Cliques in Regular Graphs
and the Core-Periphery Problem in Social Networks

Ulrik Brandes[1], Eugenia Holm[1], and Andreas Karrenbauer[2(✉)]

[1] Department of Computer & Information Science,
University of Konstanz, Konstanz, Germany
{ulrik.brandes,eugenia.holm}@uni-konstanz.de
[2] Max Planck Institute for Informatics, Saarbrücken, Germany
andreas.karrenbauer@mpi-inf.mpg.de

Abstract. The existence of a densely knit core surrounded by a loosely connected periphery is a common macro-structural feature of social networks. Formally, the COREPERIPHERY problem is to partition the nodes of an undirected graph $G = (V, E)$ such that a subset $X \subset V$, the *core*, induces a dense subgraph, and its complement $V \setminus X$, the *periphery*, induces a sparse subgraph. Split graphs represent the ideal case in which the core induces a clique and the periphery forms an independent set. The number of missing and superfluous edges in the core and the periphery, respectively, can be minimized in linear time via edit distance to the closest split graph.

We show that the COREPERIPHERY becomes intractable for standard notions of density other than the absolute number of misclassified pairs. Our main tool is a regularization procedure that transforms a given graph with maximum degree d into a d-regular graph with the same clique number by adding at most $d \cdot n$ new nodes. This is of independent interest because it implies that finding a maximum clique in a regular graph is NP-hard to approximate to within a factor of $n^{1/2-\varepsilon}$ for all $\varepsilon > 0$.

1 Introduction

In the COREPERIPHERY problem, we are given a graph $G = (V, E)$ and our goal is to find a bipartition of V into a tightly knit core and a loosely connected periphery. To formalize the COREPERIPHERY problem, we compare the given graph with the class of *split graphs*, i.e., graphs that admit a bipartition into a complete induced subgraph and a set of mutually non-adjacent vertices. Our aim is now to minimize the error, i.e., the deviation from the ideal case.

To this end, we want to simultaneously maximize the density in the core and minimize the density in the periphery.

The CorePeriphery problem is highly relevant for the analysis of social networks [2,21] in various domains [1,5,6,8,9,15,16,18–20,22–25]. The most common formalizations are due to Borgatti and Everett [2] and heuristic algorithms are used [3,4] to separate a core from its periphery. Rombach et al. [21] present a method for identifying multiple cores in a network. Holme [12] introduced a core-periphery-coefficient to measure if the network can be bisected in core and periphery. Zhang et al. [26] developed a statistically principled method, where they use a maximum-likelihood fit, for detecting a core-periphery-decomposition.

The CorePeriphery is closely related to other problems in graph theory. For example, if we omit the condition of a bipartition, we get the problem of finding a densest subgraph in a given graph. Goldberg [10] shows that this problem can be solved in polynomial time for the linear density by using an algorithm based on a network flow computation. The problem to find a vertex partition with maximal sum of the densities of the subsets is known to be NP-hard [7]. Khuller and Saha [14] give 2-approximation algorithms for computing a densest subgraph with at least k vertices for a given k.

The problems of finding large cliques and independent sets are notoriously difficult: it is NP-hard to approximate the size of the largest clique/independent set within a factor of $n^{1-\varepsilon}$ for all $\varepsilon > 0$ [13,27]. Mathieson and Szeider [17] showed that the clique problem remains $W[1]$-hard even in regular graphs. To this end, they proposed a gadget to regularize a given graph, which yields $n^{1/3-\varepsilon}$-hardness for approximating the size of the largest clique in regular graphs. To the best of our knowledge, this was the best previously known lower bound for the hardness of approximation of cliques in regular graphs.

1.1 Our Contribution

We propose a novel regularization procedure that transforms a given graph with n nodes and maximum degree d into a d-regular graph by adding $O(d \cdot n)$ nodes and $O(d^2n)$ edges without increasing the size of the largest clique provided that the given graph was not already triangle-free. This improves the construction in [17], which uses $O(d^2n)$ extra nodes. Furthermore, we show that this implies that it is NP-hard to approximate the size of the largest clique in a regular graph within a factor of $O(n^{1/2-\varepsilon})$ for all $\varepsilon > 0$. Finally we show by a new proof technique that the CorePeriphery problem is NP-hard for linear and quadratic densities.

1.2 Preliminaries

We start by a formal treatment of our problems. A graph $G = (V, E)$ is *d-regular* if all nodes in V have exactly d neighbors, i.e., all nodes in V have the same *degree $deg(v) = d$*.

We say that a graph $G' = (V', E')$ is an *induced subgraph* of the graph G, $G' \subseteq G$, if $V' \subseteq V$ and E' consists of all edges in E, which have both endnodes in V'. We also write $G' = G[V']$. If we consider the number of adjacent nodes of $v \in V'$ in a subgraph $G' = (V', E')$, we express this by $deg_{G'}(v)$.

The *complement edge set* $\bar{E} := \{\{v, w\} \subseteq V \ : \ \{v, w\} \notin E, v \neq w\}$ of $G = (V, E)$ consists of all edges which are not included in E. The *complement graph* $\bar{G} = (V, \bar{E})$ of G is a graph with the same set of vertices and the complement edge set of G.

An *ℓ-clique* K_ℓ is an induced *complete* subgraph of G with ℓ nodes, i.e., every pair of nodes in the node set of K_ℓ is connected by an edge. A famous problem in the graph theory is the MAXCLIQUE problem. The goal of it is to find a maximum clique in the given graph G, i.e., a clique with the biggest number of nodes. We call this number the *clique number* $\omega(G)$ of the graph G. If G is a regular graph, we call the problem of computing $\omega(G)$ REGULARCLIQUE.

A graph $G = (V, E)$ is called *bipartite* if the node set V can be partitioned into two not empty subsets V_1 and V_2 so that every edge in E has one endnode in V_1 and one endnode in V_2. A bipartite subgraph $K_{a,b} = (V_1' \cup V_2', E')$ in G with $|V_1'| = a$ and $|V_2'| = b$ is called *biclique*, if every node of V_1' is adjacent to every node of V_2'. A set of edges in the graph G that are mutually disjoint is called *matching* in G. If every node in G is incident to one edge of the matching then it is a *perfect* matching. A complete bipartite graph with $2n$ nodes from which the edges of a perfect matching have been removed is called *crown graph*.

2 RegularClique is NP-Hard to Approximate

To prove the hardness result, we first describe a polynomial-time algorithm to regularize a given graph $G = (V, E)$ with $|V| = n$ and $|E| = m$ and maximum degree d.

2.1 Regularization Procedure

Our goal is to augment a graph $G = (V, E)$ with maximum degree d by additional nodes and edges to obtain a d-regular graph $G_d = (V_d, E_d)$ such that $G \subseteq G_d$ and $\omega(G) = \omega(G_d)$. We assume w.l.o.g. that $|V| = 2x$ with $x \in \mathbb{Z}^+$: If the number of nodes is odd, we add an isolated node to V. This does not change $\omega(G)$ and increases n to $n + 1$, which does not harm the asymptotic statements later on.

Since we will not remove any edges, we must fill up the degree of each node until it reaches d. Then the regularized graph $G_d = (V_d, E_d)$ will contain exactly $y := d \cdot n - 2m$ new edges, each with exactly one incident node in V. We choose $r, s \in \mathbb{Z}_{\geq 0}$ so that $m = d \cdot s - r$ with $0 \leq r < d$. Thus,

$$y = 2d \cdot x - 2d \cdot s + 2r = 2d \cdot (x - s) + 2r.$$

Now we consider a crown graph with $2d$ nodes, i.e., biclique $K_{d,d}$ without a perfect matching. Every node of it has degree equal to $d - 1$. We add $x - s$ such graphs to G and connect every node of them with one of the nodes in V with degree less than d. Thereby, we may connect one node in V with one or several nodes of these auxiliary graphs until its degree is equal to d.

For the $2r$ remaining required new edges we add a further biclique $K_{d,d}$ without a matching with r edges. This auxiliary graph contains $2r$ nodes with degree $d-1$. We connect those nodes with the nodes in V whose degree is still smaller than d (Fig. 1).
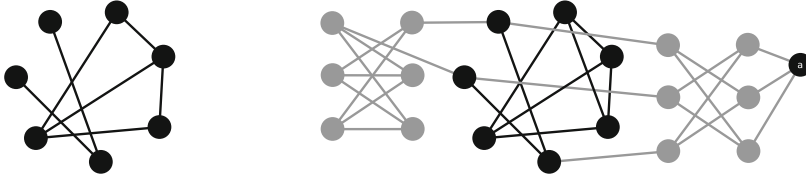
Fig. 1. A graph G with $n = 7$, $m = 8$ and $d = 3$. We add an auxiliary node a, set $x = 4$, $s = 3$ and $r = 1$. Then we construct G_d by adding one $K_{3,3}$ without a perfect matching (on the right side of the figure) and one $K_{3,3}$ without the matching with $r = 1$ edges. Finally we create new edges according to the above description.

Theorem 1. *The graph G contains a k-clique with $k \geq 4$, if and only if its regularized version G_d contains a k-clique.*

Proof. Let $k \geq 4$. Assume that G contains a k-clique. As the regularization procedure only adds nodes and edges and does not remove any of the original nodes or edges, G is an induced subgraph of G_d and thus G_d contains the same k-clique.

On the other hand, if G_d contains a k-clique K with $k \geq 4$, at most two of its nodes may be contained in $V_d \backslash V$ because the added bipartite graphs are triangle-free. But if K contains exactly two nodes from $V_d \backslash V$, then these two nodes can have at most one common neighbor in K because each of them is only incident to exactly one node in V, a contradiction to the assumption that K is a clique with at least 4 nodes. Similarly, if exactly one node from $V_d \backslash V$ is contained in K, then it is incident to exactly one other node in K, again a contradiction. Thus, K is completely contained in V, which proves the claim. □

2.2   Hardness of Approximation

The REGULARCLIQUE problem is formally defined as follows.

*Problem 1 (*REGULARCLIQUE*).* Given a regular graph G and an integer k, decide whether G contains a clique of size k.

This problem is not only NP-hard, but also very hard to approximate. To prove this, we recall the situation for MAXCLIQUE in general graphs. Zuckerman [27] derandomized a construction of Håstad [13] to obtain the following theorem.

Theorem 2 ([13, 27]). *Let $\varepsilon > 0$. Given a graph $G = (V, E)$, it is NP-hard to approximate* MaxClique *to within a factor of $|V(G)|^{1-\varepsilon}$.*

From this, we derive a similar hardness result, where we lose a $\sqrt{n}$-factor due to the blow-up of $O(dn)$ in the number of nodes with our regularization procedure. We restrict ourselves to the case $k \geq 4$ because we can decide whether there is a K_3 by enumerating all triples of nodes in $O(n^3)$ time and because the cases for $k \in \{1, 2\}$ are trivial.

Theorem 3. *It is NP-hard to approximate* RegularClique *within a factor of $n^{\frac{1}{2}-\varepsilon}$ for all $\varepsilon > 0$.*

Proof. Let $G = (V, E)$ be a given undirected graph with $n := |V|$ and $m := |E|$, and let $G_d = (V_d, E_d)$ denote its regularized version. For our considerations, we can assume that the graph G is a connected graph and therefore $m \geq n - 1 \Leftrightarrow 2m \geq 2n - 2$ and hence $2m \geq n$ for all $n \geq 2$. By the regularizing construction we get

$$N := |V_d| = n + 2d \cdot (x - s) + 2d = n + d \cdot n - 2m - 2r + 2d \leq d \cdot n \leq n^2.$$

Suppose that there exists an approximation algorithm $\mathcal{A}_{RegClique}$ for regular clique within a factor of $N^{\frac{1-\varepsilon}{2}}$ for an $\varepsilon > 0$. Then we can find a k-clique K in G_d with $k \geq \omega(G_d)/N^{\frac{1-\varepsilon}{2}}$ nodes. According to the Theorem 1, K is contained in G and $\omega(G_d) = \omega(G)$. Thus, $k \geq \omega(G)/N^{\frac{1-\varepsilon}{2}} \geq \omega(G)/n^{1-\epsilon}$. Thus, we would have an $n^{1-\varepsilon}$-approximation for MaxClique. Theorem 2 proves the statement above.   □

For the sake of presentation, we further restrict the range for k to $\{4, \ldots, d\}$. This is w.l.o.g. because d-regular graphs cannot have a clique with more than $d + 1$ nodes, the cases for $k \in \{1, 2, 3\}$ can be decided in polynomial time, as well as the case for $k = d + 1$ as the following Lemma shows.

Lemma 1. *A d-regular graph $G = (V, E)$ contains a clique K_{d+1}, if and only if G contains a connected component with $d + 1$ nodes.*

Proof. Let $K \subseteq V$ be a clique in G with $d + 1$ nodes. Since $deg_K(v) = d$ for all $v \in K$ every node in K is adjacent to all other nodes in K. Because G is a d-regular graph, there cannot be an edge $\{v, w\}$ with $v \in K$ and $w \notin K$.

On the other hand, let G contain a connected component $V' \subseteq V$ with $d + 1$ nodes. Because G is a d-regular graph, every node in V' is adjacent to d nodes. This means that every node in V' is adjacent to all nodes in V'. Thus, $G[V']$ is a clique with $d + 1$ nodes.   □

3   Application to CorePeriphery

We apply the results from the previous section to prove NP-hardness of two versions of the CorePeriphery problem. Generally speaking our aim is to

decompose a given graph $G = (V, E)$ into a core, i.e., nodes that are tightly connected, and a periphery, i.e., vertices that are loosely connected. The ideal case is a so called *split graph*. This is a graph for which there exists a bipartition of its vertices into a clique and an *independent set*, i.e., a set of nodes that not induce any edge. Hammer and Simeone [11] showed that the split graphs can be recognized in linear time.

Theorem 4 (Hammer and Simeone, 1981 [11]). *Let $G = (V, E)$ be an undirected graph with $|V| = n$ and the degree sequence $d_1 \geq \cdots \geq d_n$. Define $k := \max\{i : d_i \geq i - 1\}$. Then, G is a split graph if and only if the splittance*

$$\frac{1}{2} \left(k(k-1) - \sum_{i=1}^{k} d_i + \sum_{j=k+1}^{n} d_j \right) = 0.$$

Furthermore, if this is the case, then k is the clique number of G.

For all other graphs we try to minimize the deviation from this ideal case, i.e., to minimize the splittance. Observe that $\frac{1}{2} \left(k(k-1) - \sum_{i=1}^{k} d_i + \sum_{j=d+1}^{n} d_j \right)$ edges have to be added/deleted to make G a split graph. That is, the splittance is the number of edges that have to be added or to be deleted to obtain a bipartition of a given graph into a clique as a core and the independent set as periphery. Furthermore, Hammer and Simeone showed in [11] that the splittance of any graph can be determined in linear time.
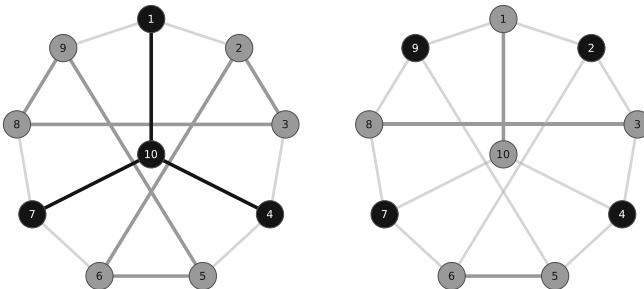
Fig. 2. Here we see two different cores (black nodes) in the Petersen graph. The split-tance of the graph is always equal to 9 if the core consists of 4 nodes, although the core induces no edges.

The fact that the splittance of a d-regular graph is equal for all cores with $d + 1$ nodes, illustrates that the splittance is not able to discriminate certain situations. For example, Fig. 2 shows two cores in the Petersen graphs that are both optimal w.r.t. the splittance, but one of them even induces an independent set — the opposite of a clique. So it is natural to ask for the size normalized deviation like how many edges on average per node must be added in the core

and be deleted in the periphery to obtain a split graph. Analogously, we can consider the problem so that we ask which fraction of edges of the clique is to be added to the core and which fraction of potential edges in periphery is to be deleted to obtain the ideal case.

To facilitate the comparison of further objective functions, we introduce the following notations. Let $G = (V, E)$ be a graph and $X \subseteq V$ a candidate for the core of G. The number of edges and non-edges in the core X are denoted by

$$c_1(X) := |\{e \in E \ : \ e \subseteq X\}|, \ c_0(X) := |\{\bar{e} \in \bar{E} \ : \ \bar{e} \subseteq X\}|$$

and similarly in the periphery $V \setminus X$:

$$p_1(X) := |\{e \in E \ : \ e \cap X = \emptyset\}|, \ p_0(X) := |\{\bar{e} \in \bar{E} \ : \ \bar{e} \cap X = \emptyset\}|.$$

We write $c_1, c_0, p_1,$ or p_0 if X is clear from the context. Using our notion the splittance is equal to $c_0 + p_1$.

Our aim is to decompose the nodes of the graph into the core X and periphery $V \setminus X$ such that the density of the subgraph induced by X is maximal and the density of the subgraph induced by $V \setminus X$ is minimal. To combine these criteria in a single objective function that mimics the splittance, we minimize $sparsity_G(X) + density_G(V \setminus X)$, where $sparsity_G(X) := density_{\bar{G}}(X)$, i.e., the sparsity is defined as density in the complement graph.

Popular density functions are the linear and the quadratic density. The *linear density* $d_1(X)$ of $X \subseteq V$ in a graph $G = (V, E)$ is defined as the average degree in the subgraph induced by X. That is,

$$d_1(X) = \frac{1}{|X|} \sum_{v \in X} deg_X(v) = \frac{2c_1}{|X|}.$$

The *quadratic density* $d_2(X)$ of $X \subseteq V$ in a graph $G = (V, E)$ is the ratio of existent edges to the number of all possible edges in the subgraph induced by X. That is,

$$d_2(G) := \frac{c_1}{\frac{|X|(|X|-1)}{2}} = \frac{2c_1}{|X|(|X| - 1)}.$$

To facilitate the discussion, we split the contribution of the sparsity of the core X and the density of the periphery $V \setminus X$ into two functions $f(X)$ and $g(X)$, respectively, such that their sum defines the objective function $h(X)$. This is summarized in the Table 1.

The function $f(X)$ counts non-edges in X, in relation to the size of X and $g(X)$ counts edges having both incident nodes in the set $V \setminus X$, in relation to the size of $V \setminus X$. These quantities yield the deviation of the core X and the periphery $V \setminus X$ to a perfect core-periphery structure, i.e., the average number of edges per node to be added to the core or to be deleted from the periphery in the case of linear density, and the percentage of missing edges in the core or surplus edges in the periphery, respectively, to make X a clique and $V \setminus X$ an independent set.

Table 1. The decomposition of the objective function into contributions of the core and the periphery for splittance, linear, and quadratic normalization.

Deviation	Absolute	Linear	Quadratic
Core X	$c_0(X)$	$f(X) := \frac{2c_0(X)}{|X|}$	$f(X) := \frac{2c_0(X)}{|X||X-1|}$
Periphery $V \setminus X$	$p_1(X)$	$g(X) := \frac{2p_1(X)}{|V|-|X|}$	$g(X) := \frac{2p_1(X)}{(|V|-|X|)(|V|-|X|-1)}$
Total deviation	$c_0(X) + p_1(X)$	$h(X) := f(X) + g(X)$	

We will show in the following that the CorePeriphery problem is NP-hard for both of these densities. For this we will use the hardness result for Regu-larClique from the previous section. The main idea of our proof is that we augment the graph by isolated nodes such that any reasonably good solution will take a clique as the core. That is, all solutions that have an incomplete subgraph as a core will have a worse objective value than taking the two endpoints of any edge as the core. It is important for our argument that the input graph is regular and therefore we will make use of our regularization procedure.

3.1   Linear Density

First we show that if the core-candidate X is not a clique, then the value of the density of X never falls below a certain value.

Lemma 2. *If a non-empty set $X \subseteq V$ does not induce a clique, then $f(X) \geq \frac{2}{d+1}$ for any $d \geq 2$.*

Proof. We first consider the case $|X| \leq d+1$. Since X does not induce a clique, the induced subgraph misses at least one edge from being complete, i.e., $c_0 \geq 1$. Thus, we have

$$f(X) = \frac{2c_0}{|X|} \geq \frac{2}{|X|} \geq \frac{2}{d+1}.$$

If $|X| \geq d+2$, we have $|X| - 1 \geq d+1$ and so X must miss more than

$$\frac{|X|(|X|-1)}{2} - \frac{d|X|}{2} \geq \frac{|X|(d+1) - d|X|}{2} = \frac{|X|}{2}$$

edges, i.e., $c_0 \geq \frac{|X|}{2}$. Thus, $f(X) = \frac{2c_0}{|X|} \geq 1 \geq 2/(d+1)$ for all $d \geq 2$.   □

The idea for proving NP-hardness is to augment a given graph by isolated nodes such that $h(X)$ is at most $1/d < 2/(d+1)$ whenever X induces a clique (even if it induces single edge) in a graph with $d \geq 2$.

Lemma 3. *A d-regular graph G contains a clique of size $k \leq d$, if and only if, $G' = (V', E)$ with the node set V' consisting of the n nodes of V and q additional isolated nodes (i.e., $|V'| = n + q$) contains a core $X \subseteq V'$ with*

$$h(X) \leq \frac{nd - 2dk + k(k-1)}{n + q - k}$$

for all $q \geq d^2 n$.

Proof. Let X be a k-clique in G and thus also in G'. Note that $f(X) = 0$. Moreover, the total number of edges is $\frac{nd}{2}$ due to the regularity and the number of edges incident to nodes in X is given by $kd - \binom{k}{2}$. Hence,

$$h(X) = g(X) = 2\frac{\frac{nd}{2} - (dk - \binom{k}{2})}{n + q - k} = \underbrace{\frac{nd - 2dk + k(k-1)}{n + q - k}}_{=:g(k)}.$$

To prove the converse direction, we show that $g(k)$ is decreasing in the range of $1 \le k \le d$. To this end, we consider the first derivative of $g(k)$, i.e.,

$$\begin{aligned} g'(k) &= \frac{(2k - 2d - 1)q + 2kn - dn - n - k^2}{(n + q - k)^2} \\ &\le -\frac{q - dn + n + 1}{(n + q - k)^2} \\ &\le -\frac{(d^2 - d + 1)n + 1}{(n + q - k)^2} < 0 \end{aligned}$$

Thus, $g(k) \le g(1) = (dn - 2d)/(n + q - 1) \le (dn - 2d)/(n + d^2 n - 1) \le 1/d < 2/(d + 1)$ for all $d \ge 2$. Hence, G contains a k-clique if and only if there is an $X \subset V'$ with

$$h(X) \le \frac{nd - 2dk + k(k-1)}{q + n - k}.$$

$\square$

As a consequence, we obtain a reduction to prove NP-hardness of linear CorePeriphery.

Theorem 5. *Solving the problem* CorePeriphery *with linear density is* NP-*complete.*

Proof. Given a d-regular graph $G = (V, E)$ and an integer $k \in \{1, \dots, d\}$, we wish to decide whether G contains a k-clique (by Lemma 1 it is sufficient to consider k-cliques with $k \le d$). We add $q = nd^2$ isolated nodes to G and thereby obtain $G' = (V', E')$ for which we compute the CorePeriphery problem. Combining Lemmas 2 and 3, the reported core will be a clique. $\square$

3.2   Quadratic Density

Lemma 4. *If $X \subseteq V$ does not induce a clique, then $f(X) \ge \frac{1}{d^2}$ for $d \ge 2$.*

Proof. We again consider the case $|X| \le d + 1$ first. Since X does not induce a clique, it contains at least one non-edge in $\overline{E}$ and thereby

$$f(X) \ge \frac{2}{|X|(|X| - 1)} \ge \frac{2}{(d + 1)d} \ge \frac{1}{d^2}.$$

If $|X| = d + \ell$ with $2 \le \ell \in \mathbb{Z}$, than there are $(d+\ell)(d+\ell-1)/2 - d(d+\ell)/2$ non-edges and thus

$$
\begin{aligned}
f(X) &\ge \frac{(d+\ell)(d+\ell-1) - d(d+\ell)}{(d+\ell)(d+\ell-1)} \\
&= \frac{d+\ell-1-d}{d+\ell-1} = 1 - \frac{d}{d+\ell-1} \\
&\ge 1 - \frac{d}{d+1} = \frac{1}{d+1}.
\end{aligned}
$$

$\square$

For proving NP-hardness for the quadratic density, we augment a given graph by q isolated nodes analogous to the case of linear density such that $h(X)$ is strictly less than $1/d^2$ whenever X induces a clique (even if it induces single edge).

Lemma 5. *A d-regular graph G contains a clique of size $k \le d$, if and only if, $G' = (V', E)$ with the node set V' consisting of the n nodes of V and q additional isolated nodes (i.e., $|V'| = n + q$) contains a core $X \subseteq V'$ with*

$$
h(X) \le \frac{nd - 2dk + k(k-1)}{(n+q-k)(n+q-k-1)}
$$

for all $q \ge dn$.

Proof. At first we assume that X is a k-clique in G and also in G'. Analogous to the proof for the linear density we obtain

$$
\begin{aligned}
h(X) = g(X) &= \underbrace{\frac{nd - 2dk + k(k-1)}{(n+q-k)(n+q-k-1)}}_{=:g(k)} \\
&= \frac{nd - k(2d - k + 1)}{q^2 + k^2 + n(n - 2k - 1) + q(2n - 2k - 1) + k} < \frac{nd}{q^2}.
\end{aligned}
$$

For $q \ge dn$ we get $g(X) < \frac{1}{dn}$ and therefore $h(X) < \frac{1}{dn} < \frac{1}{d^2}$. Recall that $h(X) \ge \frac{1}{d^2}$ if X is not a clique by Lemma 4.

Now we have to prove that a larger clique is preferred instead of a smaller one. To this end, we show that the function $g(k)$ decreases for increasing k by considering the difference

$$
g(k-1) - g(k) = \frac{-2(1 - dk - k - n - q + d - dq + kn + kq)}{(n+q-k+1)(n+q-k)(n+q-k-1)}
$$

It is easy to verify that the denominator of this difference is positive. So we have to investigate the numerator only.

$$
-2(1 - dk - k - n - q + d - dq + kn + kq) = 2(q(d - k + 1) + dk + k - kn - d + n - 1)
$$

We show that for $q \geq dn$ this term is positive.

$$
\begin{aligned}
&2(q(d-k+1) + dk + k - kn - d + n - 1) \\
&\geq 2(dn(d-k+1) + dk + k - kn - d + n - 1) \\
&\geq 2(dn(d-k+1) + d + 1 - kn - d + n - 1) \\
&\geq 2(dn(d-d+1) + d + 1 - dn - d + n - 1) \\
&= 2n > 0
\end{aligned}
$$

As a consequence a minimizer for the CorePeriphery problem for quadratic density is a clique $X \subseteq V$.   □

Theorem 6. *Solving the problem* CorePeriphery *with quadratic density is* NP*- complete.*

Proof. Given a d-regular graph $G = (V, E)$ and an integer $k \in \{1, \ldots, d\}$. The question is again whether G contains a clique with k nodes. We construct a graph $G' = (V', E')$ by adding $q = nd$ isolated nodes to G analogous to the case of linear density. Then we solve the CorePeriphery problem on G'. Combining Lemmas 4 and 5, we obtain a clique as the core.   □

References

1. Alba, R.D., Moore, G.: Elite social circles. Sociol. Methods Res. **7**(2), 167–188 (1978)
2. Borgatti, S.P., Everett, M.G.: Models of core/periphery structures. Soc. Netw. **21**(4), 375–395 (2000)
3. Borgatti, S.P., Everett, M.G., Freeman, L.C.: Ucinet for windows: software for social network analysis (2002)
4. Boyd, J.P., Fitzgerald, W.J., Beck, R.J.: Computing core/periphery structures and permutation tests for social relations data. Soc. Netw. **28**(2), 165–178 (2006)
5. Chase-Dunn, C.K.: Global Formation: Structures of the World-Economy. Rowman & Littlefield, Lanham (1998)
6. Corradino, C.: Proximity structure in a captive colony of Japanese monkeys (macaca fuscata fuscata): an application of multidimensional scaling. Primates **31**(3), 351–362 (1990)
7. Darlay, J., Brauner, N., Moncel, J.: Dense and sparse graph partition. Discrete Appl. Math. **160**(16), 2389–2396 (2012)
8. Doreian, P.: Structural equivalence in a psychology journal network. J. Am. Soc. Inf. Sci. **36**(6), 411–417 (1985)
9. Faulkner, R.R.: Music on Demand. Transaction Publishers, Piscataway (1983)
10. Goldberg, A.V.: Finding a maximum density subgraph. University of California Berkeley, CA (1984)
11. Hammer, P.L., Simeone, B.: The splittance of a graph. Combinatorica **1**(3), 275–284 (1981)
12. Holme, P.: Core-periphery organization of complex networks. Phys. Rev. E **72**(4), 046111 (2005)
13. Håstad, J.: Clique is hard to approximate within $n^{1-\epsilon}$. Acta Math. **182**, 105–142 (1999)

14. Khuller, S., Saha, B.: On finding dense subgraphs. In: Albers, S., Marchetti-Spaccamela, A., Matias, Y., Nikoletseas, S., Thomas, W. (eds.) ICALP 2009, Part I. LNCS, vol. 5555, pp. 597–608. Springer, Heidelberg (2009)
15. Krugman, P.: The self-organizing economy. Number 338.9 KRU 1996. CIMMYT (1996)
16. Laumann, E.O., Pappi, U.: Networks of collective actions, New York (1976)
17. Mathieson, L., Szeider, S.: The parameterized complexity of regular subgraph problems andgeneralizations. In: Proceedings of the Fourteenth Symposium on Computing: The Australasian Theory - Volume 77, CATS 2008, pp. 79–86, Darlinghurst, Australia. Australian Computer Society Inc. (2008)
18. Mintz, B., Schwartz, M.: Interlocking directorates and interest group formation. Am. Sociol. Rev. **46**, 851–869 (1981)
19. Mullins, N.C., Hargens, L.L., Hecht, P.K., Kick, E.L.: The group structure of cocitation clusters: a comparative study. Am. Sociol. Rev. **42**, 552–562 (1977)
20. Nemeth, R.J., Smith, D.A.: International trade and world-system structure: a multiple network analysis. Rev. (Fernand Braudel Center) **8**(4), 517–560 (1985)
21. Rombach, M.P., Porter, M.A., Fowler, J.H., Mucha, P.J.: Core-periphery structure in networks. SIAM J. Appl. Math. **74**(1), 167–190 (2014)
22. Smith, D.A., White, D.R.: Structure and dynamics of the global economy: network analysis of international trade 1965–1980. Soc. Forces **70**(4), 857–893 (1992)
23. Snyder, D., Kick, E.L.: Structural position in the world system, economic growth, 1955–1970: a multiple-network analysis of transnational interactions. Am. J. Sociol. **84**, 1096–1126 (1979)
24. Steiber, S.R.: The world system and world trade: an empirical exploration of conceptual conflicts. Sociol. Q. **20**(1), 23–36 (1979)
25. Wallerstein, I.: The Modern World-System I: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century, with a New Prologue, vol. 1. University of California Press, Berkeley (2011)
26. Zhang, X., Martin, T., Newman, M.E.: Identification of core-periphery structure in networks. Phys. Rev. E **91**(3), 032803 (2015)
27. Zuckerman, D.: Linear degree extractors and the inapproximability of max clique and chromatic number. In: Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing, STOC 2006, pp. 681–690. ACM, New York (2006)