More about sparse halves in triangle-free graphs

A. A. Razborov

Abstract. One of Erdős’s conjectures states that every triangle-free graph on \( n \) vertices has an induced subgraph on \( n/2 \) vertices with at most \( n^2/50 \) edges. We report several partial results towards this conjecture. In particular, we establish the new bound \( 27n^2/1024 \) on the number of edges in the general case. We completely prove the conjecture for graphs of girth \( \geq 5 \), for graphs with independence number \( \geq 2n/5 \) and for strongly regular graphs. Each of these three classes includes both known (conjectured) extremal configurations, the 5-cycle and the Petersen graph.

Bibliography: 21 titles.

Keywords: extremal graph theory, triangle-free graphs.

§ 1. Introduction

Throughout his long career, Erdős repeatedly [8]–[10] asked several questions united by one common theme: how far from being bipartite can a triangle-free graph be. One of them, the ‘pentagon problem’, was completely solved in [16] and [14]. Another question asks what the maximum possible \( \ell_1 \)-distance can be (which in this case is simply the number of edges deleted) from a triangle-free graph to the class of bipartite graphs. It was studied in [5], [7] and [1].

This paper is devoted to a third question, a ‘half-graph’ conjecture sometimes referred to as ‘one of Erdős’s favourites’ [19]. Given a triangle-free graph \( G \), is it always possible to remove half of its vertices so that the edge density \( |E(G)|/(2|V(G)|^2) \) becomes \( \leq 1/25 \)? In this direction, there has been more recent work done (see [6] and [18]–[20]) although the conjecture still remains wide open.

In this paper we improve on several statements from those papers and offer some new results.

Fix a triangle-free graph \( G \) on \( n \) vertices and let \( \beta(G) \) be the minimum number of edges in its half-graphs, normalized\(^1\) by \( n^2 \). The half-graph conjecture by Erdős says that \( \beta(G) \leq 1/50 \), for any triangle-free \( G \). The bound \( \beta(G) \leq 1/16 \) is obvious (attained by the random half), [6] proved that \( \beta(G) \leq 1/30 \) and [18] improved this to \( \beta(G) \leq 1/36 \).

\(^1\)It would have been much more natural to normalize by \( n^2/2 \) instead but we prefer our notation to be consistent with the literature.

AMS 2020 Mathematics Subject Classification. Primary 05C35.

© 2022 Russian Academy of Sciences (DoM) and London Mathematical Society
**Theorem 1.1.** For any triangle-free graph $G$,

$$\beta(G) \leq \frac{27}{1024}.$$ 

The number $27/1024$ here is not arbitrary, it reflects what can be achieved with a certain class of methods, and the Clebsch graph is an extremal example for the resulting extremal problem. We will comment more on it below.

The (conjectural) extremal examples in the half-graph conjecture are the pentagon $C_5$ and the Petersen graph. Moreover, the former does not contain induced matching of size 2.

**Theorem 1.2.** The half-graph conjecture is true for any (triangle-free) graph without induced matchings of size 2.

Before going any further, let us briefly discuss the proofs of these two theorems, as they bring about potentially interesting concepts and questions.

**Digression on quadrilateral counting.** Let $\rho = \rho(G)$ and $C_4 = C_4(G)$ be the edge density of $G$ and the density of quadrilaterals (copies of $C_4$) in it. They are computed in the sense of flag algebras/graph limits: $G$ is first replaced with its infinite blow-up (so that in particular copies of the path $P_3$ in $G$ and even individual edges contribute to $C_4(G)$). These two quantities are of fundamental importance in the theory of quasi-random graphs: $C_4 \geq 3\rho^4$, and an increasing sequence of graphs with the same value of $\rho(G)$ is quasi-random if and only if this inequality is asymptotically tight (see [4]).

For triangle-free graphs this bound can easily be improved to

$$C_4 \geq \frac{3\rho^4}{1 - \rho}. \quad (1)$$

This is a quantitative refinement of the statement that triangle-free graphs are not quasi-random, and thus it is natural to ask: what is the smallest value of $C_4(G)$ as a function of $\rho(G)$? In a sense, it is a dual to Erdős’s questions. The latter ask, in one form or another, how far from being bi-partite a triangle-free graph can be. The ‘quadrilateral question’, on the contrary, asks how far from quasi-random a triangle-free graph must be.

We expect this question to be extremely difficult in general. But it is very closely related to Erdős’s conjectures as has already been demonstrated in [5], §2. In our context, an easy analysis of (part of) Krivelevich’s proof, followed by a straightforward averaging gives the following.

**Proposition 1.3.**

$$\beta(G) \leq \frac{1}{8} \rho(G) - \frac{C_4(G)}{12 \rho(G)}.$$ 

Both bounds on $\beta(G)$ may then be derived from the following.
Theorem 1.4. 1. For any triangle-free graph \( G \),

\[ C_4(G) \geq \frac{3}{2} \rho(G)^2 - \frac{81}{256} \rho(G). \]

2. For any triangle-free graph \( G \) without induced matchings of size 2,

\[ C_4(G) \geq \frac{3}{2} \rho(G)^2 - \frac{6}{25} \rho(G). \]

This theorem is proved via a ‘medium size’ flag-algebraic calculation. The second bound is tight for \( \rho = 2/5 \), as it must be since the pentagon \( C_5 \) is the (conjectural) extremal example for the half-graph conjecture. The first inequality beats the trivial bound (1) for \( 0.257 \leq \rho \leq 0.366 \). It is tight for \( \rho = 5/16 \), the extremal example being the Clebsch graph, and this seems to be the only nontrivial value of \( \rho(G) \) for which we know the exact solution to the quadrilateral problem.

We further remark that our bound \( \beta(G) \leq 27/1024 \) is tight for a reasonably natural restriction of Erdős’s conjecture. More specifically, many previous results, including Proposition 1.3, are based on the following simple construction. Pick an edge \( e \in E(G) \). It naturally defines a splitting of \( V(G) \) into three parts. We somehow assign total weights to each of these three parts and then distribute the weights within each part uniformly. Then the value 27/1024 in the half-graph conjecture is the best one can achieve using this method, with the Clebsch graph being (again) an extremal example.

Let us now return to reviewing the remaining results. The half-graph conjecture is obvious if \( \rho(G) \leq 4/25 \) (once more, the random half will do the job). Keevash and Sudakov [19] relaxed this to \( \rho(G) \leq 1/6 \).

Theorem 1.5. The half-graph conjecture is true for any triangle-free graph with

\[ \rho(G) \leq \frac{33 - \sqrt{161}}{116} \approx 0.1751. \]

Based on this theorem, we prove the following.

Theorem 1.6. The half-graph conjecture is true for any triangle-free strongly regular graph.

A significant amount of activity has taken place around the critical value \( \rho = 2/5 \). Theorem 3 in [18] proved the conjecture for regular triangle-free graphs with \( \rho(G) \geq 2/5 \) and [19] removed the restriction of regularity. Norin and Yepremyan [20] improved this result by relaxing the assumption \( \rho(G) \geq 2/5 \) to \( \rho(G) \geq 2/5 - \gamma \), where \( \gamma > 0 \) is a (calculable) constant. When \( \rho(G) \) is replaced by the (normalized) minimum degree \( \delta(G) \), the bound on \( \gamma \) significantly improves and the half-graph conjecture is true whenever \( \delta(G) \geq 5/14 \) [20].

Theorem 1.7. Let \( \alpha(G) \) be the normalized (by \( n \)) independence number of \( G \), and assume that \( \alpha(G) \geq 3/8 \). Then

\[ \beta(G) \leq \frac{1}{2} \alpha(G) \left( \frac{1}{2} - \alpha(G) \right). \]
Corollary 1.8. The half-graph conjecture holds for any triangle-free graph with (normalized) maximum degree $\geq 2/5$.

Note that unlike the previous results, we do not require all vertices to have large degree, not even on average, but just one. Also, Theorem 1.7 covers the Petersen graph as well, since it has (unnormalized) independence number 4. On the negative side, we have not been able to extend it to an open neighbourhood of $2/5$ as the previous work did.

Finally, both the conjectured extremal examples (the five-vertex cycle and the Petersen graph) have girth 5.

Theorem 1.9. The half-graph conjecture holds for all triangle-free graphs of girth $\geq 5$.

The rest of the paper is organized as follows. In §2 we give all necessary definitions. In §3 we re-state our results, mostly as a matter of convenience. Section 4 is devoted to proofs, and we conclude in §5 with a few remarks and open questions.

§ 2. Preliminaries

Unless specified otherwise, all graphs $G$ in this paper are finite, simple and triangle-free. $N_G(v)$ is the neighbourhood of $v$, and $n$ will always stand for the number of vertices. For disjoint sets of vertices $X$ and $Y$, $E(X,Y)$ is the set of cross-edges between $X$ and $Y$; likewise, $E(X)$ is the set of edges induced by $X$.

A half in a graph $G$ with $n$ vertices is a function $\mu: V(G) \to [0,1]$ such that
\[ \sum_{v \in V(G)} \mu(v) = n/2. \]
We let
\[ \beta(G, \mu) \equiv \frac{1}{n^2} \sum_{(u,v) \in E(G)} \mu(u)\mu(v) \]
and
\[ \beta(G) \equiv \min(\beta(G, \mu) \mid \mu \text{ is a half in } G). \]

It is easy to see from a simple convexity argument that when $n$ is even, the minimum is attained at a $0-1$ half, in which case we also use the notation $\beta(G, A)$, $A \in \binom{V(G)}{n/2}$. When $n$ is odd, it is attained at an almost $0-1$ half $\mu$, that is $\mu(v_0) = 1/2$ for a single vertex $v_0$ and $\mu(v) \in \{0,1\}$ for all others. But the analytic form above is extremely handy in concrete constructions, as we shall see.

Conjecture 2.1 (half-graph conjecture by Erdős). $\beta(G) \leq 1/50$ for any triangle-free graph $G$.

Let $\rho(G)$ and $C_4(G)$ be the edge density and the density of quadrilaterals defined consistently with flag algebras. That is,
\[ \rho(G) \equiv \frac{2|E(G)|}{n^2}, \]
and in order to compute $C_4(G)$ we sample $v_i \in V(G)$, $i \in [4]$, uniformly and independently of each other (repetitions are allowed), form the graph on $[4]$ with the set of edges $\{ (i,j) \in \binom{[4]}{2} \mid (v_i,v_j) \in E(G) \}$ and let $C_4(G)$ be the probability that it is isomorphic to $C_4$. 

For $v \in V(G)$ we let
\[ e(v) \overset{\text{def}}{=} \frac{|N_G(v)|}{n} \]
be the relative degree of $v$ and
\[ \Delta(G) \overset{\text{def}}{=} \max(e(v) \mid v \in V(G)) \]
be the maximum degree, also relative. Likewise,
\[ \alpha(G) \overset{\text{def}}{=} \max \left( \frac{|A|}{n} \mid A \text{ is an independent set} \right) \]
is the relative independence number.

We can assume without lost of generality that $\alpha(G) \leq 1/2$ since otherwise Conjecture 2.1 is obvious. Thus,
\[ \rho(G) \leq \Delta(G) \leq \alpha(G) \leq \frac{1}{2}. \tag{2} \]

For sets of vertices $A, B, C, D, E, \ldots$ we denote their densities by $p_a, p_b, p_c, \ldots$:
\[ p_x \overset{\text{def}}{=} \frac{|X|}{n}. \]
Let $\rho_{xy}$ ($x \neq y \in \{a, b, c, d, \ldots\}$, $X \cap Y = \emptyset$) be the normalized density of cross-edges:
\[ \rho_{xy} \overset{\text{def}}{=} \frac{2|E(X, Y)|}{n^2}; \]
and, likewise,
\[ \rho_x \overset{\text{def}}{=} \frac{|E(X)|}{n^2}. \]
Finally, for $v \notin X$, let
\[ e_X(v) \overset{\text{def}}{=} \frac{|N_G(v) \cap X|}{n}. \]

§ 3. Results

In this section we collect in one place our main results stated in §1.

**Theorem 3.1.** a) For any triangle-free graph $G$,
\[ C_4(G) \geq \frac{3}{2} \rho(G)^2 - \frac{81}{256} \rho(G) \]
(the bound is tight for the Clebsch graph).

b) For any triangle-free graph $G$ without induced matchings of size 2,
\[ C_4(G) \geq \frac{3}{2} \rho(G)^2 - \frac{6}{25} \rho(G) \]
(the bound is tight for $C_5$).
Theorem 3.2. For any triangle-free graph $G$, 
\[ \beta(G) \leq \frac{27}{1024}. \]

Theorem 3.3. Conjecture 2.1 is true for any triangle-free graph without induced matchings of size 2.

Theorem 3.4. Conjecture 2.1 is true for any triangle-free graph with 
\[ \rho(G) \leq \rho_0 \overset{\rm def}{=} \frac{33 - \sqrt{161}}{116}. \]

Recall that a regular triangle-free graph $G$ is strongly regular if $|N_G(v) \cap N_G(w)|$ takes the same value $c$ for all pairs $(v, w)$ of nonadjacent distinct vertices.

Theorem 3.5. Conjecture 2.1 is true for any triangle-free strongly regular graph.

Theorem 3.6. For any triangle-free graph $G$ with $\alpha(G) \geq 3/8$ we have 
\[ \beta(G) \leq \frac{1}{2} \alpha(G) \left( \frac{1}{2} - \alpha(G) \right). \]

Corollary 3.7. Conjecture 2.1 is true for any triangle-free graph with $\alpha(G) \geq 2/5$.

Theorem 3.8. Conjecture 2.1 is true for any triangle-free graph of girth $\geq 5$.

§ 4. Proofs

In this section we prove all our results. Some of the proofs, particularly those in §§4.1 and 4.3, rely heavily on symbolic Maple computations. The corresponding worksheet, along with some supporting material, can be found at https://people.cs.uchicago.edu/~razborov/files/halves.zip.

4.1. Flag-algebraic calculations. In this section we prove Theorem 3.1. As we remarked in §2, our notation for finite graphs is consistent with flag algebras, hence it is sufficient to prove the inequalities 
\[ \frac{3}{2} \rho^2 - \frac{81}{256} \rho \leq C_4 \] (3) 
and 
\[ \frac{3}{2} \rho^2 - \frac{6}{25} \rho \leq C_4 + 2M_4 \] (4) 
(M$_4$ is the matching with two edges) in the theory $T_{\text{TF}}$ of triangle-free graphs and then apply them to the infinite (balanced) blow-up of $G$.

We do this by a straightforward Cauchy-Schwarz computation in flag algebras. Since quite a number of those have already appeared in the literature, with varying degree of informal explanation, we do ours matter-of-factly, strictly adhering to the notation of [21].

Let us start with (3); for that we need to consider triangle-free graphs on eight vertices. We have 
\[ |M_8| = 410 \quad \text{and} \quad |F^\sigma_6| = d_i, \quad \text{where} \quad d_1 = 110, \ d_2 = 81, \ d_3 = 67, \ d_4 = 46, \]
and the types $\sigma_i$ are shown in Figure 1 (with the exception of $\sigma_4$, these are the same types as employed in [15]). We enumerate flags in $F_6^{\sigma_i}$ in a rather arbitrary order as $F_6^{\sigma_i} = \{ F_1^{\sigma_i}, \ldots, F_{d_i}^{\sigma_i} \}$ and exhibit PSD matrices $Q_i$ of size $d_i \times d_i$ with rational coefficients such that

\[ \sum_{i=1}^{4} \sum_{j_1, j_2 = 1}^{d_i} [Q_i(j_1, j_2)F_{j_1}^{\sigma_i}F_{j_2}^{\sigma_i}]_{\sigma_i} \ll_8 C_4 - \frac{3}{2} \rho^2 + \frac{81}{256} \rho, \]  

(5)

where $\ll_8$ means coefficient-wise comparison after expressing both sides of this inequality as linear combinations of the elements of $M_8$.

![Figure 1. Types.](https://example.com/figure1.png)

The only further remark we want to make here is that the matrices $Q_i$ are singular and their co-ranks $d_i - \text{rk}(Q_i)$ are equal to 2, 2, 5 and 4, respectively.

This reflects the fact (and makes an excellent sanity check for our calculations) that the Clebsch graph $G_{\text{Clebsch}}$ is an extremal configuration for inequality (5). Hence every strict homomorphism $\sigma_i \rightarrow G_{\text{Clebsch}}$ gives rise to an element in the kernel of $Q_i$. The actual computation is deferred to https://people.cs.uchicago.edu/~razborov/files/halves.zip.

The inequality (4) is proved similarly, but this time we need only graphs on six vertices; on the other hand, instead of $\sigma_3$ we need the type $E$. We have

\[ |M_6| = 38 \quad \text{and} \quad |F_6^{\sigma_i}| = d_i, \quad \text{where} \quad d_1 = 12, \ d_2 = 10, \ d_4 = 7, \]

and also $|M_6^E| = 10$. The computation has the form

\[ \sum_{j_1, j_2 = 1}^{10} [R_E(j_1, j_2)F_{j_1}^E F_{j_2}^E]_E + \sum_{i \in \{1,2,4\}} \sum_{j_1, j_2 = 1}^{d_i} [R_i(j_1, j_2)F_{j_1}^{\sigma_i} F_{j_2}^{\sigma_i}]_{\sigma_i} \]

\[ \ll_6 C_4 + 2M_4 - \frac{3}{2} \rho^2 + \frac{6}{25} \rho. \]

The coefficient 2 in front of $M_4$ is rather arbitrary; we did not attempt to optimize on it. As this inequality is tight on $C_5$, the matrices $R_E$, $R_1$, $R_2$ and $R_4$ must also be singular, and they have indeed co-ranks 1, 1, 1 and 3, respectively.
4.2. Absolute upper bounds on $\beta(G)$. In this section we establish Theorems 3.2 and 3.3. As already mentioned, they immediately follow from Theorem 3.1 and Proposition 1.3, so it only remains to prove the latter. This is simply part of Krivelevich’s argument [18], slightly re-phrased, but we include it here for the sake of completeness.

Let us start by considering an individual edge $(v_1, v_2) \in E(G)$. Denote

$$A_i \overset{\text{def}}{=} N_{G}(v_i) \quad \text{and} \quad e_i \overset{\text{def}}{=} e(v_i) \quad (= p_a);$$

recall that $e_i \leq 1/2$ by (2). Let

$$p_i \overset{\text{def}}{=} \frac{1}{2} - e_i \frac{1}{1 - e_1 - e_2},$$

so that $p_1 + p_2 = 1$, and let $B = V(G) \setminus (A_1 \cup A_2)$. For $i = 1, 2$ define the half $\mu_i$ by

$$\mu_i(v) \overset{\text{def}}{=} \begin{cases} 1 & \text{if } v \in A_i, \\ p_i & \text{if } v \in B, \\ 0 & \text{if } v \in A_{3-i}. \end{cases}$$

Then

$$2\beta(G) \leq 2\beta(G, \mu_1) = p_1 \rho_{a_1, b} + p_1^2 \rho_b$$

and

$$2\beta(G) \leq 2\beta(G, \mu_2) = p_2 \rho_{a_2, b} + p_2^2 \rho_b. \quad (7)$$

Multiplying the $i$th inequality here by $p_{3-i}$ and adding them together, we get

$$2\beta(G) \leq p_1 p_2 (\rho_{a_1, b} + \rho_{a_2, b} + \rho_b) = p_1 p_2 (\rho - \rho_{a_1, a_2}) \leq \frac{1}{4} (\rho - C_4^E(v_1, v_2)),$$

where we have denoted $\rho_{a_1, a_2}$ by $C_4^E(v_1, v_2)$ to stress that this is the contribution of $(v_1, v_2)$ to $C_4(G)$. Finally, averaging this over all edges we obtain

$$2\rho \beta(G) \leq \frac{1}{4} \lVert \rho - C_4^E \rVert_E = \frac{1}{4} \left( \rho^2 - \frac{2}{3} C_4 \right),$$

which is precisely Proposition 1.3.

Theorems 3.2 and 3.3 are proved.

4.3. Sparse graphs. In this section we prove Theorem 3.4. As in the previous work [19], the analysis splits into two cases: $\Delta(G) \geq 1/4$ and $\Delta(G) \leq 1/4$.

The first case is taken care of by the following variant of Proposition 1.3.

Lemma 4.1.

$$\beta(G) \leq \frac{\rho(G)(1 - 2\Delta(G))}{8(1 - \Delta(G))^2}. \quad (8)$$

**Proof.** Pick $v \in V(G)$ with $e(v) = \Delta \overset{\text{def}}{=} \Delta(G)$, and let $A \overset{\text{def}}{=} N_{G}(v)$ (so that $p_a = \Delta$) and $B = N_{G}(v) \setminus A$. 

Construct the following halves $\mu_0$ and $\mu_1$:

$$
\mu_0(w) \overset{\text{def}}{=} \begin{cases} 
0 & \text{if } w \in A, \\
\frac{1}{2(1 - \Delta)} & \text{if } w \in B
\end{cases}
\quad \text{and} \quad
\mu_1(w) \overset{\text{def}}{=} \begin{cases} 
1 & \text{if } w \in A, \\
\frac{1}{2 - \Delta} & \text{if } w \in B.
\end{cases}
$$

Then

$$
2\beta(G) \leq 2\beta(G, \mu_0) = \frac{\rho_b}{4(1 - \Delta)^2}
$$

and

$$
2\beta(G) \leq 2\beta(G, \mu_1) = \frac{1}{2 - \Delta} \rho_{ab} + \left(\frac{1}{2 - \Delta}\right)^2 \rho_b.
$$

Multiplying (9) by $1 - 2\Delta$ and adding it to (8), we obtain

$$
4(1 - \Delta)\beta(G) \leq \frac{1 - 2\Delta}{2(1 - \Delta)}(\rho_{ab} + \rho_b) = \frac{1 - 2\Delta}{2(1 - \Delta)}\rho(G).
$$

The lemma is proved.

Now, the function $(1 - 2\Delta)/(8(1 - \Delta^2))$ is decreasing for $\Delta \in [1/4, 1/2]$, hence $\Delta(G) \geq 1/4$ implies that $\beta(G) \leq \rho(G)/9$ and then Theorem 3.4 follows since $\rho_0 \leq 9/50$.

The case $\Delta(G) \leq 1/4$ is more difficult. As in the proof of Proposition 1.3, let us first consider an individual edge $(v_1, v_2) \in E(G)$ (but this time we will not randomize over this choice but will pick it up in a way to be specified later). We will re-use the notation $e_1$, $A_1$, $B$ and $p_i$ from that proof so that we still have the bounds (6) and (7). But now the condition $\Delta(G) \leq 1/4$ allows us to form one more half

$$
\mu_0(G) \overset{\text{def}}{=} \begin{cases}
1 & \text{if } v \in A_1 \cup A_2, \\
q & \text{if } v \in B,
\end{cases}
$$

where

$$
p_0 \overset{\text{def}}{=} \frac{1/2 - e_1 - e_2}{1 - e_1 - e_2}.
$$

This leads to the extra bound

$$
2\beta(G) \leq \rho_{a_1a_2} + p_0(\rho_{a_1b} + \rho_{a_2b}) + p_0^2 \rho_b.
$$

We are now looking for nonnegative coefficients $\alpha_0$, $\alpha_1$ and $\alpha_2$ such that multiplying (10), (6) and (7) by them, respectively, and adding up the results, we equalize the coefficients in front of $\rho_{a_1b}$ and $\rho_{a_1a_2}$, as well as $\rho_{a_2b}$ and $\rho_b$. For that purpose we set

$$
\alpha_0 \overset{\text{def}}{=} (1 - 2e_1)^2(1 - 2e_2),
\alpha_1 \overset{\text{def}}{=} (1 - 2e_1)(1 - 2e_2),
\alpha_2 \overset{\text{def}}{=} 2(1 - 2e_1)e_2.
$$
Then (see the Maple worksheet)

\[ 4(1 - 2e_1)(1 - e_1 - e_2 + 2e_1e_2)\beta(G) \leq \alpha_0(\rho_{a_1a_2} + \rho_{a_1b}) + \gamma(\rho_{a_2b} + \rho_b) \]

\[ = (\alpha_0 - \gamma)(\rho_{a_1a_2} + \rho_{a_1b}) + \gamma(\rho_{a_1a_2} + \rho_{a_1b} + \rho_{a_2b} + \rho_b) \]

where

\[ \gamma \overset{\text{def}}{=} \frac{(1 - 2e_1)(1 - 2e_2)(1 - 4e_1 + 4e_1^2 + 4e_1e_2)}{2(1 - e_1 - e_2)}. \]

Note for the record that

\[ \alpha_0 - \gamma = \frac{(1 - 2e_1)(1 - 2e_2)(1 - 2e_1 - 2e_2)}{2(1 - e_1 - e_2)} \geq 0, \]

since \( e_1, e_2 \leq \Delta(G) \leq 1/4 \). Hence we need an upper bound on \( \rho_{a_1a_2} + \rho_{a_1b} \).

For that purpose we now specify \( v_1 \) and \( v_2 \). The vertex \( v_1 \) is chosen as the vertex of the maximum degree so that \( e_1 = \Delta \). We choose \( v_2 \) to have the maximum degree among all vertices in \( N_G(v_1) \). The latter choice gives us the estimate \( \rho_{a_1a_2} + \rho_{a_1b} \leq 2\Delta e_2 \) since \( \rho_{a_1a_2} + \rho_{a_1b} \) is simply the overall density of edges incident to \( A_1 \). Putting all this together, we arrive at the estimate

\[ \beta(G) \leq f(\rho, \Delta, e_2) \overset{\text{def}}{=} \frac{(1 - 2e_2)(4\Delta^2 - 4\Delta^2e_2 + 4\Delta\rho e_2 - 4\Delta e_2^2 - 4\Delta\rho + 2\Delta e_2 + \rho)}{8(1 + 2\Delta e_2 - \Delta - e_2)(1 - \Delta - e_2)}. \]

Let us also recall that we have the constraints

\[ 0 \leq \rho \quad \text{and} \quad e_2 \leq \Delta \leq \frac{1}{4}. \]

This optimization problem is a bit nasty to be fully analyzed, that is, to give an analytical estimate of \( \beta(G) \) in terms of \( \rho(G) \). Instead, we compute

\[ \frac{1}{50} - f(\rho, \Delta, e_2) = \frac{Q(\rho, \Delta, e_2)}{200(1 - \Delta - e_2)(1 - \Delta - e_2 + 2\Delta e_2)}, \]

where \( Q \) is a polynomial. Our goal is to show that \( \rho \leq \rho_0 \) implies that \( Q(\rho, \Delta, e_2) \geq 0 \).

We note that \( Q(\rho_0, \rho_0, \rho_0) = 0 \) and that the individual degrees of \( Q \) in \( \rho, \Delta \) and \( e_2 \) are 1, 2 and 3, respectively.

We first compute

\[ \frac{\partial Q}{\partial \rho} = -25(1 - 2e_2)(4\Delta^2 + 4\Delta e_2 - 4\Delta + 1) \leq -25(1 - 2e_2)(1 - 2\Delta)^2 \leq 0. \]

Hence it is sufficient to prove that

\[ Q_1(\Delta, e_2) \overset{\text{def}}{=} Q(\min(\rho_0, \Delta), \Delta, e_2) \geq 0. \]

\( Q_1 \) is no longer smooth in \( \Delta \) but it is still a cubic polynomial in \( e_2 \). We consider two cases: \( e_2 \leq \rho_0 \) and \( e_2 \geq \rho_0 \).
If \( e_2 \leq \rho_0 \), then we consider the Taylor coefficients at \( e_2 = \rho_0 \):

\[
\frac{1}{r!} \left. \frac{\partial^r Q_1(\Delta, e_2)}{(\partial e_2)^r} \right|_{e_2=\rho_0}, \quad r = 0, 1, 2, 3,
\]

and it turns out (see the Maple worksheet) that they are nonnegative for even \( r \) and negative for odd \( r \). The required inequality \( Q_1(\Delta, e_2) \geq 0 \) follows.

In the second case \( e_2 \geq \rho_0 \) we also have \( \Delta \geq \rho_0 \) and hence \( Q_1(\Delta, e_2) = Q(\rho_0, \Delta, e_2) \) is a quadratic polynomial in \( \Delta \in [e_2, 1/4] \). It should be noted that it can either be convex or concave. But in either case the required inequality \( Q_1(\Delta, e_2) \geq 0 \) follows from

\[ Q_1(e_2, e_2) \geq 0, \quad Q_1\left(\frac{1}{4}, e_2\right) \geq 0 \quad \text{and} \quad \frac{\partial Q_1}{\partial \Delta} \bigg|_{\Delta=e_2} \geq 0. \]

Theorem 3.4 is proved.

4.4. Strongly regular case. In this section we prove Theorem 3.5. For some brief background on triangle-free strongly regular (TFSR in what follows) graphs we follow [2].

Except for the complete bipartite graphs \( K_{n,n} \) (for which Erdős’s conjecture is vacuously true), there are seven known examples of TFSR graphs; the cycle \( C_5 \), the Petersen graph and the Clebsch graph being the smallest. The obvious parameters of a TFSR graph \( G \) are \( n, k \) (the degree of a vertex) and \( c \) (the number of common neighbours of a pair of non-adjacent vertices). They are actually related by

\[
n = 1 + \frac{k}{c} (k - 1 + c).
\]

From now on we assume that \( G \) is different from \( K_{n,n} \) and that it is different from \( C_5 \). Then the quantity \( s \defeq \sqrt{c^2 + 4(k - c)} \) is an integer, and the only positive eigenvalue of the adjacency matrix different from \( k \) is given by \( q \defeq (s - c)/2 \). It is also an integer such that

\[
1 \leq c \leq q(q + 1).
\] (11)

Furthermore, we have

\[
k = (q + 1)c + q^2,
\]

hence \( k \) and \( n \) are rational functions in \( c \) and \( q \). In particular, we compute

\[
\rho(G) = \frac{k}{n} = \frac{c(qc + q^2 + c)}{(qc + q^2 + 2c + q)(qc + q^2 + c - q)} \defeq Q(q, c).
\]

Let us now analyze this expression. Firstly,

\[
\frac{\partial Q}{\partial c} = \frac{q(q + 1)(c^2q^2 + 2cq^3 + q^4 + c^2q - q^3 - c^2 - 2qc)}{(qc + q^2 + 2c + q)^2(qc + q^2 + c - q)^2} \geq 0;
\]

hence \( Q \) is increasing in \( c \).
Next, let

\[ Q_1(q) \overset{\text{def}}{=} Q(q, q(q + 1)) = \frac{q^2 + 3q + 1}{q(q + 3)^2} \]

(this corresponds to so-called Krein graphs). Then

\[ Q_1(q)' = -\frac{q^3 + 3q^3 + 3q + 3}{q^2(q + 3)^3} < 0, \]

hence \( Q_1(q) \) is decreasing. As \( Q_1(4) = 29/196 < \rho_0 \), the proof of Theorem 3.5 boils down to the three cases \( q = 1, 2, 3 \): all the others are taken care of by Theorem 3.4.

When \( q = 1 \), we have either the Petersen graph \((c = 1)\) or the Clebsch graph \((c = 2)\). Conjecture 2.1 for the Clebsch graph is verified by the half \((N_G(u) \cup N_G(v)) \setminus \{u, v\}\), where \((u, v)\) is an arbitrary edge.

When \( q = 2 \), we have \( Q(2, 1) = 7/50 < \rho_0 \) (this is the Hoffman-Singleton graph) hence it is sufficient to consider the cases \( 2 \leq c \leq 6 \). The well-known ‘arithmetic conditions’ rule out \( c \in \{3, 5\} \) (see [2], Table 1), and the three other cases correspond precisely to the remaining known TFSR graphs: Gewirtz, \( M_{22} \) and Higman-Sims (they are unique for their values of \( c \) and \( q \); see [12], [3] and [11]).

For the Gewirtz graph, we pick up four vertices \( u_1, u_2, u_3 \) and \( u_4 \) spanning an induced matching with two edges and consider the half (see the Maple worksheet) induced by the set

\[ \bigcup_{i=1}^{4} N_G(u_i) \setminus \{u_1, u_2, u_3, u_4\}. \]

It spans 51 edges, which proves \( \beta(G) \leq 0.017 \).

When \( G \) is the \( M_{22} \) graph, we similarly let

\[ A \overset{\text{def}}{=} \bigcup_{i=1}^{3} N_G(u_i) \setminus \{u_1, u_2, u_3\}, \]

where \((u_1, u_2) \in E(G)\) and \( u_3 \notin N_G(u_1) \cup N_G(u_2) \). Then \(|A| = 38\) and \(|E(A)| = 109\). Moreover, there exists a vertex \( v \notin A \) such that \(|N_G(v) \cap A| = 9\). Adding ‘half’ of the vertex \( v \) to \( A \) we obtain a half proving \( \beta(G) \leq 0.0192 \).

For the Higman-Sims graph we present an ad hoc half achieving \( \beta(G) \leq 1/50 - 10^{-4} \). It was found by a simple optimization program, remarkably suggesting that this bound is actually tight. If it is true (we did not attempt to verify the claim with a rigorous argument), then the Higman-Sims graph comes very close to the bound in Erdős’s conjecture.

Finally, when \( q = 3 \), we have \( Q(3, 11) = 583/3350 < \rho_0 \). Hence the only case to consider is \( q = 3 \) and \( c = 12 \), that is, a hypothetical 57-regular Krein graph on 324 vertices. A simple solution is to note that such a graph is known not to exist [13], [17]. Let us, however, sketch another argument due to Grzesik and Volec (unpublished), which is in our opinion more instructive and may be of independent interest.

As we already noticed, in the bound (10) the quantity \( \rho_b \) can be eliminated via the identity \( \rho = \rho_{a_1a_2} + \rho_{a_1b} + \rho_{a_2b} + \rho_b \). If \( G \) is also known to be regular, then
More about sparse halves in triangle-free graphs

$p_0 = (1/2 - 2\rho)/(1 - 2\rho)$ and $\rho_{a,b}$ can be also eliminated using $\rho_{a,b} + \rho_{a_1a_2} = 2\rho^2$. Plugging all this into (10) and averaging over all choices of the edge $(v_1, v_2)$, as in §4.2, we arrive at the bound

$$\beta(G) \leq \frac{(2/3)C_4 + \rho^2(1 - 4\rho)}{8\rho(1 - 2\rho)^2},$$

(12)

which holds for any regular (triangle-free) graph $G$.

Now, if $G$ is also strongly regular, then $C_4(G)$ can be easily calculated as

$$C_4(G) = \frac{3}{n^3}(k^2 + c^2(n - k - 1))$$

(recall from §2 that $C_4(G)$ also counts degenerate cycles). Substituting this into (12) we obtain

$$\beta(G) \leq \frac{c(cq + q^2 - c)}{8q(q + 1)(c + q)(c + q - 1)}.$$

In particular, when $q = 3$ and $c = 12$, we have $\beta(G) \leq 11/560$.

Theorem 3.5 is proved.

### 4.5. Graphs with large independence number

In this section we prove Theorem 3.6; as we noted in §1, for $\alpha \geq 2/5$ it generalizes several previously known results.

It will be convenient to assume that $n$ is even: this can always be achieved by replacing each vertex with two identical twins. Let $A \subseteq V(G)$ be an independent set with $p_a = \alpha \geq 3/8$. We build a larger set $B \supseteq A$ by recursively adding to it vertices that bring with them only a few edges. More exactly, apply the following simple algorithm:

$$B := A$$

while $|B| < \frac{n}{2}$ and $\exists v \notin B \left( e_B(v) \leq \frac{1}{2} - \alpha \right)$

do $B := B \cup \{v\}$.

If this algorithm terminates as a result of $B$ reaching size $n/2$, then $\beta(G, B) \leq (1/2 - \alpha)^2$ which is $\leq (\alpha/2)(1/2 - \alpha)$ since $\alpha \geq 3/8 > 1/3$, and we are done. Hence we can assume without lost of generality that the algorithm stops when the required vertex $v$ no longer exists. Thus, we now have a set $B$ such that:

$$\begin{cases}
p_b \in \left[\alpha; \frac{1}{2}\right], \\
\rho_b \leq 2\left(\frac{1}{2} - \alpha\right)(p_b - \alpha), \\
\forall v \notin B \left( e_B(v) > \frac{1}{2} - \alpha \right).
\end{cases}$$

(13)

Let

$$C \overset{\text{def}}{=} \left\{ v \notin B \mid e_B(v) > \frac{p_b}{2} \right\}.$$
Then $C$ is independent (as any two vertices of $C$ have a common neighbour in $B$). Hence $p_e \leq \alpha$ from the definition of $\alpha(G)$. This allows us to choose a set of vertices $D$ disjoint from both $B$ and $C$ and such that $p_d = 1 - \alpha - p_b$. We now consider two cases, depending on whether there exists a vertex in $D$ that has many neighbours in $D$ or not.

Case 1: there exists $v \in B$ such that $e_D(v) \geq 1/2 - p_b$.

Take arbitrarily $E \subseteq N_G(v) \cap D$ with $p_e = 1/2 - p_b$; note that $E \subseteq N_G(v)$ is independent. Also, for every $v \in E$ we have $e_B(v) \leq p_b/2$ since $E \subseteq D$. Then we have (note the absence of the coefficient 2 in the last term)

$$2\beta(G, B \cup E) \leq 2\left(\frac{1}{2} - \alpha\right)(p_b - \alpha) + p_b\left(\frac{1}{2} - p_b\right).$$

The right-hand side is a concave quadratic function in $p_b$, with maximum at $p_b = 3/4 - \alpha$, which is $\leq \alpha$ since we have assumed that $\alpha \geq 3/8$. Hence, since $p_b \geq \alpha$, we can plug in $p_b := \alpha$ and this completes the analysis of Case 1.

Case 2: for any $v \in B$ we have $e_D(v) \leq 1/2 - p_b$.

This case is slightly more elaborate. Let us first fix an individual $v_0 \in B$ (we will later average over this choice). Let $E \overset{\text{def}}{=} N_G(v_0) \cap D$ (so that $p_e \leq 1/2 - p_b$) and $F \overset{\text{def}}{=} D \setminus N_G(v_0)$; thus, $D = E \cup F$ with $E$ independent. Consider the half

$$\mu(v) \overset{\text{def}}{=} \begin{cases} 1 & \text{if } v \in B \cup E, \\ p & \text{if } v \in F, \\ 0 & \text{in all other cases,} \end{cases}$$

where $$p = \frac{1/2 - p_b - p_e}{p_f}.$$ Then

$$2\beta(G, \mu) = \rho_b + \rho_{be} + p\rho_{bf} + p\rho_{ef} + p^2 \rho_f.$$

The bound on $\rho_b$ is given by (13), and we have $\rho_{bf} \leq p_b p_f$; the coefficient 2 is absent for the same reasons as above. For $\rho_{ef}$ we use the trivial bound $\rho_{ef} \leq 2p_e p_f$ and, finally, $\rho_f \leq p_f^2 / 2$ simply because $G$ is triangle-free. Plugging all this into the above bound (we leave $\rho_{be}$ alone for the time being), we obtain

$$2\beta(G, \mu) \leq 2\left(\frac{1}{2} - \alpha\right)(p_b - \alpha) + \rho_{be} + p_b\left(\frac{1}{2} - p_b - p_e\right) + 2p_e\left(\frac{1}{2} - p_b - p_e\right) + \frac{1}{2}(\frac{1}{2} - p_b - p_e)^2$$

$$= 2\left(\frac{1}{2} - \alpha\right)(p_b - \alpha) + \frac{1}{2}\left(\frac{1}{2} - p_b - p_e\right)\left(\frac{1}{2} + p_b + 3p_e\right) + \rho_{be}. \quad (14)$$

In this bound, $p_e$ and $\rho_{be}$ are the only quantities that depend on the choice of $v_0 \in B$, and we now randomize over all such choices.

The bound (14) is concave in $p_e$ hence we may simply replace $p_e$ with its expected value $\rho_{bd}/2p_b$. 


As for \( \rho_{be} \), pick \( w \in_R D \) uniformly at random; then using a standard double counting we see that

\[
\mathbb{E}[\rho_{be}] = \frac{2pd}{pb} \mathbb{E}[e_B(w)^2].
\]

But we also know that

\[
\frac{1}{2} - \alpha \leq e_B(w) \leq \frac{pb}{2},
\]

where the first inequality comes from (13) while the second follows from \( D \cap C = \emptyset \). Moreover,

\[
\mathbb{E}[e_B(w)] = \frac{\rho_{bd}}{2pd}.
\]

Estimating the second moment in the standard way we obtain

\[
\mathbb{E}[e_B(w)^2] \leq \frac{\rho_{bd}}{2pd} \left( \frac{1}{2} - \alpha + \frac{pb}{2} \right) - \frac{pb}{2} \left( \frac{1}{2} - \alpha \right).
\]

Finally, plugging all our findings into (14) we obtain

\[
\beta \leq Q(\alpha, pb, \rho_{bd}) \defeq \frac{8\alpha^2 p_b^2 - 24\alpha p_b^3 - 4p_b^4 + 4\alpha p_b^2 - 8\alpha pb_{rd} + 12p_b^3 - 4p_b^2 \rho_{bd} - 3p_b^2 + 6p_b \rho_{bd} - 3\rho_{bd}^2}{16p_b^2}
\]

(see the Maple worksheet).

\( Q \) is quadratic concave in \( \rho_{bd} \) and, as before, \( \rho_{bd} \leq pbpd = pb(1 - \alpha - pb) \) since \( D \cap C = \emptyset \). Moreover,

\[
\frac{\partial Q}{\partial \rho_{bd}} \bigg|_{\rho_{bd}=pbpd} = \frac{pb - \alpha}{8pb} \geq 0.
\]

Hence

\[
Q(\alpha, pb, \rho_{bd}) \leq Q(\alpha, pb, pb(1 - \alpha - pb)) = \frac{13}{16} \alpha^2 - \frac{9}{8} \alpha pb - \frac{3}{16} p_b^2 - \frac{1}{4} \alpha + \frac{1}{2} pb \defeq Q_1(\alpha, pb).
\]

Finally, \( Q_1 \) is quadratic concave in \( pb \) and

\[
\frac{\partial Q_1}{\partial pb} \bigg|_{pb=\alpha} = \frac{1 - 3\alpha}{2} < 0
\]

(as \( \alpha \geq 3/8 \)). Since \( pb \geq \alpha \), we obtain

\[
Q_1(\alpha, pb) \leq Q_1(\alpha, \alpha) = \frac{\alpha}{2} \left( \frac{1}{2} - \alpha \right).
\]

This completes the proof of Theorem 3.6.
4.6. Graphs of girth $\geq 5$. In this section we prove Theorem 3.8, and for this particular proof we resort to absolute sizes of the sets involved rather than their densities. The reason is that the girth assumption does not survive blowing up a graph, and this makes the density-based language unnatural.

So we fix a triangle-free graph $G$ with $g(G) \geq 5$, $|V(G)| = n$, and let $v_0 \in V(G)$ be a vertex of the maximum degree $k$. We may assume that $k \leq (n - 1)/2$ (otherwise the result is trivial) and also that $G$ is a minimal counterexample to Erdős’s conjecture, that is, $\beta(G^*) \leq 1/50$ for any proper induced subgraph $G^*$ of $G$. We let

$$A \overset{\text{def}}{=} N_G(v_0) \quad \text{and} \quad B \overset{\text{def}}{=} V(G) \setminus (\{v_0\} \cup A).$$

Then $g(G) \geq 5$ implies that

$$\forall v \in B \quad (|N_G(v) \cap A| \leq 1). \tag{15}$$

We now apply the minimality assumption to the induced subgraph $G|_B$. This gives us a function $\nu: B \to [0,1]$ such that

$$\begin{cases}
\sum_{v \in B} \nu(v) = \frac{n - k - 1}{2}, \\
\sum_{(u,v) \in E(B)} \nu(u)\nu(v) \leq \frac{(n - k - 1)^2}{50}.
\end{cases} \tag{16}$$

We use it to define a half $\mu$ in the whole graph $G$ as follows:

$$\mu(v) \overset{\text{def}}{=} \begin{cases} 
0 & \text{if } v = v_0, \\
1 & \text{if } v \in A, \\
p
\nu(v) & \text{if } v \in B,
\end{cases}$$

where

$$p \overset{\text{def}}{=} \frac{n - 2k}{n - k - 1}.$$ 

Then we have

$$\beta(G, \mu) \leq \frac{1}{n^2} \left( \frac{(n - 2k)^2}{50} + \sum_{(u,v) \in E(A,B)} \mu(u)\mu(v) \right). \tag{17}$$

We employ two different methods for bounding the term $\sum_{(u,v) \in E(A,B)} \mu(u)\mu(v)$. Firstly, by (15) we have

$$\sum_{(u,v) \in E(A,B)} \mu(u)\mu(v) \leq \sum_{v \in B} \mu(v) = \frac{n}{2} - k,$$

and thus

$$\beta(G, \mu) \leq \frac{1}{n^2} \left( \frac{(n - 2k)^2}{50} + \frac{n}{2} - k \right)$$

$$= \frac{1}{50} - \frac{1}{5n^2} \left( n(4k - 25) + 50k - 4k^2 \right). \tag{18}$$
We now start a case analysis.

Case 1: \( k \geq 7 \).

In this case, since \( n \geq 2k + 1 \), we have

\[
n(4k - 25) + 50k - 4k^2 \geq (2k + 1)(4k - 25) + 50k - 4k^2 = 4k^2 + 4k - 25 \geq 0,
\]

and we are done by (18).

Case 2: \( k \leq 6 \).

This time, the same condition \( n(4k - 25) + 50k - 4k^2 < 0 \) (which can be assumed without lost of generality) provides a new lower bound on \( n \)

\[
n \geq \left\lceil \frac{2k(25 - 2k)}{25 - 4k} \right\rceil. \tag{19}
\]

To obtain an upper bound on \( n \), we estimate the term \( \sum_{(u, v) \in E(A, B)} \mu(u)\mu(v) \) as \((k - 1)k\), simply because the degree of any vertex in \( A \) is \( \leq k \), and all of them are adjacent to \( v_0 \not\in B \). Thus

\[
\beta(G, \mu) \leq \frac{1}{n^2} \left( \frac{(n - 2k)^2}{50} + (k - 1)k \right) = \frac{1}{50} - \frac{k}{25n^2}(2n + 25 - 27k).
\]

Hence we can also assume that

\[
n \leq \left\lceil \frac{27}{2}(k - 1) \right\rceil, \tag{20}
\]

which immediately rules out the case \( k = 1 \). Also, (19) and (20) rule out the case \( k = 6 \) as well, which leaves us with the possibilities \( k = 2, 3, 4, 5 \) and 80 potential values for the pair \((k, n)\).

Instead of trying to do the remaining analysis manually, we employ a different strategy. Namely, we record our argument in the form of ‘unprocessed’ (and recursive) bounds, without attempting to simplify them, and then we simply feed the formulae into Maple to finish the job.

To start with, let

\[
C \overset{\text{def}}{=} V(G) \setminus (A \cup N_G(A))
\]

be the set of vertices at distance \( \geq 3 \) from \( v_0 \); note that

\[
|C| \geq n - k^2 - 1.
\]

Let \( R(3, u) \) be the off-diagonal Ramsey number; we will only use the following well-known small values:

\[
R(3, 0) = 0, \quad R(3, 1) = 1, \quad R(3, 2) = 3, \quad R(3, 3) = 6, \quad R(3, 4) = 9, \quad R(3, 5) = 14.
\]

For every \( u \in [0, \lceil n/2 - k \rceil] \) such that \( R(3, u) \leq n - k - 1 \) (\( = |B| \)) we are going to derive its own bound \( \beta(G) \leq \beta_u \), and then we minimize over all choices of \( u \). So let us fix for the time being some \( u \) with the above properties.
Pick a subset $B_u \in \binom{B}{R(3,u)}$ with the only restriction that it contains as many vertices in $C$ as possible. Then we have

$$|E(A, B_u)| = |B_u \setminus C| = R(3,u) - |C| \leq R(3,u) - (n - k^2 - 1), \quad (21)$$

where $x \div y \overset{\text{def}}{=} \max(0, x - y)$. Finally, let $B'_u \subseteq B_u$ be an independent subset of size $u$ existing by the definition of the Ramsey numbers. Further analysis splits into two more cases.

**Case 2.1: $u = \lceil n/2 - k \rceil$.**

If $n$ is even, we take the half $A \cup B'_u$. If $n$ is odd, we can assume without loss of generality that $B'_u \cap N_G(A) \neq \emptyset$ (as otherwise we are done). Let $\mu$ be the half obtained from $A \cup B'_u$ by removing ‘half’ a vertex in $B'_u \cap N_G(A)$; this will give us an extra saving of ‘half’ an edge.

We have two different estimates on $|E(A, B'_u)|$: one follows from (21) and, on the other hand we, like before, have the trivial bound $|E(A, B'_u)| \leq n/2 - k \leq u$ coming from (15). Summarizing,

$$\beta(G) \leq \beta_u \overset{\text{def}}{=} \frac{1}{n^2} \left( \min(u, R(3,u) - (n - k^2 - 1)) \div \frac{1}{2}(n \mod 2) \right), \quad (22)$$

$$u = \left\lceil \frac{n}{2} - k \right\rceil.$$

Let us stress that this bound is defined only when $R(3, \lceil n/2 - k \rceil) \leq n - k - 1$.

**Case 2.2: $u < \lceil n/2 - k \rceil$.**

In this case we have

$$\beta(G) \leq \beta_u \overset{\text{def}}{=} \gamma(k, n, k + u, \min(u, R(3,u) - (n - k^2 - 1))), \quad (23)$$

where the function $\gamma(k, n, t, e) \ (t \leq n/2)$ abstracts our situation as follows:

$$\gamma(k, n, t, e) \overset{\text{def}}{=} \max_{G, A} \min_{\mu|A=1} \beta(G, \mu), \quad t \leq \left\lfloor \frac{n}{2} \right\rfloor.$$

Here $G$ runs over all graphs with $n$ vertices and $\Delta(G) \leq k$, $A$ runs over all sets of vertices with $|A| = t$, and $|E(A)| \leq e$ and $\mu$ runs over all halves containing $A$. What remains is to give sufficiently good (for our purposes) recursive bounds on $\gamma$.

First of all, when $n$ is even and $t = n/2$, we clearly have

$$\gamma\left(k, n, \frac{n}{2}, e\right) = \frac{e}{n^2}, \quad n \text{ is even.} \quad (24)$$

Next, assume that $n$ is odd and $t = (n - 1)/2$. Fix the worst-case $G$ and $A$, and let $e^* \leq e$ be the actual number of edges in $G|A$. Then $|E(A, V(G) \setminus A)|$ has at most $kt - 2e^*$ edges. Hence there exists a vertex $v \notin A$ with

$$|N_G(V) \cap A| \leq \left\lfloor \frac{kt - 2e^*}{n - t} \right\rfloor. \quad (25)$$

Adding ‘half’ of this vertex to $A$, we conclude that

$$\gamma(k, n, t, e) \leq \max_{0 \leq e^* \leq e} \frac{1}{n^2} \left( e^* + \frac{1}{2} \left( \frac{kt - 2e^*}{n - t} \right) \right), \quad n \text{ is odd, } t = \frac{n - 1}{2}. \quad (26)$$

\footnote{We do not need the half-edge saving from the previous case.}
Similarly, for smaller values of $t$ we apply recursion by letting $A := A \cup \{v\}$, where $v$ is the vertex satisfying (25). This gives us

$$\gamma(k, n, t, e) \leq \gamma\left(k, n, t + 1, \max_{0 \leq e^* \leq e} \left( e^* + \left\lfloor \frac{kt - 2e^*}{n-t} \right\rfloor \right) \right), \quad t < \left\lfloor \frac{n}{2} \right\rfloor. \quad (27)$$

This completes our description of $\beta_u$ in Case 2.2.

Finally, the ‘master formula’ now reads as

$$\beta(G) \leq \min \left\{ \beta_u \mid 0 \leq u \leq \left\lfloor \frac{n}{2} - k \right\rfloor \land R(3, u) \leq n - k - 1 \right\}. \quad (28)$$

The bounds (28), (22) and (23), along with the recursive estimates (24), (26) and (27) on the auxiliary function $\gamma$ suffice to complete the analysis of the 80 remaining cases (see the Maple worksheet for details).

Theorem 3.8 is proved.

§ 5. Conclusion

In this paper we have proved several partial results on Erdős’s half-graph conjecture. While they make this conjecture even more plausible, it still remains wide open. The same is true for the last of Erdős’s conjectures on this subject: prove that any triangle-free graph on $n$ vertices can be made bi-partite by removing at most $n^2/25$ edges.

As for intermediate, and probably more accessible, goals, we would like to ask to extend Theorem 3.6 to a neighbourhood of the critical value $\alpha = 2/5$, that is, prove the half-graph conjecture for triangle-free graphs $G$ with $\alpha(G) \geq 2/5 - \varepsilon$ for a fixed constant $\varepsilon > 0$. As we noted above, such an improvement is known for the minimum degree and the average degree (see [18] and [19]).

We have highlighted the extremal problem of finding the minimal density of quadrilaterals in triangle-free graphs with given edge density and have given its applications to the sparse half problem. Since this quantity can be viewed (actually, in a quite precise sense) as the measure of non-randomness in a graph, perhaps it might be worth studying in its own right.

Acknowledgement. I would like to thank Andrzej Grzesik and Jan Volec for pointing out to me the references [13] and [17] and for sharing with me the alternative argument sketched at the end of § 4.4.

Bibliography

[1] J. Balogh, F. C. Clemen and B. Lidický, Max cuts in triangle-free graphs, arXiv: 2103.14179.
[2] N. Biggs, Strongly regular graphs with no triangles, arXiv:0911.2160.
[3] A. E. Brouwer, “The uniqueness of the strongly regular graph on 77 points”, J. Graph Theory 7:4 (1983), 455–461.
[4] F. R. K. Chung, R. L. Graham and R. M. Wilson, “Quasi-random graphs”, Combinatorica 9:4 (1989), 345–362.
[5] P. Erdős, R. Faudree, J. Pach and J. Spencer, “How to make a graph bipartite”, J. Combin. Theory Ser. B 45:1 (1988), 86–98.
[6] P. Erdős, R. J. Faudree, C. C. Rousseau and R. H. Schelp, “A local density condition for triangles”, *Discrete Math.* **127**:1–3 (1994), 153–161.

[7] P. Erdős, E. Győri and M. Simonovits, “How many edges should be deleted to make a triangle-free graph bipartite?”, *Sets, graphs and numbers* (Budapest 1991), Colloq. Math. Soc. János Bolyai, vol. 60, North-Holland, Amsterdam 1992, pp. 239–263.

[8] P. Erdős, “Problems and results in graph theory and combinatorial analysis”, *Proceedings of the 5th British combinatorial conference* (Univ. Aberdeen, Aberdeen 1975), Congressus Numerantium, vol. 15, Utilitas Math., Winnipeg, MB 1976, pp. 169–192.

[9] P. Erdős, “On some problems in graph theory, combinatorial analysis and combinatorial number theory”, *Graph theory and combinatorics* (Cambridge 1983), Academic Press, London 1984, pp. 1–17.

[10] P. Erdős, “Some old and new problems in various branches of combinatorics”, *Discrete Math.* **165/166** (1997), 227–231.

[11] A. Gewirtz, “Graphs with maximal even girth”, *Canadian J. Math.* **21** (1969), 915–934.

[12] A. Gewirtz, “The uniqueness of \(g(2, 2, 10, 56)\)”, *Trans. New York Acad. Sci.*, II Ser. **31**:6 (1969), 656–675.

[13] A.L. Gavrilyuk and A.A. Makhnev, “On Krein graphs without triangles”, *Dokl. Akad. Nauk* **403**:6 (2005), 727–730; English transl. in *Dokl. Math.* **72**:1 (2005), 591–594.

[14] J. Hatami, J. Hladký, D. Kráľ, S. Norine and A. Razborov, “Non-three-colourable common graphs exist”, *Combin. Probab. Comput.* **21**:5 (2012), 734–742.

[15] J. Hatami, J. Hladký, D. Kráľ, S. Norine and A. Razborov, “On the number of pentagons in triangle-free graphs”, *J. Combin. Theory Ser. A* **120**:3 (2013), 722–732.

[16] P. Keevash and B. Sudakov, “Sparse halves in triangle-free graphs”, *J. Combin. Theory Ser. B* **96**:4 (2006), 614–620.

[17] S. Norin and L. Yepremyan, “Sparse halves in dense triangle-free graphs”, *J. Combin. Theory Ser. B* **115** (2015), 1–25.

[18] A.A. Razborov, “Flag algebras”, *J. Symbolic Logic* **72**:4 (2007), 1239–1282.