where the first and last inequality are by Proposition 11 and the second inequality is via (3). ◀

▶ Remark 16. We described two algorithms that achieve an approximation factor of $\frac{1}{2}$. Would returning the best among the sets returned by the two algorithms achieve a factor that is better than $\frac{1}{2}$? In the full version, we construct a non-trivial family of instances on which both algorithms have an approximation factor of at most $\frac{1}{2}$. We emphasize that we seek non-trivial instances – in particular, instances in which the optimum value is arbitrary (i.e., grows) and is not a fixed constant.

5 Conclusion

MAXCORE and DSG are polynomial-time solvable densest subgraph problems with numerous applications. p-MEAN DSG, introduced by Veldt, Benson, and Kleinberg [47], captures both these special cases and provides a unified way to generate subgraphs with different density properties. p-MEAN DSG is polynomial-time solvable for $p = -\infty$ and for $p \geq 1$. In this work, we addressed the complexity and algorithmic aspects of the problem for $p \in (-\infty, 1)$.

We showed that p-MEAN DSG is NP-hard for $p \in (-\infty, 1)$ and WEIGHTED p-MEAN DSG is APX-hard for every fixed constant $p \in (0, 1)$. Our hardness results motivate the need for approximation algorithms for $p \in (-\infty, 1)$. We gave a simple 1/2-approximation for p-MEAN DSG for all $p \in (-\infty, 1)$. Our approximation algorithms also extend to WEIGHTED p-MEAN DSG with the same approximation guarantee in a natural manner.

There are two interesting directions for future work. Firstly, is p-MEAN DSG (or WEIGHTED p-MEAN DSG) APX-hard for every $p \in (-\infty, 1)$? Our APX-hardness results hold for every fixed constant $p \in (0, 1)$. Extending our approach to show APX-hardness for fixed constant $p \in (-\infty, 0)$ requires extending the proof of Theorem 10 to $p < 0$. The technical barrier to extending is the third case in the proof. Secondly, can we improve the approximability of p-MEAN DSG for $p \in (-\infty, 1)$? In contrast to the densest subgraph problem, the non-linearity of the objective function of p-MEAN DSG makes it difficult to develop mathematical programming relaxations. We leave it here as an interesting open problem.

References

1 Reid Andersen and Kumar Chellapilla. Finding dense subgraphs with size bounds. In *International Workshop on Algorithms and Models for the Web-Graph*, pages 25–37. Springer, 2009.

2 Albert Angel, Nick Koudas, Nikos Sarkas, Divesh Srivastava, Michael Svendsen, and Srikanta Tirthapura. Dense subgraph maintenance under streaming edge weight updates for real-time story identification. *The VLDB journal*, 23:175–199, 2014.

3 Yuichi Asahiro, Kazuo Iwama, Hisao Tamaki, and Takeshi Tokuyama. Greedily finding a dense subgraph. *Journal of Algorithms*, 34(2):203–221, 2000.

4 Gary D Bader and Christopher WV Hogue. An automated method for finding molecular complexes in large protein interaction networks. *BMC bioinformatics*, 4(1):1–27, 2003.

5 Bahman Bahmani, Ashish Goel, and Kamesh Munagala. Efficient primal-dual graph algorithms for MapReduce. In *International Workshop on Algorithms and Models for the Web-Graph*, pages 59–78. Springer, 2014.

6 Aditya Bhaskara, Moses Charikar, Eden Chlamtac, Uriel Feige, and Aravindan Vijayaraghavan. Detecting high log-densities: an $o(n^{1/4})$ approximation for densest k-subgraph. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 201–210, 2010.

7 Digvijay Boob, Yu Gao, Richard Peng, Saurabh Sawlani, Charalampos Tsourakakis, Di Wang, and Junxing Wang. Flowless: Extracting densest subgraphs without flow computations. In *Proceedings of The Web Conference 2020*, pages 573–583, 2020.

8 Digvijay Boob, Saurabh Sawlani, and Di Wang. Faster width-dependent algorithm for mixed packing and covering LPs. *Advances in Neural Information Processing Systems 32 (NIPS 2019)*, 2019.

9 Moses Charikar. Greedy approximation algorithms for finding dense components in a graph. In *International Workshop on Approximation Algorithms for Combinatorial Optimization*, pages 84–95. Springer, 2000.

10 Chandra Chekuri, Kent Quanrud, and Manuel R Torres. Densest subgraph: Supermodularity, iterative peeling, and flow. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1531–1555. SIAM, 2022.

11 Chandra Chekuri and Manuel R. Torres. On the generalized mean densest subgraph problem: Complexity and algorithms. *CoRR*, abs/2306.02172, 2023. `doi:10.48550/arXiv.2306.02172`.

12 Li Chen, Rasmus Kyng, Yang P. Liu, Richard Peng, Maximilian Probst Gutenberg, and Sushant Sachdeva. Maximum flow and minimum-cost flow in almost-linear time. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 612–623, 2022.

13 Maximilien Danisch, T.-H. Hubert Chan, and Mauro Sozio. Large scale density-friendly graph decomposition via convex programming. In *Proceedings of the 26th International Conference on World Wide Web*, pages 233–242, 2017.

14 Yon Dourisboure, Filippo Geraci, and Marco Pellegrini. Extraction and classification of dense communities in the web. In *Proceedings of the 16th international conference on World Wide Web*, pages 461–470, 2007.

15 Xiaoxi Du, Ruoming Jin, Liang Ding, Victor E Lee, and John H Thornton Jr. Migration motif: a spatial-temporal pattern mining approach for financial markets. In *Proceedings of the 15th ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135–1144, 2009.

16 Martin Farach-Colton and Meng-Tsung Tsai. Computing the degeneracy of large graphs. In *LATIN 2014: Theoretical Informatics: 11th Latin American Symposium, Montevideo, Uruguay, March 31–April 4, 2014. Proceedings 11*, pages 250–260. Springer, 2014.

17 András Faragó. A general tractable density concept for graphs. *Mathematics in Computer Science*, 1(4):689–699, 2008.

18 András Faragó and Zohre R Mojaveri. In search of the densest subgraph. *Algorithms*, 12(8):157, 2019.

19 Uriel Feige, David Peleg, and Guy Kortsarz. The dense k-subgraph problem. *Algorithmica*, 29(3):410–421, 2001.

20 Eugene Fratkin, Brian T Naughton, Douglas L Brutlag, and Serafim Batzoglou. MotifCut: regulatory motifs finding with maximum density subgraphs. *Bioinformatics*, 22(14):e150–e157, 2006.

21 Giorgio Gallo, Michael D Grigoriadis, and Robert E Tarjan. A fast parametric maximum flow algorithm and applications. *SIAM Journal on Computing*, 18(1):30–55, 1989.

22 Michael R Garey and David S Johnson. *Computers and intractability*, volume 174. freeman San Francisco, 1979.

23 Naveen Garg, Sanjeev Khanna, and Amit Kumar. Hardness of approximation for orienteering with multiple time windows. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2977–2990. SIAM, 2021.

24 David Gibson, Ravi Kumar, and Andrew Tomkins. Discovering large dense subgraphs in massive graphs. In *Proceedings of the 31st international conference on Very large data bases*, pages 721–732, 2005.

25 Aristides Gionis and Charalampos E Tsourakakis. Dense subgraph discovery: KDD 2015 tutorial. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2313–2314, 2015.

26 Anupam Gupta, Euiwoong Lee, Jason Li, Pasin Manurangsi, and Michał Włodarczyk. Losing treewidth by separating subsets. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1731–1749. SIAM, 2019.

27 Elfarouk Harb, Kent Quanrud, and Chandra Chekuri. Faster and scalable algorithms for densest subgraph and decomposition. In *Advances in Neural Information Processing Systems*, 2022.

28 Elfarouk Harb, Kent Quanrud, and Chandra Chekuri. Convergence to lexicographically optimal base in a (contra)polymatroid and applications to densest subgraph and tree packing. In *31st Annual European Symposium on Algorithms*, volume 274, pages 56:1–56:17, 2023.

29 Haiyan Hu, Xifeng Yan, Yu Huang, Jiawei Han, and Xianghong Jasmine Zhou. Mining coherent dense subgraphs across massive biological networks for functional discovery. *Bioinformatics*, 21(suppl_1):i213–i221, 2005.

30 Shuguang Hu, Xiaowei Wu, and TH Hubert Chan. Maintaining densest subsets efficiently in evolving hypergraphs. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 929–938, 2017.

31 Yingsheng Ji, Zheng Zhang, Xinlei Tang, Jiachen Shen, Xi Zhang, and Guangwen Yang. Detecting cash-out users via dense subgraphs. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 687–697, 2022.

32 Viggo Kann. Maximum bounded 3-dimensional matching is max snp-complete. *Information Processing Letters*, 37(1):27–35, 1991.

33 Yasushi Kawase and Atsushi Miyauchi. The densest subgraph problem with a convex/concave size function. *Algorithmica*, 80(12):3461–3480, 2018.

34 Ravi Kumar, Jasmine Novak, and Andrew Tomkins. Structure and evolution of online social networks. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 611–617, 2006.

35 Tommaso Lanciano, Francesco Bonchi, and Aristides Gionis. Explainable classification of brain networks via contrast subgraphs. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 3308–3318, 2020.

36 Tommaso Lanciano, Atsushi Miyauchi, Adriano Fazzone, and Francesco Bonchi. A survey on the densest subgraph problem and its variants. *arXiv preprint arXiv:2303.14467*, 2023.

37 Victor E Lee, Ning Ruan, Ruoming Jin, and Charu Aggarwal. A survey of algorithms for dense subgraph discovery. In *Managing and Mining Graph Data*, pages 303–336. Springer, 2010.

38 Fragkiskos D Malliaros, Christos Giatsidis, Apostolos N Papadopoulos, and Michalis Vazirgiannis. The core decomposition of networks: Theory, algorithms and applications. *The VLDB Journal*, 29:61–92, 2020.

39 Pasin Manurangsi. Almost-polynomial ratio ETH-hardness of approximating densest k-subgraph. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 954–961, 2017.

40 Jean-Claude Picard and Maurice Queyranne. A network flow solution to some nonlinear 0-1 programming problems, with applications to graph theory. *Networks*, 12(2):141–159, 1982.

41 Alexander Schrijver. *Combinatorial optimization: polyhedra and efficiency*, volume 24. Springer Science & Business Media, 2003.

42 Victor Spirin and Leonid A Mirny. Protein complexes and functional modules in molecular networks. *Proceedings of the national Academy of sciences*, 100(21):12123–12128, 2003.

43 Luca Trevisan. Non-approximability results for optimization problems on bounded degree instances. In *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, pages 453–461, 2001.

44 Charalampos Tsourakakis. The k-clique densest subgraph problem. In *Proceedings of the 24th international conference on world wide web*, pages 1122–1132, 2015.

45 Charalampos Tsourakakis and Tianyi Chen. Dense subgraph discovery: Theory and application (Tutoral at SDM 2021), 2021. URL: https://tsourakakis.com/dense-subgraph-discovery-theory-and-applications-tutorial-sdm-2021/.

46 Charalampos E Tsourakakis. A novel approach to finding near-cliques: The triangle-densest subgraph problem. *arXiv preprint arXiv:1405.1477*, 2014.

47 Nate Veldt, Austin R Benson, and Jon Kleinberg. The generalized mean densest subgraph problem. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, pages 1604–1614, 2021.

48 Si Zhang, Dawei Zhou, Mehmet Yigit Yildirim, Scott Alcorn, Jingrui He, Hasan Davulcu, and Hanghang Tong. Hidden: hierarchical dense subgraph detection with application to financial fraud detection. In *Proceedings of the 2017 SIAM International Conference on Data Mining*, pages 570–578. SIAM, 2017.

A Missing proofs for NP-hardness

We include the missing proofs for $p \in (0,1)$ from Section 2 here.

A.1 Proof of Theorem 2

▶ **Theorem 2.** *EXACT ℓ-COVER is* NP*-complete for every integer $\ell \geq 3$.*

Proof. We recall that EXACT 3-COVER is NP-complete [22]. We reduce EXACT ℓ-COVER to EXACT $(\ell+1)$-COVER for $\ell \geq 3$. Consider an instance of EXACT ℓ-COVER with ground set $\mathcal{U}$ of cardinality ℓn and a family $\mathcal{S} \subseteq 2^{\mathcal{U}}$ of subsets each of which has cardinality ℓ. Let $x_1, \ldots, x_n$ be n new elements that are not in $\mathcal{U}$. We create an instance of EXACT $(\ell+1)$-COVER as follows:

- Let $\mathcal{U}' := \mathcal{U} \cup \{x_1, \ldots, x_n\}$ be the ground set. We have that $|\mathcal{U}'| = |\mathcal{U}| + n = (\ell+1) \cdot n$.
- Let $\mathcal{S}' := \{S \cup \{x_i\} : S \in \mathcal{S}, 1 \leq i \leq n\}$. Each set in $\mathcal{S}'$ has cardinality $\ell + 1$.

If there exists an exact ℓ-cover $\{S_{i_1}, S_{i_2}, \ldots, S_{i_n}\}$ of $\mathcal{U}$, then $\{S_{i_1} \cup \{x_1\}, S_{i_2} \cup \{x_2\}, \ldots, S_{i_n} \cup \{x_n\}\}$ is an exact $(\ell+1)$-cover of $\mathcal{U}$. If an exact ℓ-cover of $\mathcal{U}$ does not exist, then an exact $(\ell+1)$-cover of $\mathcal{U}'$ does not exist. Hence, NP-completeness of EXACT 3-COVER implies NP-completeness of EXACT ℓ-COVER for every $\ell \geq 3$. ◀

A.2 Technical Lemmas for Hardness Results

The following inequalities will be used when proving the hardness results.

▶ **Lemma 17.** *Let $p, x \in (0,1)$. Then,*

1. $(1-x)^p < 1 - px$,
2. $(1-x)^p < 1 - px - \frac{p(1-p)}{2}x^2$,
3. $(1+x)^p < 1 + px$, *and*
4. $(1+x)^p < 1 + px - \frac{p(1-p)}{2}x^2 + \frac{p(1-p)(2-p)}{6}x^3$.

Proof. For the first inequality, let $f_1(x) := (1-x)^p - (1-px)$. Then, $f_1'(x) = p \cdot (1 - (1-x)^{p-1}) < 0$, which implies that $f_1(x) < f_1(0) = 0$.

For the second inequality, let $f_2(x) := (1-x)^p - (1 - px - \frac{p(1-p)}{2}x^2)$. Then, $f_2'(x) = p \cdot (1 + (1-p)x - (1-x)^{p-1})$ and $f_2''(x) = p \cdot (1-p) \cdot (1 - (1-x)^{p-2}) < 0$. Hence, $f_2'(x) < f_2'(0) = 0$, which implies that $f_2(x) < f_2(0) = 0$.

For the third inequality, let $f_3(x) := (1+x)^p - (1+px)$. Then, $f_3'(x) = p \cdot ((1+x)^{p-1} - 1) < 0$, which implies that $f_3(x) < f_3(0) = 0$.

For the fourth inequality, let $f_4(x) := (1+x)^p - (1 + px - \frac{p(1-p)}{2}x^2 + \frac{p(1-p)(2-p)}{6}x^3)$. Then, $f_4'(x) = p \cdot ((1+x)^{p-1} - (1 - (1-p)x + \frac{(1-p)(2-p)}{2}x^2))$. Also, $f_4''(x) = p \cdot (1-p) \cdot (1 - (2-p)x - (1+x)^{p-2})$ and $f_4'''(x) = p(1-p)(2-p) \cdot ((1+x)^{p-3} - 1) < 0$, which implies that $f_4''(x) < f_4''(0) = 0$. Thus, $f_4'(x) < f_4'(0) = 0$ and $f_4(x) < f_4(0) = 0$. ◀

A.3 Proof of Lemma 5

Lemma 5 follows from the following stronger lemma. The stronger version will be useful in proving APX-hardness.

▶ **Lemma 18.** *For every $p \in (0,1)$, there exists a positive value $\eta > 0$ and an integer $\ell \geq 3$, both of which depend only on p, such that the following two inequalities hold:*

$$\left(1 - \frac{1}{2\ell}\right)^p < 1 - \frac{1 - 1/2^p}{\ell+1} - \eta \quad \text{and} \quad \left(1 + \frac{1}{2\ell}\right)^p < 1 + \frac{1 - 1/2^p}{\ell+1} - \eta.$$

Proof. Set $\ell_0 := \frac{p \cdot 2^{p-1}}{2^p - p \cdot 2^{p-1} - 1}$. Then, we have $\frac{p}{2\ell_0} = \frac{1-1/2^p}{\ell_0+1}$. We note that

$$\ell_0 - \frac{5}{2} = \frac{p \cdot 2^{p-1}}{2^p - p \cdot 2^{p-1} - 1} - \frac{5}{2} = \frac{2^{1-p} + \frac{7}{5}p - 2}{2^p - p \cdot 2^{p-1} - 1} \cdot \frac{5}{2} \cdot 2^{p-1} = \frac{2^{1-p} - \frac{7}{5}(1-p) - \frac{3}{5}}{2^p - p \cdot 2^{p-1} - 1} \cdot \frac{5}{2} \cdot 2^{p-1} > 0.$$

The last inequality holds since $2^x - \frac{7}{5}x - \frac{3}{5}$ is a convex function and evaluates to zero at $x = 1$ with negative derivative. This implies that $2^{1-p} - \frac{7}{5}(1-p) - \frac{3}{5} > 0$. Consequently,

$$\ell_0 > \frac{5}{2}. \tag{4}$$

We also note that

$$\begin{aligned} 2 - (1-p)\ell_0 &= 2 - (1-p) \cdot \frac{p \cdot 2^{p-1}}{2^p - p \cdot 2^{p-1} - 1} \quad (\text{since } \ell_0 = \frac{p \cdot 2^{p-1}}{2^p - p \cdot 2^{p-1} - 1}) \\ &= \frac{2^p \cdot (2 - \frac{3}{2}p + \frac{1}{2}p^2 - 2^{1-p})}{2^p - p \cdot 2^{p-1} - 1} \\ &= \frac{2^p \cdot (1 + (1-p) + \frac{1}{2}(-p)(1-p) - 2^{1-p})}{2^p - p \cdot 2^{p-1} - 1} < 0. \end{aligned}$$

The last inequality above is because $2^x > 1 + x + \frac{1}{2}x(x-1)$ for all $x \in (0, 1)$. This implies that

$$(1-p)\ell_0 > 2. \tag{5}$$

Hence, if we can prove that the two inequalities of the lemma hold for every $\ell \in [\ell_0 - \frac{1}{2}, \ell_0 + \frac{1}{2}]$, then it implies the lemma. We define two functions $f_1 : (0, +\infty) \to \mathbb{R}$ and $f_2 : (0, +\infty) \to \mathbb{R}$ as

$$f_1(\ell) = (1 - \frac{1}{2\ell})^p + \frac{1 - 1/2^p}{\ell + 1} - 1 \quad \text{and} \quad f_2(\ell) = (1 + \frac{1}{2\ell})^p - \frac{1 - 1/2^p}{\ell + 1} - 1.$$

By setting

$$\eta := \frac{1}{2} \cdot \min \left\{ \frac{\frac{1}{2}p \cdot (2\ell_0 + 3)}{\ell_0 \cdot (2\ell_0 + 3) \cdot (2\ell_0 + 1)^2} \cdot ((1-p)\ell_0 - 2), \frac{p}{2\ell_0 - 1} \cdot \frac{16\ell_0 - 16}{6\ell_0(2\ell_0 + 1)(2\ell_0 - 1)^2} \right\},$$

which is larger than 0 by inequalities (4) and (5), we will prove that $f_1(\ell) < -\eta$ and $f_2(\ell) < -\eta$ for every $\ell \in [\ell_0 - \frac{1}{2}, \ell_0 + \frac{1}{2}]$. We note that

$$\eta < \frac{\frac{1}{2}p \cdot (2\ell_0 + 3)}{\ell_0 \cdot (2\ell_0 + 3) \cdot (2\ell_0 + 1)^2} \cdot ((1-p)\ell_0 - 2) \qquad \text{and} \tag{6}$$

$$\eta < \frac{p}{2\ell_0 - 1} \cdot \frac{16\ell_0 - 16}{6\ell_0(2\ell_0 + 1)(2\ell_0 - 1)^2}. \tag{7}$$

By differentiating f_1 with respect to ℓ, we have

$$\begin{aligned}
\frac{d}{d\ell}f_1(\ell) &= \frac{p}{2\ell^2\cdot(1-\frac{1}{2\ell})^{1-p}} - \frac{1-1/2^p}{(\ell+1)^2}\\
&= \frac{p\cdot(\ell+1)^2 - 2(1-1/2^p)\ell^2\cdot(1-\frac{1}{2\ell})^{1-p}}{2\ell^2\cdot(1-\frac{1}{2\ell})^{1-p}\cdot(\ell+1)^2}\\
&= \frac{2(1-1/2^p)\cdot\ell^2}{2\ell^2\cdot(1-\frac{1}{2\ell})^{1-p}\cdot(\ell+1)^2}\cdot\left(\frac{p}{2(1-1/2^p)}\cdot(1+\frac{1}{\ell})^2-(1-\frac{1}{2\ell})^{1-p}\right)\\
&= \frac{2(1-1/2^p)\cdot\ell^2}{2\ell^2\cdot(1-\frac{1}{2\ell})^{1-p}\cdot(\ell+1)^2}\cdot\left(\frac{\ell_0}{\ell_0+1}\cdot(1+\frac{1}{\ell})^2-(1-\frac{1}{2\ell})^{1-p}\right)\\
&\quad (\text{since } \frac{p}{2\ell_0} = \frac{1-1/2^p}{\ell_0+1})\\
&> \frac{2(1-1/2^p)\cdot\ell^2}{2\ell^2\cdot(1-\frac{1}{2\ell})^{1-p}\cdot(\ell+1)^2}\cdot\left(\frac{\ell_0}{\ell_0+1}\cdot(1+\frac{1}{\ell})^2-(1-\frac{1-p}{2\ell})\right)\\
&\quad (\text{since } (1-\frac{1}{2\ell})^{1-p} < 1-\frac{1-p}{2\ell} \text{ according to Lemma 17})\\
&= \frac{2(1-1/2^p)\cdot\ell^2}{2\ell^2\cdot(1-\frac{1}{2\ell})^{1-p}\cdot(\ell+1)^2}\cdot\frac{\ell_0\cdot(\ell+1)^2-(\ell_0+1)\cdot\ell^2+\frac{1-p}{2}(\ell_0+1)\cdot\ell}{(\ell_0+1)\cdot\ell^2}\\
&> \frac{2(1-1/2^p)\cdot\ell^2}{2\ell^2\cdot(1-\frac{1}{2\ell})^{1-p}\cdot(\ell+1)^2}\cdot\frac{\ell_0\cdot(\ell+1)^2-(\ell_0+1)\cdot\ell^2}{(\ell_0+1)\cdot\ell^2}\\
&= \frac{2(1-1/2^p)\cdot\ell^2}{2\ell^2\cdot(1-\frac{1}{2\ell})^{1-p}\cdot(\ell+1)^2}\cdot\frac{\ell_0(2\ell+1)-\ell^2}{(\ell_0+1)\cdot\ell^2}\\
&\ge \frac{2(1-1/2^p)\cdot\ell^2}{2\ell^2\cdot(1-\frac{1}{2\ell})^{1-p}\cdot(\ell+1)^2}\cdot\frac{(\ell-\frac{1}{2})(2\ell+1)-\ell^2}{(\ell_0+1)\cdot\ell^2} \quad (\text{since } \ell_0 \ge \ell-\frac{1}{2})\\
&= \frac{2(1-1/2^p)\cdot\ell^2}{2\ell^2\cdot(1-\frac{1}{2\ell})^{1-p}\cdot(\ell+1)^2}\cdot\frac{\ell^2-\frac{1}{2}}{(\ell_0+1)\cdot\ell^2}\\
&> 0. \quad (\text{since } \ell \ge \ell_0 - \frac{1}{2} > 2 \text{ according to inequality (4)})
\end{aligned}$$

Hence, the function $f_1(\ell)$ is strictly increasing for $\ell \in [\ell_0 - \frac{1}{2}, \ell_0 + \frac{1}{2}]$. This implies that

$$\begin{aligned}
f_1(\ell) \le f_1(\ell_0+\frac{1}{2}) &= (1-\frac{1}{2\ell_0+1})^p + \frac{1-1/2^p}{\ell_0+\frac{3}{2}} - 1\\
&= \frac{\frac{\ell_0+1}{2\ell_0}\cdot p}{\ell_0+\frac{3}{2}} + (1-\frac{1}{2\ell_0+1})^p - 1 \quad (\text{since } \frac{p}{2\ell_0} = \frac{1-1/2^p}{\ell_0+1})\\
&< \frac{p\cdot(\ell_0+1)}{\ell_0\cdot(2\ell_0+3)} - \frac{p}{2\ell_0+1} - \frac{p(1-p)}{2}\cdot\frac{1}{(2\ell_0+1)^2}\\
&\quad (\text{since } (1-x)^p < 1-px-\frac{p(1-p)}{2}x^2 \text{ according to Lemma 17})\\
&= p\cdot\frac{(\ell_0+1)(2\ell_0+1)^2-\ell_0(2\ell_0+3)(2\ell_0+1)-\frac{1-p}{2}\ell_0(2\ell_0+3)}{\ell_0\cdot(2\ell_0+3)\cdot(2\ell_0+1)^2}\\
&= p\cdot\frac{(2\ell_0+1)-\frac{1-p}{2}\ell_0(2\ell_0+3)}{\ell_0\cdot(2\ell_0+3)\cdot(2\ell_0+1)^2}\\
&< \frac{\frac{1}{2}p\cdot(2\ell_0+3)}{\ell_0\cdot(2\ell_0+3)\cdot(2\ell_0+1)^2}\cdot(2-(1-p)\ell_0)\\
&< -\eta. \quad (\text{by inequality (6)})
\end{aligned}$$

By differentiating f_2 with respect to ℓ, we have

$$
\begin{aligned}
\frac{d}{d\ell} f_2(\ell) &= -\frac{p}{2\ell^2 \cdot (1+\frac{1}{2\ell})^{1-p}} + \frac{1-1/2^p}{(\ell+1)^2} \\
&= \frac{-p\cdot(\ell+1)^2 + 2(1-1/2^p)\ell^2 \cdot (1+\frac{1}{2\ell})^{1-p}}{2\ell^2 \cdot (1+\frac{1}{2\ell})^{1-p} \cdot (\ell+1)^2} \\
&= \frac{2(1-1/2^p)\cdot \ell^2}{2\ell^2 \cdot (1+\frac{1}{2\ell})^{1-p} \cdot (\ell+1)^2} \cdot \left(-\frac{p}{2(1-1/2^p)} \cdot (1+\frac{1}{\ell})^2 + (1+\frac{1}{2\ell})^{1-p}\right) \\
&= \frac{2(1-1/2^p)\cdot \ell^2}{2\ell^2 \cdot (1+\frac{1}{2\ell})^{1-p} \cdot (\ell+1)^2} \cdot \left(-\frac{\ell_0}{\ell_0+1} \cdot (1+\frac{1}{\ell})^2 + (1+\frac{1}{2\ell})^{1-p}\right) \\
&\quad \text{(since } \frac{p}{2\ell_0} = \frac{1-1/2^p}{\ell_0+1}\text{)} \\
&< \frac{2(1-1/2^p)\cdot \ell^2}{2\ell^2 \cdot (1+\frac{1}{2\ell})^{1-p} \cdot (\ell+1)^2} \cdot \left(-\frac{\ell_0}{\ell_0+1} \cdot (1+\frac{1}{\ell})^2 + (1+\frac{1-p}{2\ell})\right) \\
&\quad \text{(since } (1+\frac{1}{2\ell})^{1-p} < 1 + \frac{1-p}{2\ell} \text{ according to Lemma 17)} \\
&= \frac{2(1-1/2^p)\cdot \ell^2}{2\ell^2 \cdot (1+\frac{1}{2\ell})^{1-p} \cdot (\ell+1)^2} \cdot \frac{(\ell_0+1)\cdot \ell^2 + \frac{1-p}{2}(\ell_0+1)\cdot \ell - \ell_0 \cdot (\ell+1)^2}{(\ell_0+1)\cdot \ell^2} \\
&< \frac{2(1-1/2^p)\cdot \ell^2}{2\ell^2 \cdot (1+\frac{1}{2\ell})^{1-p} \cdot (\ell+1)^2} \cdot \frac{(\ell_0+1)\cdot \ell^2 + \frac{1}{2}(\ell_0+1)\cdot \ell - \ell_0 \cdot (\ell+1)^2}{(\ell_0+1)\cdot \ell^2} \\
&= \frac{2(1-1/2^p)\cdot \ell^2}{2\ell^2 \cdot (1+\frac{1}{2\ell})^{1-p} \cdot (\ell+1)^2} \cdot \frac{\ell^2 - \frac{3}{2}\ell_0\ell + \frac{1}{2}\ell - \ell_0}{(\ell_0+1)\cdot \ell^2} \\
&\le \frac{2(1-1/2^p)\cdot \ell^2}{2\ell^2 \cdot (1+\frac{1}{2\ell})^{1-p} \cdot (\ell+1)^2} \cdot \frac{\ell(\ell_0+\frac{1}{2}) - \frac{3}{2}\ell_0\ell + \frac{1}{2}\ell - \ell_0}{(\ell_0+1)\cdot \ell^2} \quad \text{(since } \ell \le \ell_0 + \frac{1}{2}\text{)} \\
&= \frac{2(1-1/2^p)\cdot \ell^2}{2\ell^2 \cdot (1+\frac{1}{2\ell})^{1-p} \cdot (\ell+1)^2} \cdot \frac{-\frac{1}{2}\ell_0\ell + \ell - \ell_0}{(\ell_0+1)\cdot \ell^2} \\
&\le \frac{2(1-1/2^p)\cdot \ell^2}{2\ell^2 \cdot (1+\frac{1}{2\ell})^{1-p} \cdot (\ell+1)^2} \cdot \frac{-\frac{1}{2}\ell_0\ell + \frac{1}{2}}{(\ell_0+1)\cdot \ell^2} \quad \text{(since } \ell \le \ell_0 + \frac{1}{2}\text{)} \\
&< 0. \quad \text{(since } \ell_0 > \frac{5}{2} \text{ and } \ell \ge \ell_0 - \frac{1}{2} > 2 \text{ according to inequality (4))}
\end{aligned}
$$

Hence, function $f_2(\ell)$ is strictly decreasing for $\ell \in [\ell_0 - \frac{1}{2}, \ell_0 + \frac{1}{2}]$. This implies that

$$
\begin{aligned}
f_2(\ell) \le f_2(\ell_0 - \frac{1}{2}) &= (1+\frac{1}{2\ell_0-1})^p - \frac{1-1/2^p}{\ell_0+\frac{1}{2}} - 1 \\
&= -\frac{\frac{\ell_0+1}{2\ell_0}\cdot p}{\ell_0+\frac{1}{2}} + (1+\frac{1}{2\ell_0-1})^p - 1 \quad \text{(since } \frac{p}{2\ell_0} = \frac{1-1/2^p}{\ell_0+1}\text{)} \\
&< -\frac{\frac{\ell_0+1}{2\ell_0}\cdot p}{\ell_0+\frac{1}{2}} + \left(1 + \frac{p}{2\ell_0-1} - \frac{p(1-p)}{2(2\ell_0-1)^2} + \frac{p(1-p)(2-p)}{6(2\ell_0-1)^3}\right) - 1 \\
&\quad \text{(since } (1+x)^p < 1 + px - \frac{p(1-p)}{2}x^2 + \frac{p(1-p)(2-p)}{6}x^3 \\
&\quad \text{according to Lemma 17)} \\
&= \frac{p}{2\ell_0-1} \cdot \left(\frac{1}{\ell_0(2\ell_0+1)} - \frac{1-p}{2(2\ell_0-1)} + \frac{(1-p)(2-p)}{6(2\ell_0-1)^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
&< \frac{p}{2\ell_0 - 1} \cdot \left(\frac{1}{\ell_0(2\ell_0+1)} - \frac{1-p}{2(2\ell_0-1)} + \frac{1-p}{3(2\ell_0-1)^2} \right) \\
&= \frac{p}{2\ell_0 - 1} \cdot \frac{6(2\ell_0-1)^2 - (1-p)\cdot(3\ell_0(2\ell_0+1)(2\ell_0-1) - 2\ell_0(2\ell_0+1))}{6\ell_0(2\ell_0+1)(2\ell_0-1)^2} \\
&= \frac{p}{2\ell_0 - 1} \cdot \frac{6(2\ell_0-1)^2 - (1-p)\ell_0\cdot(12\ell_0^2 - 4\ell_0 - 5)}{6\ell_0(2\ell_0+1)(2\ell_0-1)^2} \\
&< \frac{p}{2\ell_0 - 1} \cdot \frac{6(2\ell_0-1)^2 - 2\cdot(12\ell_0^2 - 4\ell_0 - 5)}{6\ell_0(2\ell_0+1)(2\ell_0-1)^2} \\
&\qquad \text{(since } (1-p)\ell_0 > 2 \text{ according to inequality (5)} \\
&\qquad \text{and } 12\ell_0^2 - 4\ell_0 - 5 = (6\ell_0-5)(2\ell_0+1) > 0) \\
&= \frac{p}{2\ell_0 - 1} \cdot \frac{-16\ell_0 + 16}{6\ell_0(2\ell_0+1)(2\ell_0-1)^2} \\
&< -\eta. \quad \text{(by inequality (7))}
\end{aligned}
$$

Thus, for every $\ell \in [\ell_0 - \frac{1}{2}, \ell_0 + \frac{1}{2}]$, we have $f_1(\ell) < -\eta$ and $f_2(\ell) < -\eta$. ◀