The Maximum Number of Edges in a Minimal Graph of Diameter 2

Zoltán Füredi*
MATHEMATICAL INSTITUTE
OF THE HUNGARIAN ACADEMY OF SCIENCES
1364 BUDAPEST
P.O.B. 127, HUNGARY

ABSTRACT

A graph $\mathcal{G}$ of diameter 2 is minimal if the deletion of any edge increases its diameter. Here the following conjecture of Murty and Simon is proved for $n > n_0$. If $\mathcal{G}$ has n vertices then it has at most $\lfloor n^2/4 \rfloor$ edges. The only extremum is the complete bipartite graph.

1. PRELIMINARIES

A graph $\mathcal{G}$ is a pair $(V(\mathcal{G}), E(\mathcal{G}))$ (or for short (V, E)) where E (the *edge-set*) is a set of pairs of V. (V is called *vertex-set*.) Let S be a subset of vertices. Then $\mathcal{G}[S]$ denotes the subgraph induced by S, and $\mathcal{G}[A, B]$ stands for the induced bipartite subgraph (for $A \cap B = \emptyset$). $\mathcal{K}[A, B]$ denotes the *complete bipartite* graph with parts A and B, and $\mathcal{K}[a, b]$ stands for a complete bipartite graph with $|A| = a$, $|B| = b$. $\mathcal{K}[S]$ is the complete graph with vertex set S. The *neighborhood* of a vertex v is denoted by $N_{\mathcal{G}}(v)$ (or sometimes briefly by $N(v)$), i.e., $N(v) =: \{u \in V: \{u, v\} \in E\}$. Note that $v \notin N(v)$. The size of $N(v)$ is called the *degree* of v, $\deg_{\mathcal{G}}(v)$. The $\deg_{\mathcal{G}}(v, Y)$ stands for $|N(v) \cap Y|$, the number of edges connecting v to Y. The graph $\mathcal{G}$ has *diameter* 2 if it is not the complete graph and for each two vertices $u, v \in V$ either $\{u, v\}$ is an edge of $\mathcal{G}$, or $N(u) \cap N(v) \neq \emptyset$ (or both). $\mathcal{G}$ is called a *minimal* graph of diameter 2 if its diameter is 2, and the deletion of any of its edges spoils this property. Plesník [7] observed that all known minimal graphs of diameter 2 on n vertices have no more than $n^2/4$ edges, and

*New address: Dept. Math., Univ. Illinois, Urbana, IL 61801, USA. Research supported in part by the Hungarian National Science Foundation, Grant No. 1812. This paper was written while the author visited the Institute for Mathematics and its Applications, University of Minnesota, Minneapolis.

Journal of Graph Theory, Vol. 16, No. 1, 81–98 (1992)
 CCC 0364-9024/92/010081-18$04.00

that the complete bipartite graphs are minimal graphs of diameter 2. Independently, Murty and Simon (see in [2]) stated these as the following conjecture:

Conjecture 1.1. If $\mathcal{G}$ is a minimal graph of diameter 2 on n vertices, then $|E(\mathcal{G})| \leq \lfloor n^2/4 \rfloor$. Equality holds if and only if $\mathcal{G}$ is the complete bipartite graph $\mathcal{K}[\lfloor n/2 \rfloor, \lceil n/2 \rceil]$.

Paul Erdös (private communication) informed me, that this conjecture goes back to the 1960s to the work of O. Ore, but the author failed to uncover an exact reference. Let $\mathcal{G}$ be a minimal graph of diameter 2 with n vertices. Plesník [7] proved that $|E(\mathcal{G})| < 3n(n-1)/8$. Caccetta and Häggkvist [2] obtained $|E(\mathcal{G})| < 0.27n^2$. Fan [6] proved affirmatively the first part of the Conjecture 1.1 for $n \leq 24$ and for $n = 26$. For $n \geq 25$ he obtained

$$|E(\mathcal{G})| < \frac{1}{4}n^2 + \frac{n^2 - 16.2n + 56}{320} < 0.2532n^2.$$

An incorrect proof was published [10] in 1984.

Theorem 1.2. Conjecture 1.1 is true for $n > n_0$.

The value of n_0 is explicitly computable, but the proof given here yields a vastly huge number (a tower of 2's of height about 10^{14}).

This paper is organized as follows. In Section 2, a lemma is proved about the number of disjoint neighborhoods in an arbitrary graph. In Section 3, we prove that $|E(\mathcal{G})| < (1 + o(1))n^2/4$ holds for all n. The main idea of the proof is that we delete some $o(n^2)$ edges of $\mathcal{G}$ such that the remaining graph, $\mathcal{G}_0$, has only at most $n^2/4$ edges. In this step we utilize a result of Ruzsa and Szemerédi [8] about triangle-free, 3-uniform hypergraphs. In Sections 4–6 we put back the deleted edges. Then after a lengthy argument, where we repeatedly use the structure of $\mathcal{G}_0$, we conclude that the conjecture is true for sufficiently large n. During the proof we suppose that $|E(\mathcal{G})| \geq (\frac{1}{4} - \delta)n^2$, and prove more and more common properties of $\mathcal{G}$ and a complete bipartite graph $\mathcal{K}[C, D]$ with $C \cup D = V(\mathcal{G})$. In Section 4 it is shown that crit $\mathcal{G}$ contains a huge bipartite graph with edge density $1 - o(1)$. Section 5 contains the proof that almost all vertices in D are connected to almost all vertices of C, i.e., $\mathcal{G}$ is almost a complete bipartite graph. In Section 6 we finish the proof. In Section 7 we have some closing remarks on further open problems.

2. THE NUMBER OF DISJOINT NEIGHBORHOODS IN A GRAPH

Let $\mathcal{F}$ be an arbitrary graph on n vertices. Define the set of pairs with disjoint neighborhoods as follows:

$$E(\text{disj } \mathcal{F}) =: \{\{u,v\}: N_{\mathcal{F}}(u) \cap N_{\mathcal{F}}(v) = \emptyset\}.$$

We treat disj $\mathcal{F}$ as a graph with vertex-set $V(\mathcal{F})$.

Lemma 2.1. $|E(\mathcal{F})| + |E(\text{disj } \mathcal{F})| \le \lfloor n^2/2 \rfloor$.

For the complete bipartite graph $\mathcal{K}[\lfloor n/2 \rfloor, \lceil n/2 \rceil]$ equality holds in Lemma 2.1. There are other extremal examples, e.g., a matching of size $\lfloor n/2 \rfloor$, or for $n = 5$ the disjoint union of an edge and a path of length 2, etc. Moreover, if $\mathcal{F}$ is vertex-disjoint union of complete graphs, then $|E(\mathcal{F})| + |E(\text{disj } \mathcal{F})| \ge \binom{n}{2}$, i.e., it is very close to the optimal one.

Proof. We use induction on n. The cases $n = 1, 2$ are trivial. Suppose that the vertex x has maximum degree, i.e., $|N_{\mathcal{F}}(x)| = \max \deg_{\mathcal{F}}(v)$. If $N_{\mathcal{F}}(x) = \emptyset$, then the left-hand side in Lemma 2.1 is $\binom{n}{2} < \lfloor n^2/2 \rfloor$. So we may suppose that there exists a $y \in N(x)$. For every $z \in N(x) \backslash \{y\}$ we have $x \in N(y) \cap N(z) \neq \emptyset$, hence

$$\deg_{\text{disj}\,\mathcal{F}}(y) \le n - \deg_{\mathcal{F}}(x), \tag{2.2}$$

and by definition

$$\deg_{\mathcal{F}}(y) \le \deg_{\mathcal{F}}(x). \tag{2.3}$$

Summing up (2.2) and (2.3) we have

$$\deg_{\text{disj}\,\mathcal{F}}(y) + \deg_{\mathcal{F}}(y) \le n. \tag{2.4}$$

We distinguish between two subcases.

(1) Suppose first that there exists a $y_0 \in N(x)$ such that the left-hand side of (2.4) is only at most $n - 1$. Let $\mathcal{F}'$ be the graph obtained from $\mathcal{F}$ by deleting the vertex y_0 and the edges through y_0. Obviously

$$|E(\mathcal{F})| = |E(\mathcal{F}')| + \deg_{\mathcal{F}}(y_0), \tag{2.5}$$

and it is easy to see that

$$|E(\text{disj } \mathcal{F})| \le |E(\text{disj } \mathcal{F}')| + \deg_{\text{disj}\,\mathcal{F}}(y_0). \tag{2.6}$$

Summing up (2.5) and (2.6), then using the induction hypothesis for $\mathcal{F}'$ and the assumption for y_0, we obtain that

$$|E(\mathcal{F})| + |E(\text{disj } \mathcal{F})| \le \lfloor (n-1)^2/2 \rfloor + (n-1) \le \lfloor n^2/2 \rfloor.$$

(2) Suppose now that equality holds in (2.4) for every $y \in N(x)$. Then equality holds in (2.2) for all $y \in N(x)$, which implies that the complete bipartite graph $\mathcal{K}[N(x), V(\mathcal{F}) \backslash N(x)]$ is a subgraph of disj $\mathcal{F}$. Consequently,

there is no edge of $\mathcal{F}$ in $N(x)$. Equality holds in (2.3) as well, so $\mathcal{H}[N(x), V(\mathcal{F})\backslash N(x)]$ is a subgraph of $\mathcal{F}$, too. Hence $\mathcal{F} = \mathcal{H}[N(x), V(\mathcal{F})\backslash N(x)]$. Finally, for this graph the left-hand side in Lemma 2.1 is at most $2\lfloor n^2/4 \rfloor$. ∎

3. THE PROOF OF $\max|E(\mathcal{G})| = \frac{1}{4}(1 + o(1))n^2$

Let $\mathcal{G}$ denote a minimal graph of diameter 2 with n vertices. Define the set of *critical pairs* as follows. $\{u, v\} \in E(\text{crit } \mathcal{G})$ if there is a *unique* path of length at most 2 with end points u and v. Call this unique path *critical* path and denote it by $P(u, v)$. There are two cases.

If $P(u, v)$ consists of only a single edge, then we call it *type I.* If $P(u, v)$ consists of two edges, then we call them *type II.* It is possible that an edge of $\mathcal{G}$ has both types. But the minimality of $\mathcal{G}$ ensures that every edge has at least one type, i.e., every edge belongs to a critical path. For an edge $E \in E(\mathcal{G})$, denote $m(E)$ the *multiplicity* of E, i.e., the number of critical paths in which the edge E appears.

Lemma 3.1. For any $m > 0$ the number of edges of $\mathcal{G}$ with multiplicity at least m is at most $n(n - 1)/m$.

Proof. The total sum of multiplicities is at most twice the number of critical pairs, i.e., it is at most $2\binom{n}{2}$. ∎

An upper bound on the number of light paths. Let m be an arbitrary positive number. A critical path is called *light* if it has two edges, and both have multiplicity less than m. We are going to give an upper bound (depending on n and m) for the number of light paths. To do this we recall some definitions and results from the extremal hypergraph theory.

A *3-graph* (or 3-uniform hypergraph) $\mathcal{H}$ is a pair $(V(\mathcal{H}), E(\mathcal{H}))$, where $V(\mathcal{H})$ is a finite set (the set of *vertices*), and $E(\mathcal{H})$ is a set of 3-element subsets of $V(\mathcal{H})$ (the set of *edges*). $\mathcal{H}$ is called *linear* if every two distinct edges intersect in at most 1 element. Three edges of a hypergraph form a *triangle* if they pairwise intersect, but no vertex is contained in all the three of them. For example, a triangle in a linear 3-graph is isomorphic to $\{\{1,2,3\}, \{3,4,5\}, \{5,6,1\}\}$. Denote by $\text{RSz}(n)$ the largest number of edges in a triangle-free, linear 3-graph over n vertices. Ruzsa and Szemerédi proved the following theorem:

Theorem 3.2 [8]. $\text{RSz}(n) = o(n^2)$.

(Actually, they also proved that $\text{RSz}(n)$ is larger than n^{2-c} for all positive c, but we need the upper bound only.)

Lemma 3.3 The number of light paths is less than $27m\mathrm{RSz}(n)$.

Proof. Define the 3-graph $\mathcal{H}^1$ with vertex-set $V(\mathcal{G})$ as the set of 3-element sets determined by the light critical paths of $\mathcal{G}$. Consider a light critical path $P(u,v) =: \{\{u,c\},\{c,v\}\}$. The critical pair $\{u,v\}$ does not appear in any other triples from $\mathcal{H}^1$, so there are at most $2(m-1)$ further triples intersecting $\{u,c,v\}$ in 2 elements. Keeping the triple $\{u,c,v\}$ and deleting those from $\mathcal{H}^1$ that intersect it in 2 elements, then continuing this process until no two triples left with intersection size 2, one obtains a linear hypergraph $\mathcal{H}^2$ such that

$$E(\mathcal{H}^2) \subset E(\mathcal{H}^1) \quad \text{and} \quad |E(\mathcal{H}^2)| \geq |E(\mathcal{H}^1)|/(2m-1). \tag{3.4}$$

A 3-graph $\mathcal{H}$ is called *3-partite,* if one can partition its vertex-set $V(\mathcal{H}) = V_1 \cup V_2 \cup V_3$ such that for every edge $E \in E(\mathcal{H})$ and for all i $(1 \leq i \leq 3)$ one has $|E \cap V_i| = 1$. Erdös and Kleitman proved the following simple but important fact:

Fact 3.5 [5]. Let $\mathcal{H}$ be an arbitrary r-graph. Then one can find an r-partite subhypergraph $\mathcal{H}'$ of it such that

$$|E(\mathcal{H}')| \geq \frac{r!}{r^r}|E(\mathcal{H})|.$$

Applying Fact 3.5 to $\mathcal{H}^2$, one obtains a 3-partite, linear hypergraph $\mathcal{H}^3$ with parts V_1, V_2, V_3, such that

$$|E(\mathcal{H}^3)| \geq \frac{2}{9}|E(\mathcal{H}^2)|. \tag{3.6}$$

Let $P(u,v)$ be a critical path with edges $\{u,c\}$ and $\{c,v\}$. The vertex c is called the *center* of the triple $\{u,c,v\}$. Without loss of generality we may suppose that at least 1/3 of the triples of $\mathcal{H}^3$ have its center in V_2. This means, that there is a subhypergraph $\mathcal{H}^4$ of $\mathcal{H}^3$ such that

$$|E(\mathcal{H}^4)| \geq \frac{1}{3}|E(\mathcal{H}^3)|, \tag{3.7}$$

and with the additional property that if $\{v_1, v_2, v_3\}$ is a triple of $\mathcal{H}^4$ with $v_i \in V_i$ then $\{v_1, v_3\}$ is its critical pair.

Proposition 3.8. $\mathcal{H}^4$ is triangle-free.

Proof. Suppose, to the contrary, that three triples P_1, P_2, P_3 of $\mathcal{H}^4$ form a triangle. Then $P_1 \cup P_2 \cup P_3$ intersects V_i $(1 \leq i \leq 3)$ in at least 2 elements. As $|P_1 \cup P_2 \cup P_3| = 6$, we obtain that each V_i contains exactly two vertices from the triangle. Let $(P_1 \cup P_2 \cup P_3) \cap V_i =: \{a_i, b_i\}$ and $P_1 =: \{a_1, a_2, a_3\}$. Without loss of generality, we may suppose that P_i intersects P_1 in a_i, i.e., $P_2 =: \{b_1, a_2, b_3\}$ and $P_3 =: \{b_1, b_2, a_3\}$. Then (b_1, b_2, a_3) and (b_1, a_2, a_3) are two disjoint paths from b_1 to a_3, which contradicts the earlier constraint that $\{b_1, a_3\}$ is a critical pair. ∎

The End of the Proof of Lemma 3.3. Proposition 3.8 and Theorem 3.2 imply that $|E(\mathcal{H}^4)| \leq \mathrm{RSz}(n)$, and (3.4), (3.6), and (3.7) imply that $|E(\mathcal{H}^1)| \leq 27m|E(\mathcal{H}^4)|$. ∎

The asymptotic upper bound on $|E(\mathcal{G})|$. Let $m =: \sqrt{n^2/54\mathrm{RSz}(n)}$. Note that $m =: m(n)$ tends to infinity according to Theorem 3.2. Delete all edges of $\mathcal{G}$ whose multiplicity is at least m, and those edges that appear in a light critical path. Denote the obtained graph by $\mathcal{G}_0$. Lemmas 3.1 and 3.3 imply the following upper bound on the number of deleted edges:

$$|E(\mathcal{G})| \leq |E(\mathcal{G}_0)| + \frac{n(n-1)}{m} + 54m\mathrm{RSz}(n) < |E(\mathcal{G}_0)| + \frac{2n^2}{m}. \qquad (3.9)$$

Deleting these edges from $\mathcal{G}$, we have destroyed all critical paths of length 2. In other words, if (u, c, v) is a critical path in $\mathcal{G}$ (with critical pair $\{u, v\}$), then the neighborhoods of u and v in $\mathcal{G}_0$ are disjoint. This implies that $E(\text{crit } \mathcal{G}) \subset E(\text{disj } \mathcal{G}_0)$, i.e.,

$$|E(\text{crit } \mathcal{G})| \leq |E(\text{disj } \mathcal{G}_0)|. \qquad (3.10)$$

As the edge-set of $\mathcal{G}$ is the union of critical paths, and after the deletion every $V(P(u, v))$ contains at most one edge of $\mathcal{G}_0$, we conclude that the number of edges in $\mathcal{G}_0$ is not more than the number of critical pairs in $\mathcal{G}$, i.e.,

$$|E(\mathcal{G}_0)| \leq |E(\text{crit } \mathcal{G})|. \qquad (3.11)$$

The inequalities (3.10) and (3.11) imply, together with Lemma 2.1, that

$$|E(\mathcal{G}_0)| \leq \frac{1}{2}(|E(\mathcal{G}_0)| + |E(\text{disj } \mathcal{G}_0)|) \leq n^2/4. \qquad (3.12)$$

Finally, (3.12) and (3.9) give

Corollary 3.13. $|E(\mathcal{G})| \leq (n^2/4) + (2n^2/m) = (1 + o(1))(n^2/4)$. ∎

4. crit $\mathcal{G}$ HAS A GIANT BIPARTITE SUBGRAPH

We continue the proof started in the previous section. Suppose that $|E(\mathcal{G})| > (\frac{1}{4} - \delta)n^2$ for some $\delta \leq 10^{-6}$. As $\mathrm{RSz}(n)/n^2 \to 0$ there exists an n_0

such that for $n > n_0$ one has

$$\mathrm{RSz}(n) \leq 10^{-13} n^2. \tag{4.1}$$

(The proof given in [8] implies that n_0 is not larger than a power tower of 2's of height 10^{14}.) Inequality (4.1) implies that $m > 400{,}000$. Define ε such that $\varepsilon^2/4 > \delta + (2/m)$, e.g., $\varepsilon = 1/200$.

We have four graphs, $E(\mathcal{G}_0) \subset E(\mathcal{G})$ and $E(\mathrm{crit}\ \mathcal{G}) \subset E(\mathrm{disj}\ \mathcal{G}_0)$. Equation (3.9) implies that

$$|E(\mathcal{G}_0)| \geq |E(\mathcal{G})| - \frac{2}{m} n^2 > (1 - \varepsilon^2)\frac{n^2}{4}. \tag{4.2}$$

The huge values of n_0 and m were needed to satisfy the inequalities $\varepsilon \leq 0.005$ and (4.2). In the proof we use only these two constraints. The value of n_0 probably can be lowered.

First we formulate the fact that crit $\mathcal{G}$ and disj $\mathcal{G}_0$ are *close* to each other. Equations (4.2) and (3.11), then (3.10), and finally (3.12) imply that

$$(1 - \varepsilon^2)\frac{n^2}{4} < |E(\mathrm{crit}\ \mathcal{G})| \leq |E(\mathrm{disj}\ \mathcal{G}_0)| < (1 + \varepsilon^2)\frac{n^2}{4}. \tag{4.3}$$

Hence we have $|E(\mathrm{disj}\ \mathcal{G}_0)| - |E(\mathrm{crit}\ \mathcal{G})| < \varepsilon^2 n^2/2$. Let $A_4 =: \{u \in V\text{: not satisfying (4.4)}\}$:

$$\deg_{\mathrm{disj}\,\mathcal{G}_0}(u) - \deg_{\mathrm{crit}\,\mathcal{G}}(u) < \varepsilon n. \tag{4.4}$$

Note that the left-hand side of (4.4) is always nonnegative. This implies that

$$\begin{aligned} |A_4|\varepsilon n &\leq \sum_{u \in A_4} (\deg_{\mathrm{disj}\,\mathcal{G}_0}(u) - \deg_{\mathrm{crit}\,\mathcal{G}}(u)) \\ &\leq 2(|E(\mathrm{disj}\ \mathcal{G}_0)| - |E(\mathrm{crit}\ \mathcal{G})|) < \varepsilon^2 n^2, \end{aligned}$$

i.e., $|A_4| < \varepsilon n$.

Proposition. For all but less than εn vertices $u \in V$, the following holds:

$$\deg_{\mathcal{G}_0}(u) + \deg_{\mathrm{disj}\,\mathcal{G}_0}(u) > (1 - \varepsilon)n. \tag{4.5}$$

Proof. Let A_5 denote the set of exceptional vertices, i.e., the set of vertices $u \in V$, which does not fulfill (4.5). Suppose, to the contrary, that $|A_5| \geq \varepsilon n$, and let $B \subset A_5$, $|B| =: b = \lceil \varepsilon n \rceil$. Delete B. Then for the obtained graph $\mathcal{G}_0 \backslash B$ we can apply Lemma 2.1.

$$|E(\mathcal{G}_0)| + |E(\text{disj } \mathcal{G}_0)|$$
$$\leq |E(\mathcal{G}_0 \backslash B)| + |E(\text{disj}(\mathcal{G}_0 \backslash B))| + \sum_{u \in B} \deg_{\mathcal{G}_0}(u) + \deg_{\text{disj } \mathcal{G}_0}(u)$$
$$\leq (n - b)^2/2 + b\lfloor (1 - \varepsilon)n \rfloor .$$

Here the right-hand side is at most $\frac{1}{2}(1 - \varepsilon^2)n^2$ and the left-hand side is at least $2|E(\mathcal{G}_0)|$ (by (3.12)). This contradicts (4.2). ∎

Let v be a vertex with maximum degree in $\mathcal{G}_0$, i.e., $d =: \deg_{\mathcal{G}_0}(v)$, and for all u we have $\deg_{\mathcal{G}_0}(u) \leq d$. Denote $N_{\mathcal{G}_0}(v)$ by D, and its complement $V(\mathcal{G}) \backslash D$ by C. By (4.2) we have

$$d > \frac{n}{2}(1 - \varepsilon^2). \tag{4.6}$$

No edge of disj $\mathcal{G}_0$ is contained in D; hence

$$\deg_{\text{disj } \mathcal{G}_0}(y) \leq n - d \tag{4.7}$$

holds for all $y \in D$. This and (4.5) implies that

$$\deg_{\mathcal{G}_0}(y) > d - \varepsilon n \tag{4.8}$$

holds for all $y \in D \backslash A_5$.

Proposition 4.9. $d < 0.75n$.

Proof. Equation (4.2) and the above inequality (4.7) imply

$$\frac{n^2}{4}(1 - \varepsilon) < |E(\mathcal{G}_0)| \leq |E(\text{disj } \mathcal{G}_0)| \leq \frac{1}{2}\left(\sum_{x \in C \cup D} \deg_{\text{disj } \mathcal{G}_0}(x) \right)$$
$$\leq \frac{1}{2}((n - d)(n - 1) + d(n - d)) < \frac{1}{2}(n^2 - d^2).$$

This leads a contradiction if $d \geq 0.75n$ and ε is small. ∎

Moreover, (4.5) imply that $\deg_{\text{disj } \mathcal{G}_0}(y) > n - d - \varepsilon n$ for all $y \in V \backslash A_5$. As disj $\mathcal{G}_0 | D$ is empty, we have the following lower bound for the number of disj $\mathcal{G}_0$ edges connecting y to C:

$$\deg_{\text{disj } \mathcal{G}_0}(y, C) > n - d - \varepsilon n \tag{4.10}$$

holds for all $y \in D \backslash A_5$. In other words disj $\mathcal{G}_0[C, D]$ is almost a complete bipartite graph.

Consider the bipartite subgraph of crit $\mathcal{G}$, induced by C and D, i.e., the edge-set is defined by $E(\text{crit } \mathcal{G}(C, D)) =: \{\{x, y\} \in E(\text{crit } \mathcal{G}): x \in C, y \in D\}$. The inequalities (4.10) and (4.4) imply

$$|N_{\text{crit } \mathcal{G}}(y) \cap C| > n - d - 2\varepsilon n \tag{4.11}$$

holds for all $y \in D \backslash (A_4 \cup A_5)$. So crit $\mathcal{G}[C, D]$ is almost a complete bipartite graph, as well.

Let A_6 be the set of vertices $x \in C$ having at most $6\varepsilon n$ crit $\mathcal{G}$ neighbors in D.

Proposition 4.12. $|A_6| < 5\varepsilon n$.

Proof. Equation (4.11) implies that

$$\begin{aligned}|A_6|(d - 6\varepsilon n) &\leq \sum_{x \in A_6} (d - \deg_{\text{crit } \mathcal{G}}(x, D)) \leq \sum_{x \in C} (d - \deg_{\text{crit } \mathcal{G}}(x, D)) \\ &= \sum_{y \in D} (n - d - \deg_{\text{crit } \mathcal{G}}(y, C)) \leq |A_4 \cup A_5|(n - d) \\ &\quad + (d - |A_4 \cup A_5|)2\varepsilon n \leq 2\varepsilon n^2 - 4\varepsilon^2 n^2.\end{aligned}$$

This and (4.6) imply 4.12. ∎

5. $\mathcal{G}$ HAS A GIANT BIPARTITE SUBGRAPH

It is impossible that for some vertex u both

$$|N_{\mathcal{G}_0}(u) \cap C| \geq 2\varepsilon n \quad \text{and} \quad |N_{\mathcal{G}_0}(u) \cap D| \geq 2\varepsilon n \tag{5.1}$$

hold. Suppose, to the contrary, that (5.1) holds for some $u \in V$. Since $|A_4 \cup A_5| < 2\varepsilon n$, there is y from $N_{\mathcal{G}_0}(u) \cap D \backslash (A_4 \cup A_5)$. By (4.11), at least $n - d - 2\varepsilon n$ edges of crit $\mathcal{G}$ adjacent to y go into C. Noting that $|N_{\mathcal{G}_0}(u) \cap C| \geq 2\varepsilon n$, there exists an edge $\{x, y\}$ of crit $\mathcal{G}$ with $x \in N_{\mathcal{G}_0}(u) \cap C$. But then (x, u, y) is the critical path belonging to the critical pair $\{x, y\}$, which contradicts the definition of $\mathcal{G}_0$. ∎

Call a vertex u of *type C* (or *D*) if it has at least $2\varepsilon n$ $\mathcal{G}_0$ neighbors in C (in D, respectively). Eventually, a vertex with a small $\mathcal{G}_0$ degree has no type. But as every $\mathcal{G}_0$ degree in $D \backslash A_5$ is between d and $d - \varepsilon n$, by (4.8), we obtain that they have types. The aim of this section is to show the following lemma, which leads to the fact that $\mathcal{G}$ is (almost) a complete bipartite graph. Let D_0 be the set of vertices $y \in D \backslash A_5$ with type D, $d_0 =: |D_0|$.

Lemma 5.2. $|D_0| < 3\varepsilon n$.

Proof. Suppose, to the contrary, that $|D_0| \geq 3\varepsilon n$. The first step of the proof of 5.2 is to show that

$$|D_0| > d - 5\varepsilon n. \tag{5.3}$$

Proof of (5.3). Consider the bipartite graph $\mathscr{G}_0(D_0, D\backslash(D_0 \cup A_5))$. If $y \in D\backslash(D_0 \cup A_5)$ (i.e., it has type C), then $|N_{G_0}(y) \cap D_0| \leq |N_{\mathscr{G}_0}(y) \cap D| < 2\varepsilon n$. Hence

$$|E(\mathscr{G}_0[D_0, D\backslash(D_0 \cup A_5)])| \leq |D\backslash(D_0 \cup A_5)|2\varepsilon n. \tag{5.4}$$

On the other hand, every point in $y \in D_0$ has type C, i.e., $\deg_{\mathscr{G}_0}(y, C) < 2\varepsilon n$. So y has more than $\deg_{\mathscr{G}_0}(y) - 2\varepsilon n$ $\mathscr{G}_0$ neighbors in D. Hence, by (4.8),

$$\deg_{\mathscr{G}_0}(y, D) > |D| - 3\varepsilon n \tag{5.5}$$

for all $y \in D_0$. So y has more than $|D\backslash(D_0 \cup A_5)| - 3\varepsilon n$ neighbors in $D\backslash(D_0 \cup A_5)$. Thus,

$$|D_0|(|D| - |D_0 \cup A_5| - 3\varepsilon n) < |E(\mathscr{G}_0[D_0, D\backslash(D_0 \cup A_5)])|. \tag{5.6}$$

Rearranging (5.4) and (5.6), and using the fact $|A_5 \cap D| < \varepsilon n$, we have

$$d_0 3\varepsilon n > (d - d_0 - \varepsilon n)(d_0 - 2\varepsilon n).$$

This implies (5.3) since $d_0 \geq 3\varepsilon n$ and $d > 18\varepsilon n$. ∎

Consider the induced graph $\mathscr{G}_0 | D_0$. Equation (5.5) implies that every vertex from D_0 has at least $|D_0| - 3\varepsilon n$ $\mathscr{G}_0$ neighbors in D_0, and $d_0 - 3\varepsilon n > 2d_0/3$, by (5.3). Hence every two vertices of D_0 have at least $d_0/3$ common neighbors in $\mathscr{G}_0$. So in this case D_0 (and by (5.3) D) induce almost a complete graph (in $\mathscr{G}_0$). Consequently, D_0 does not contain any edge from disj $\mathscr{G}_0$, from crit $\mathscr{G}$ and there is no edge of $\mathscr{G}$ in D_0 of type I.

Consider the induced bipartite subgraph $\mathscr{G}[C\backslash A_6, D_0]$. (Warning! $\mathscr{G}$ and not $\mathscr{G}_0$.)

Proposition 5.7. Let $x \in C\backslash A_6$. Then $\deg_{\mathscr{G}}(x, D_0) \leq 1$, i.e., there is at most 1 edge of $\mathscr{G}$ from x to D_0.

Proof of 5.7. Suppose that there are 2 such edges of $\mathscr{G}$ $\{x, y_1\}$ and $\{x, y_2\}$ with $y_1, y_2 \in D_0$. We have that $|D\backslash N_{\mathscr{G}}(y_1) \cap N_{\mathscr{G}}(y_2)|$ is at most $6\varepsilon n$ by (5.5). So there exists a critical edge $\{x, z\} \in E(\text{crit } \mathscr{G})$ with $z \in D \cap N_{\mathscr{G}}(y_1) \cap N_{\mathscr{G}}(y_2)$, as $x \notin A_6$. Then (x, y_1, z) and (x, y_2, z) are two distinct paths in $\mathscr{G}$. But this contradicts to the criticality of $\{x, z\}$. ∎

Let F be the set of those vertices in D_0 that are not connected to $C\backslash A_6$ in $\mathcal{G}$.

Claim 5.8. $|F| < 23\varepsilon n$.

Proof of 5.8. We obtain 5.8 by estimating $|E(\mathcal{G}[F])|$. Let $|F| =: f$ and suppose $f > 0$. The number of $\mathcal{G}_0$ edges in F is at least

$$\frac{1}{2} f(f - 3\varepsilon n), \tag{5.9}$$

by (5.5). The set F (and D_0) contains no critical pair. So each of the edges of $\mathcal{G}|F$ has type II, so they belong to critical paths of length 2. Hence

$$|E(\mathcal{G}[F])| \le \sum_{z \in F} \deg_{\text{crit}\,\mathcal{G}}(z). \tag{5.10}$$

If $\{y, z\}$ is an edge of $\mathcal{G}$ contained in F, then let (x, y, z) be a critical path containing it, $\{x, z\} \in \text{crit}\,\mathcal{G}$ and $\{x, y\} \in E(\mathcal{G})$. By definition, $C\backslash A_6$ is not connected (in $\mathcal{G}$) to F, so we have that $x \notin C\backslash A_6$ (and $x \notin D_0$). Now we can easily give an upper bound on the number of critical pairs $\{x, z\}$, such that $x \notin D_0 \cup (C\backslash A_6)$ and $z \in F$. This is at most

$$(|A_6| + |D\backslash D_0|)\,|F| < 10\varepsilon n f \tag{5.11}$$

by 4.12 and (5.3). Then (5.9)–(5.11) imply $f < 23\varepsilon n$. ∎

The end of the proof of 5.2. Denote $D_0\backslash(F \cup A_4)$ by D_1. For every $y \in D_1$, let $x =: x(y)$ be a vertex of $C\backslash A_6$ such that $\{x, y\} \in E(\mathcal{G})$. By Proposition 5.7 these second end points are all distinct. Let $C_1 =: \{x(y): y \in D_1\}$. Then (5.3) and Claim 5.8 imply that

$$\begin{aligned} \frac{n}{2}(1 + \varepsilon^2) > n - d \ge |C_1| \\ = |D_1| = d_0 - f - |A_4| > (d - 5\varepsilon n) - 23\varepsilon n - \varepsilon \\ > \frac{n}{2} - 30\varepsilon n. \end{aligned} \tag{5.12}$$

Let $|D_1| =: d_1$. Consider two arbitrary edges $\{x, y\}$ and $\{x', y'\}$ of $\mathcal{G}$ between C_1 and D_1 with $x, x' \in C_1$ and $y, y' \in D_1$. If $\{x', y\}$ is a critical pair, then either $\{x, x'\}$ or $\{y, y'\}$ is not in $E(\mathcal{G})$. This, and the structure of $\mathcal{G}[C_1, D_1]$, imply that

$$\deg_{\mathcal{G}}(x, C_1) + \deg_{\mathcal{G}}(y, D_1) + \deg_{\text{crit}\,\mathcal{G}}(y, C_1) \le 2(d_1 - 1).$$

By (4.11) we have that $y \in D_1$ has at least $d_1 - 2\varepsilon n$ crit $\mathcal{G}$ neighbors in C_1. This and the above inequality imply

$$\deg_{\mathcal{G}}(x, C_1) + \deg_{\mathcal{G}}(y, D_1) \le d_1 - 2 + 2\varepsilon n. \tag{5.13}$$

Summing up (5.13) for the d_1 edges of $\mathcal{G}$ connecting C_1 and D_1, we obtain that

$$|E(\mathcal{G}[C_1 \cup D_1])| = d_1 + |E(\mathcal{G}[C_1])| + |E(\mathcal{G}[D_1])| \le \frac{1}{2} d_1(d_1 + 2\varepsilon n). \tag{5.14}$$

It is obvious that the number of edges of $\mathcal{G}_0$ not included in $C_1 \cup D_1$ is not more than

$$(n - 2d_1)d. \tag{5.15}$$

Finally, the sum of (5.14) and (5.15) gives

$$|E(\mathcal{G}_0)| \le \frac{1}{2} d_1^2 + (\varepsilon n - 2d)d_1 + nd =: g(d_1).$$

As the function $g(d_1)$ is monotone decreasing for $d_1 < d$, and $d_1 \ge d - 29\varepsilon n$ by (5.12), we have

$$\begin{aligned} |E(\mathcal{G}_0)| \le g(d - 29\varepsilon n) &= -\frac{3}{2} d^2 + (1 + 30\varepsilon) nd + 391.5\varepsilon^2 n^2 \\ &< -1.5d^2 + 1.15nd + 0.01n^2. \end{aligned}$$

We used that $\varepsilon \le 0.005$. Here the right-hand side is less than $0.24n^2$ for all real d. This contradicts (4.2). ▮

6. THE END OF THE PROOF

Claim 6.1. Every $y \in D \setminus (A_4 \cup A_5 \cup D_0)$ is connected by type I edges of $\mathcal{G}$ to all but less than $6\varepsilon n$ vertices in C, i.e., $\deg_{\text{type I}}(y, C) > |C| - 6\varepsilon n$.

Note that Lemma 5.2 (and (4.4) and (4.5)) imply that $|A_4 \cup A_5 \cup D_0| < 5\varepsilon n < |D|$.

Proof. The $\deg_{\mathcal{G}_0}(y) > d - \varepsilon n$ by (4.8), as $y \notin A_5$. $|N_{\mathcal{G}_0}(y) \cap D| < 2\varepsilon n$ by (5.1), as $y \notin D_0 \cup A_5$. Hence

$$\deg_{\mathcal{G}_0}(y, C) > d - 3\varepsilon n. \tag{6.2}$$

As $d > n - d - \varepsilon n$ by (4.6), we have that the right-hand side of (6.2) is at least $|C| - 4\varepsilon n$. Moreover, $\deg_{\text{crit}\,\mathcal{G}}(y, C) > |C| - 2\varepsilon n$ by (4.11), as $y \notin (A_4 \cup A_5)$. So y is connected to more than $|C| - 6\varepsilon n$ vertices in C with $\mathcal{G}_0$ (so $\mathcal{G}$) edges, such that these edges are also critical pairs. That is, these are edges of type I. ∎

As a byproduct, (6.2) implies that $n - d = |C| > d - 3\varepsilon n$; hence

$$d < \frac{n}{2} + \frac{3}{2}\varepsilon n. \tag{6.3}$$

Let D_7 be the set of vertices $y \in D$ with at least $\frac{1}{4}n + \varepsilon n$ type I neighbors in C, i.e.,

$$D_7 =: \left\{ y \in D\colon \deg_{\text{type I}}(y, C) \geq \frac{n}{4} + \varepsilon n \right\}.$$

Analogously, let

$$C_7 =: \left\{ x \in C\colon \deg_{\text{type I}}(x, D) \geq \frac{n}{4} + \varepsilon n \right\}.$$

As ε is small, Claim 6.1 implies that $D_7 \supset D \backslash (A_4 \cup A_5 \cup D_0)$, hence

$$|D \backslash D_7| < 5\varepsilon n. \tag{6.4}$$

Again Claim 6.1 and a simple counting argument yields that

$$|C \backslash C_7| < 22\varepsilon n.$$

Indeed, like in (4.11), we have

$$\begin{aligned} |C \backslash C_7| \left(d - \frac{n}{4} - \varepsilon n \right) &\leq \sum_{x \in C \backslash C_7} (d - \deg_{\text{type I}}(x, D)) \\ &\leq \sum_{x \in C} (d - \deg_{\text{type I}}(x, D)) \\ &= \sum_{y \in D} (n - d - \deg_{\text{type I}}(y, C)) \\ &\leq |A_4 \cup A_5 \cup D_0|(n - d) \\ &\quad + (d - |A_4 \cup A_5 \cup D_0|) 6\varepsilon n \leq 5\varepsilon n^2 \\ &\quad + \varepsilon n d - 30\varepsilon^2 n^2. \end{aligned}$$ ∎

Let $A_7 =: (C \backslash C_7) \cup (D \backslash D_7)$. Summarizing the above inequalities,

$$|A_7| < 27\varepsilon n. \tag{6.5}$$

From now on we will not deal with $\mathscr{G}_0$; we return to investigate directly $\mathscr{G}$. The following proposition is implied by Claim 6.1 in the same way as (5.1) is implied by (4.11):

Proposition 6.6. For every $u \in V$ either $|N_{\mathscr{G}}(u) \cap C| < 6\varepsilon n$ or $|N_{\mathscr{G}}(u) \cap D| < 5\varepsilon n$ holds (or both). ∎

Split A_7 into three parts. Let A_8 consist of those vertices of A_7 whose degree is less than $\frac{1}{4}n + 26\varepsilon n$. Let C_8 (D_8) consist of those vertices of $A_7 \backslash A_8$ that have at least $6\varepsilon n$ $\mathscr{G}$ neighbors in $D \backslash A_7$ (in $C \backslash A_7$, respectively). Note that C_8 is not necessarily a subset of C. Proposition 6.6 implies that $C_8 \cap D_8 = \emptyset$, and A_8, C_8, and D_8 form a partition of A_7.

For $x \in C_8$ $\deg(x) \geq (n/4) + 26\varepsilon n$ and $|N(x) \cap C| < 6\varepsilon n$, so we have

$$\deg(x, D) > \frac{n}{4} + 20\varepsilon n, \quad \text{and similarly,} \quad \deg(y, C) > \frac{n}{4} + 20\varepsilon n \tag{6.7}$$

for every $y \in D_8$. (Here and from now on deg and N simply means $\deg_{\mathscr{G}}$ and $N_{\mathscr{G}}$ unless otherwise stated.)

Our next aim is to give an upper bound for $|E(\mathscr{G})|$ using the above partition $V = C_7 \cup C_8 \cup D_7 \cup D_8 \cup A_8$. Obviously, we have that the number of edges adjacent to A_8 is

$$|\{e \in E(\mathscr{G}) \colon e \cap A_8 \neq \emptyset\}| \leq \left(\frac{n}{4} + 26\varepsilon n\right)|A_8|. \tag{6.8}$$

For brevity use the notations $C' =: C_7 \cup C_8$, $c' =: |C'|$ and $D' =: D_7 \cup D_8$, $d' =: |D'|$. As for arbitrary $a, b \in C'$ we have $|N(a, D)| + |N(b, D)| > \frac{1}{2}n + 2\varepsilon n \geq |D| + 2$, we have

Proposition 6.9. C' (and similarly D') contains no critical pair. ∎

As for arbitrary $a \in C_7$ and $b \in C'$, we have

$$|N_{\text{type I}}(a, D)| + |N_{\mathscr{G}}(b, D)| > \frac{1}{2}n + 2\varepsilon n \geq |D| + 2.$$

This implies

Proposition 6.10. There is no $\mathscr{G}$ edge connecting C_7 to C_8, and there is no edge in C_7. Similarly, $E(\mathscr{G}[D_7]) = \emptyset$, and $E(\mathscr{G}[D_7, D_8]) = \emptyset$.

Classify the edges of $\mathcal{G}$ in $C' \cup D'$ as follows:

(i) First of all we have the edges connecting C' and D'.
(ii/C) In this class we have those edges $\{a, b\}$ that are contained in C' (so in C_8 by 6.10) and are part of a critical path (a, b, c) with $c \in D'$.
(ii/D) The definition is analogous to (ii/C), i.e., $\{a, b\} \in E(\mathcal{G})$ belongs to this class if $a, b \in D_8$ and there is a critical path (a, b, c) with $c \in C'$.
(iii) The rest of the edges are in $C' \cup D'$.

First we prove an upper bound for the edges of type (iii). Consider an edge $\{a, b\}$ of type (iii). Say it is included in C_8. As a and b have a lot of common neighbors in D' (by 6.9), the type of $\{a, b\}$ is II. Then it is a part of a critical path (a, b, c) of length 2, and by definition, $c \notin D'$. C' does not contain the critical pair $\{a, c\}$ by 6.9, so $c \notin C'$, too. So $\{a, b\}$ belongs to a critical pair $\{a, c\}$ with $c \in A_8$. As $a \in A_7 \backslash A_8$, the number of such critical pairs is bounded above by $|A_7 \backslash A_8|\,|A_8|$; hence

$$\#(iii) \leq (|A_7| - |A_8|)\,|A_8|. \tag{6.11}$$

For each edge $\{a, b\}$ from the classes (ii/C) or (ii/D), one can associate a critical pair $\{a, c\}$ such that one of a and c lies in C' and the other lies in D', but $\{a, c\}$ is not an edge of $\mathcal{G}$. The pair $\{a, c\}$ is not associated to another edge of type (ii), so in this way we have that the number of edges in class (ii) is not more than the number of nonedges between C' and D'. In other words, the number of edges of types (i) and (ii) is at most $|D'|\,|C'|$. This and (6.8) and (6.11) give

$$|E(\mathcal{G})| \leq d'c' + \left(\frac{n}{4} + 26\varepsilon n + |A_7| - |A_8|\right)|A_8|. \tag{6.12}$$

Here $\varepsilon < 1/4$, so $|A_7| < (n/4) + 26\varepsilon n$ by (6.5). This implies that

$$\left(\frac{n}{4} + 26\varepsilon n + |A_7| - |A_8|\right)|A_8| \leq \left(\frac{n}{4} + 26\varepsilon n\right)|A_7| \leq d'(n - c' - d').$$

In the last step we used that $d' > d - 5\varepsilon n$ by 6.1, and this is larger than $(n/2) - 6\varepsilon n > (n/4) + 26\varepsilon n$. So the right-hand side of (6.12) is at most $d'(n - d') \leq \lfloor n^2/4 \rfloor$, as desired. ∎

Equality can hold in (6.12) only if $n - c' - d' = 0$, i.e., $A_8 = \emptyset$. Then there is no edge of type (iii) by (6.11).

Moreover, every nonedge between C' and D' must be a critical pair. There is no edge between C_7 and C_8 (and between D_7 and D_8) by 6.10, so for every critical pair $e \in E(\text{crit}\ \mathcal{G})$ that is a nonedge, we have $e \cap (C_8 \cup D_8) \neq \emptyset$. So $\mathcal{K}[C_7, D_7]$ is a subgraph of $\mathcal{G}$.

Suppose that $|E(\mathcal{G}[C_8])| \geq |E(\mathcal{G}[D_8])|$; the other case is similar. Let $P =: \{u \in C_8\colon \deg_{\mathcal{G}[C_8]}(u) > 0\}, p =: |P|$. Then (6.5) gives

$$|E(\mathcal{G}[C_8])| + |E(\mathcal{G}[D_8])| \leq p(p-1) < 27\varepsilon np. \tag{6.13}$$

Let $x \in P$ be chosen arbitrarily, and let $\{x, y\}$ be an edge of $\mathcal{G}$ contained in C_8. Then, by (6.4) and (6.7), $\deg(x, D_7) > \deg(x, D) - 5\varepsilon n > (n/4) + 15\varepsilon n$. Similarly, $\deg(y, D) > (n/4) + 15\varepsilon n$, so

$$|N(x) \cap N(y) \cap D_7| \geq 2\left(\frac{n}{4} + 15\varepsilon n\right) - d > 28\varepsilon n.$$

This yields at least $28\varepsilon n$ edges $\{x, z\}$, $z \in D'$ of type II. So the number of edges of type II between C_8 and D' is at least $28\varepsilon np$. Each of such edge is a part of a critical path of length two, with a critical pair between C' and D' (by 6.9). So the number of nonedges between C' and D' is much more than the right-hand side of (6.13), if $p > 0$. Thus $|E(\mathcal{G})| \geq \lfloor n^2/4 \rfloor$ implies that $p = 0$. That is, $\mathcal{G}$ is a bipartite graph, and then a complete bipartite one. ∎

7. REMARKS AND PROBLEMS

We can construct a large nonbipartite minimal graph $\mathcal{M}$ of diameter 2 as follows: Let $V(\mathcal{M}) =: X \cup Y \cup \{z\}$, where $|X| =: \lfloor (n-1)/2 \rfloor$, $|Y| =: \lceil (n-1)/2 \rceil$, and let $x \in X, y \in Y$. The graph $\mathcal{M}$ obtained from the complete bipartite graph $\mathcal{K}[X, Y]$ by deleting the edge $\{x, y\}$ and adding the edges $\{x, z\}$ and $\{z, y\}$. With a little more effort, the above proof gives the following slightly stronger statement:

Theorem 7.1. Suppose that $\mathcal{G}$ is a minimal graph of diameter 2 over n elements, $n > n_0$. If $|E(\mathcal{G})| \geq \lfloor (n-1)^2/4 \rfloor + 1$, then either $\mathcal{G}$ is a complete bipartite graph, or it is isomorphic to $\mathcal{M}$. ∎

Let now $\mathcal{G}$ be an arbitrary graph with n vertices. Let k be an integer and define $\mathrm{disj}_k\, \mathcal{G}$ as follows: The pair $\{x, y\}$ belongs to $E(\mathrm{disj}_k\, \mathcal{G})$ if they have at most k common neighbors, i.e., $|N(x) \cap N(y)| \leq k$. In this way $\mathrm{disj}\, \mathcal{G}$ defined above is just $\mathrm{disj}_0\, \mathcal{G}$. If we use directly the Szemerédi lemma [9] instead of Theorem 3.2, then we can obtain the following statement, which was the essence of the proof presented in Section 3:

Theorem 7.2. Let k be a fixed integer. Then from any graph $\mathcal{G}$ on n vertices one can remove $o(n^2)$ edges such that the following holds: If x and y had at most k common neighbors in $\mathcal{G}$, then in the obtained new graph $\mathcal{G}_k$ they have no common neighbor anymore, i.e., $E(\mathrm{disj}_k\, \mathcal{G}) \subset E(\mathrm{disj}\, \mathcal{G}_k)$. ∎

The following would be a powerful sharpening of the above theorem. For simplicity we state only the case $k = 1$.

Conjecture 7.3. One can remove $o(n^2)$ edges from any graph $\mathcal{G}$ on n vertices such that the following holds for the obtained graph $\mathcal{G}'$: For every x and $y \in V$ either $N_{\mathcal{G}'}(x) \cap N_{\mathcal{G}'}(y) = \emptyset$, or $|N_{\mathcal{G}'}(x) \cap N_{\mathcal{G}'}(y)| > 1$.

Recently Duke and Rödl [4] have had some remarkable results in this direction. (They investigated bipartite graphs only.)

The following conjecture generalizes our main theorem:

Conjecture 7.4. Let $\mathcal{G}$ be a graph over n vertices and suppose that every two vertex is connected by at least k paths of length at most 2. Suppose further that $\mathcal{G}$ is minimal with respect this property. Then $|E(\mathcal{G})| \le (k - 1)(n - k + 1) + \lfloor (n - k + 1)^2/4 \rfloor$.

Here the extremal graph would be complete 3-colored graph with parts of sizes $\lfloor (n - k + 1)/2 \rfloor$, $\lceil (n - k + 1)/2 \rceil$ and $k - 1$. Caccetta and Häggkvist raised the following conjectures, which also generalize Conjecture 1.1:

Conjecture 7.5 [2]. If $\mathcal{G}$ is a minimal graph of diameter 2, then $\bar{d} \le |V(\mathcal{G})|$, where $\bar{d}$ denotes the average edge-degree in $\mathcal{G}$, i.e.,

$$\bar{d} = \sum_{\{x,y\} \in E(\mathcal{G})} (\deg(x) + \deg(y))/|E(\mathcal{G})| = \sum_{x \in V(\mathcal{G})} (\deg(x))^2/|E(\mathcal{G})|.$$

Conjecture 7.6 [2]. If $\mathcal{G}$ is a minimal graph of diameter k, with $k > 2$, then $|E(\mathcal{G})| \le (1 + o(1))n^2/2(k + 1)^2$.

The conjectured extremal graph consists of two complete bipartite graphs $\mathcal{K}[A_0, A_1]$ and $\mathcal{K}[A_{k-1}, A_k]$ where $|A_i| \sim n/(k + 1)$, and $|A_1| =: |A_{k-1}|$, and $|A_1|$ disjoint path of length $k - 2$ connecting the points of A_1 to A_{k-1}. The method presented in this paper does not seem to be applicable in proving Conjecture 7.5, but it may be useful for attacking the last one. To find further problems (and results) about diameter critical graphs, one can see, e.g., [3] or [1].

Remark 7.7. Maybe it is worth noting that this proof is the first application of Szemerédi's Regularity Lemma, where an exact result is obtained (at least for $n > n_0$).

ACKNOWLEDGEMENTS

This research was supported in part by the Hungarian National Science Foundation under Grant No. 1812. The author is grateful to the referees for their helpful comments.

References

[1] B. Bollobás, *Extremal Graph Theory.* Academic Press, London (1978).
[2] L. Caccetta and R. Häggkvist, On diameter critical graphs. *Discrete Math.* **28** (1979) 223–229.
[3] F. R. K. Chung, Diameters of communication networks. *Proc. Symp. Appl. Math. AMS* **34** (1986) 1–18.
[4] R. Duke and V. Rödl, The Erdös-Ko-Rado theorem for small families, *J. Combinat. Theory, Ser. A*, to appear.
[5] P. Erdös and D. J. Kleitman, On coloring graphs to maximize the proportion of multicolored k-edges. *J. Combinat. Theory* **5** (1968) 164–169.
[6] G. Fan, On diameter 2-critical graphs. *Discrete Math.* **67**(1987) 235–240.
[7] J. Plesník, Critical graphs of given diameter. *Acta Fac. Rerum Natur. Univ. Comenianae Math.* **30** (1975) 71–93.
[8] I. Z. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles. *Combinatorics,* Proc. Colloq. Math. Soc. J. Bolyai 18, Vol. II, North-Holland, Amsterdam—New York (1978), pp. 939–945.
[9] E. Szemerédi, Regular partitions of graphs. *Problèmes Combinatoires et Théorie des Graphes, Colloques Internationaux C.N.R.S.,* No. 260, Paris (1978), pp. 399–401.
[10] J. M. Xu, Proof of a conjecture of Simon and Murty. *J. Math. Res. Exposition* **4** (1984) 85–86 (in Chinese) [Corrigendum. Ibid. **5** (1985) p. 38.]