Combinatorics, Probability and Computing (2017) 26, 525–535. © Cambridge University Press 2017
doi:10.1017/S0963548316000407

The Minimum Number of Triangular Edges and a Symmetrization Method for Multiple Graphs

ZOLTÁN FÜREDI[1†] and ZEINAB MALEKI[2]

[1] Alfréd Rényi Institute of Mathematics, 13–15 Reáltanoda Street, 1053 Budapest, Hungary
(e-mail: z-furedi@illinois.edu)

[2] Department of Mathematical Sciences, Isfahan University of Technology, Isfahan 84156-83111, Iran
(e-mail: zmaleki@math.iut.ac.ir)

Received 12 November 2014; revised 29 October 2016; first published online 23 January 2017

This work was done while the authors visited the Department of Mathematics and Computer Science, Emory University, Atlanta, GA, USA.

We give an asymptotic formula for the minimum number of edges contained in triangles among graphs with n vertices and e edges. Our main tool is a generalization of Zykov's symmetrization method that can be applied to several graphs simultaneously.

2010 *Mathematics subject classification*: Primary 05C35
Secondary 05C22, 05D99

1. Graphs with few triangular edges

Erdős, Faudree and Rousseau [3] showed that a graph on n vertices and at least $\lfloor n^2/4 \rfloor + 1$ edges has at least $2\lfloor n/2 \rfloor + 1$ edges in triangles. To see that this result is sharp, consider the graph obtained by adding one edge to the larger side of the complete bipartite graph $K_{\lceil n/2 \rceil, \lfloor n/2 \rfloor}$. We consider a more general problem, where the number of edges may be larger than $\lfloor n^2/4 \rfloor + 1$. Given a graph G, let $\mathrm{Tr}(G)$ denote the number of edges of G contained in triangles, and let $\mathrm{Tr}(n,e) := \min\{\mathrm{Tr}(G) : |V(G)| = n, e(G) = e\}$. With this notation the above result of Erdős, Faudree and Rousseau can be reformulated as

$$\mathrm{Tr}(n, \lfloor n^2/4 \rfloor + 1) = 2\lfloor n/2 \rfloor + 1. \tag{1.1}$$

Note that $\mathrm{Tr}(n,e) = 0$ whenever $e \leqslant n^2/4$, because in that case there exist triangle-free (even bipartite) graphs with n vertices and e edges. To avoid trivialities, we usually implicitly assume that $e > n^2/4$.

† Research was supported in part by grant K104343 from the Hungarian National Research, Development and Innovation Office (NKFIH), by the Simons Foundation Collaboration, grant 317487, and by the European Research Council Advanced Investigators, grant 267195.

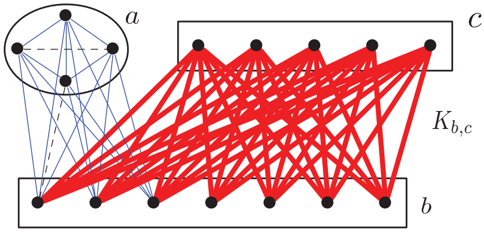


Figure 1. A graph from $\mathcal{G}(a,b,c)$.

Given integers a, b and c, ($a \geqslant 2$), we define a family of graphs $\mathcal{G}(a,b,c)$ as follows (see Figure 1). The vertex set V of a graph G in this class has a partition $V = A \cup B \cup C$ where $|A| = a$, $|B| = b$ and $|C| = c$, such that B and C are independent sets, $B \cup C$ induces a complete bipartite graph $K_{b,c}$, the vertices of C have neighbours only in B, and $G[A]$ and $G[A,B]$ are 'almost complete graphs', namely, they span more than $\binom{|A|-1}{2} + |A||B|$ edges. The edges of $G[B,C]$ are the non-triangular edges.

Given integers $n \geqslant 3$ and $n^2/4 < e \leqslant \binom{n}{2}$, we define a class of graphs $\mathcal{G}(n,e)$ with many non-triangular edges as follows. Let $\mathcal{G}(n,e)$ denote the set of graphs with n vertices and e edges that belong to a class $\mathcal{G}(a,b,c)$. Define $g(n,e)$ as $\min\{\mathrm{Tr}(G) : G \in \mathcal{G}(n,e)\}$. We have

$$\mathrm{Tr}(n,e) \leqslant g(n,e) = \min\left\{ e - bc : a+b+c = n,\, a,b,c \in \mathbb{N} \cup \{0\},\, \binom{a}{2} + ab + bc \geqslant e \right\}. \quad (1.2)$$

We believe that one can extend the theorem of Erdős, Faudree and Rousseau [3] as follows.

Conjecture 1.1. *Suppose that G is an n-vertex graph with e edges, such that $e > n^2/4$ and it has the minimum number of triangular edges, that is,* $\mathrm{Tr}(G) = \mathrm{Tr}(n,e)$. *Then* $G \in \mathcal{G}(n,e)$.

In particular, we conjecture that $\mathrm{Tr}(n,e) = g(n,e)$. We prove a slightly weaker result.

Theorem 1.2. *For $e > n^2/4$ we have* $g(n,e) - (3/2)n \leqslant \mathrm{Tr}(n,e) \leqslant g(n,e)$.

Our main tool, presented in Section 2, is a new symmetrization method that generalizes previous results by Motzkin and Straus such that they can be applied to more than one graph simultaneously.

In Section 3 we use the new symmetrization method to prove a lemma about triangular edges of a given graph. In Section 4, using the lemma of Section 3, we complete the proof of Theorem 1.2. In Section 5 we introduce more problems for future research, some of which may be solved using our methods (see [5]).

2. The symmetrization method

In this section, we describe Zykov's symmetrization process [10]. It starts with a K_p-free graph G with vertex set $\{v_1, \ldots, v_n\}$, and at each step takes two non-adjacent vertices v_i and v_j such that $\deg(v_i) > \deg(v_j)$ and replaces all edges incident to v_j by new edges incident to v_j and to the neighbourhood $N(v_i)$. We do the same if $\deg(v_i) = \deg(v_j)$, $N(v_i) \neq N(v_j)$ and $i < j$.

https://doi.org/10.1017/S0963548316000407 Published online by Cambridge University Press

Symmetrization does not increase the size of the largest clique and does not decrease the number of edges. When the process terminates it yields a complete multipartite graph with at most $p-1$ parts.

Using this process, Zykov [10] gave an alternative proof of Turán's theorem which states that the number of edges of a K_p-free graph is at most as large as in a complete $(p-1)$-partite graph with almost equal parts.

At first sight, it seems that this method cannot be used directly to determine $\mathrm{Tr}(n,e)$ because we need to increase simultaneously the number of edges and the number of non-triangular edges. In the rest of the section this method will be generalized to settings involving more than one graph.

Let us recall a continuous version of Zykov's symmetrization method, due to Motzkin and Straus [8]. Given a graph G with vertex set $\{v_1,\dots,v_n\}$, define a real polynomial

$$f(G,\mathbf{x}) := \sum\{x_i x_j : v_i v_j \in E\}.$$

Define a simplex

$$S_n := \Big\{\mathbf{x} \in \mathbb{R}^n : \forall x_i \geqslant 0 \text{ and } \sum x_i = 1\Big\}.$$

Let $f(G) := \max\{f(G,\mathbf{x}) : \mathbf{x} \in S_n\}$. Motzkin and Straus [8] provided an alternative proof of an asymptotic version of Turán's theorem by observing a remarkable connection between the clique number, $\omega(G)$, and $f(G)$. They proved that $f(G) = (\omega-1)/(2\omega)$. Their main tool was a continuous version of Zykov's symmetrization, as follows.

Theorem 2.1 (Motzkin and Straus [8]). *Given a graph G on n vertices and a vector $\mathbf{x} \in S_n$, there exists $\mathbf{y} \in S_n$ such that $f(G,\mathbf{x}) \leqslant f(G,\mathbf{y})$ and* $\mathrm{support}(\mathbf{y})$ *induces a complete subgraph.*

We generalize this result so that it can be applied to several graphs simultaneously.

Theorem 2.2. *Let G be a graph on n vertices and let $G_1, G_2, \dots, G_d$ be subgraphs of G with the same vertex set. For every $\mathbf{x} \in S_n$ there exists a subset $K \subseteq V(G)$ and a vector $\mathbf{y} \in S_n$ with support K such that $f(G_i,\mathbf{x}) \leqslant f(G_i,\mathbf{y})$ for every $1 \leqslant i \leqslant d$ and $\alpha(G[K]) \leqslant d$.*

To prove Theorem 2.2 we need the following lemma.

Lemma 2.3. *Suppose that $\mathbf{a}_1,\dots,\mathbf{a}_d \in \mathbb{R}^{d+1}$. Then there exists a non-zero vector $\mathbf{z} \in \mathbb{R}^{d+1}$ such that $\mathbf{a}_i^T\mathbf{z} \geqslant 0$ for every $1 \leqslant i \leqslant d$ and the sum of the coordinates is 0, namely $\sum_{1 \leqslant i \leqslant d+1} z_i = 0$.*

Proof. Let $\mathbf{j} \in \mathbb{R}^{d+1}$ be the all-1 vector and define the matrix A as $(\mathbf{a}_1,\dots,\mathbf{a}_d,\mathbf{j})$. If $\det(A)=0$, then there are non-trivial solutions of $A^T\mathbf{z} = \mathbf{0}$. If $\det(A) \neq 0$, define $\mathbf{a} := (1,\dots,1,0)^T \in \mathbb{R}^{d+1}$. There is a unique solution $\mathbf{z}$ of $A^T\mathbf{z} = \mathbf{a}$. Clearly $\mathbf{z} \neq 0$, so we are done. □

A remark on keeping equalities. Note that if we take $\mathbf{a} := \mathbf{e}_\ell \in \mathbb{R}^{d+1}$ (the ℓth unit vector) for some $1 \leqslant \ell \leqslant d$ instead of $\mathbf{a} := (1,\dots,1,0)^T \in \mathbb{R}^{d+1}$ in the second half of the above proof (*i.e.* in the case $\det(A) \neq 0$), then we can obtain a sharper version of Lemma 2.3. Namely, there exists a

https://doi.org/10.1017/S0963548316000407 Published online by Cambridge University Press

non-zero vector $\mathbf{z} \in \mathbb{R}^{d+1}$ such that $\sum_{1\leqslant i\leqslant d+1} z_i = 0$ and $\mathbf{a}_\ell^T\mathbf{z} \geqslant 0$, but $\mathbf{a}_i^T\mathbf{z} = 0$ for every $1 \leqslant i \leqslant d$, $i \neq \ell$.

Proof of Theorem 2.2. Let $\mathbf{y} \in S_n$ be a vector whose support has minimum size among vectors $\mathbf{y}' \in S_n$ satisfying $f(G_i,\mathbf{x}) \leqslant f(G_i,\mathbf{y}')$ for every $1 \leqslant i \leqslant d$. If $\{v_1,v_2,\dots,v_{d+1}\} \subseteq \operatorname{support}(\mathbf{y})$ is an independent set, then for any $\mathbf{z} = (z_1,\dots,z_{d+1},0,0,\dots)^T \in \mathbb{R}^n$, $t \in \mathbb{R}$, and $1 \leqslant i \leqslant d$ we have $f(G_i,\mathbf{y}+t\mathbf{z}) = f(G_i,\mathbf{y}) + t(\mathbf{a}_i^T\mathbf{z})$ for some $\mathbf{a}_i \in \mathbb{R}^{d+1}$. Here $\mathbf{a}_i$ depends only on G_i and $\mathbf{y}$, not on $\mathbf{z}$ or t. Apply Lemma 2.3 to obtain a non-zero vector $\mathbf{z} = (z_1,\dots,z_{d+1},0,0,\dots)^T$ with $\sum_{1\leqslant i\leqslant d+1} z_i = 0$ and $\mathbf{a}_i^T\mathbf{z} \geqslant 0$ for $1 \leqslant i \leqslant d$. Choosing an appropriate $t > 0$ we have $\mathbf{y}+t\mathbf{z} \in S_n$ and $\operatorname{support}(\mathbf{y}+t\mathbf{z}) \subseteq \operatorname{support}(\mathbf{y}) - \{v_j\}$ for some $1 \leqslant j \leqslant d+1$. This is a contradiction, so $\mathbf{y}$ has the desired property. □

3. Maximizing the weight of non-triangular edges in a weighted graph

In this section, by using the power of Theorem 2.2, we prove a lemma concerning the weight of non-triangular edges in a weighted graph as follows.

Lemma 3.1. *Let G_1 be a graph on n vertices and G_2 be a subgraph of G_1 whose edges are some of the non-triangular edges of G_1, $E(G_2) \neq \emptyset$. For every $\mathbf{x} \in S_n$ there exists a subset $K \subseteq V$ and a vector $\mathbf{y} \in S_n$ with support K such that $f(G_1,\mathbf{x}) \leqslant f(G_1,\mathbf{y})$ and $f(G_2,\mathbf{x}) \leqslant f(G_2,\mathbf{y})$. Furthermore, the graph $H := G_1[K]$ contains exactly one edge e of G_2 and $H \setminus V(e)$ is a complete graph.*

Proof. By Theorem 2.2, we know that there is a vector $\mathbf{y} \in S_n$ such that $f(G_1,\mathbf{x}) \leqslant f(G_1,\mathbf{y})$, $f(G_2,\mathbf{x}) \leqslant f(G_2,\mathbf{y})$ and $\alpha(H) \leqslant 2$. Let $V = \{v_1,\dots,v_n\}$ and $\mathbf{y} = (y_1,\dots,y_n)$ be such a vector whose support has minimal size. We claim that $K := \operatorname{support}(\mathbf{y})$ satisfies the required properties. First we show that the structure of $G_2[K]$ is rather simple; then we show that by finding an appropriate $\mathbf{y}'$ one can further reduce K if $G_2[K]$ has two or more edges.

Recall that $(\partial/\partial z_k)f$ stands for the partial derivative of the function $f(z_1,z_2,\dots,z_n)$ with respect to the variable z_k. Suppose that v_k and $v_h \in K$ are non-adjacent vertices such that

$$\frac{\partial}{\partial y_k} f(G_1,\mathbf{y}) \geqslant \frac{\partial}{\partial y_h} f(G_1,\mathbf{y}) \quad \text{and} \quad \frac{\partial}{\partial y_k} f(G_2,\mathbf{y}) \geqslant \frac{\partial}{\partial y_h} f(G_2,\mathbf{y}). \tag{3.1}$$

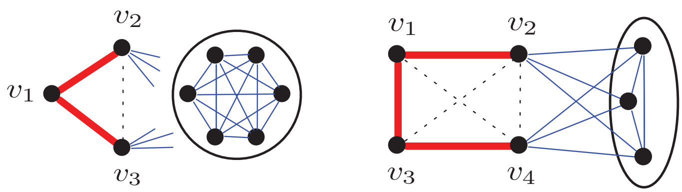
In other words,

$$\sum\{y_\ell : v_k v_\ell \in E(G_i[K])\} \geqslant \sum\{y_\ell : v_h v_\ell \in E(G_i[K])\} \quad \text{for } i = 1,2.$$

Define the vector $\mathbf{y}' \in S_n$ by

$$y'_\ell = \begin{cases} y_k + y_h & \ell = k, \\ 0 & \ell = h, \\ y_\ell & \text{otherwise.} \end{cases}$$

We have $f(G_i,\mathbf{y}) \leqslant f(G_i,\mathbf{y}')$ for $i \in \{1,2\}$ and $\operatorname{support}(\mathbf{y}') = K \setminus \{v_h\}$, a contradiction. We conclude that such v_k and v_h do not exist.



(a) Case 1a (b) Case 1b

Figure 2. The structure of H in Cases 1a and 1b. The G_2 edges are bold.

Without loss of generality, we may assume that v_1v_2 is a G_2-edge of H. From now on, in this section if we talk about 'edges', 'degrees', *etc.*, then we always mean H-edges, degree in H, *etc.*, unless otherwise stated.

If $f(G_1,\mathbf{y}) \leqslant 1/4$ then define $\mathbf{y}' = (1/2, 1/2, 0, \ldots, 0)$. We obtain

$$f(G_2,\mathbf{y}) \leqslant f(G_1,\mathbf{y}) \leqslant 1/4 = f(G_1,\mathbf{y}') = f(G_2,\mathbf{y}').$$

This implies $K = \{1,2\}$, and we are done. So from now on, we suppose that $f(G_1,\mathbf{y}) > 1/4$. Then the Motzkin–Straus theorem implies that the graph H is not triangle-free.

Claim 3.2. *No two edges of $G_2[K]$ are adjacent.*

Proof of Claim 3.2. Assume to the contrary that v_1v_2 and $v_1v_3 \in E(H)$ are G_2 edges. We claim that

$$v_2 \text{ and } v_3 \text{ are non-adjacent, } \deg(v_1) = 2, \text{ and } H \setminus \{v_1,v_2,v_3\} \text{ is a complete graph.} \tag{3.2}$$

Indeed, v_2 and v_3 are non-adjacent, otherwise the triangle $v_1v_2v_3$ contains G_2 edges. Suppose to the contrary that $|N(v_1)| > 2$, that is, there exists a vertex $v_4 \neq v_2, v_3$ such that $v_1v_4 \in E(H)$. Since $\alpha(H) \leqslant 2$ and $v_2v_3 \notin E(H)$, without loss of generality, $v_3v_4 \in E(H)$. Then the triangle $v_1v_3v_4$ contains a G_2 edge (namely v_1v_3), a contradiction, so we must have $N(v_1) = \{v_2, v_3\}$. Finally, the condition $\alpha(H) \leqslant 2$ implies that $K \setminus (N(v_1) \cup \{v_1\})$ induces a complete graph (see Figure 2).

The statement (3.2) already implies that the structure of G_2 edges is rather simple in H. Using condition (3.1) and other techniques, we reach a contradiction, considering three possible cases.

Case 1a. Assume that there is no G_2 edge connecting $\{v_1,v_2,v_3\}$ to $K \setminus \{v_1,v_2,v_3\}$.

Then

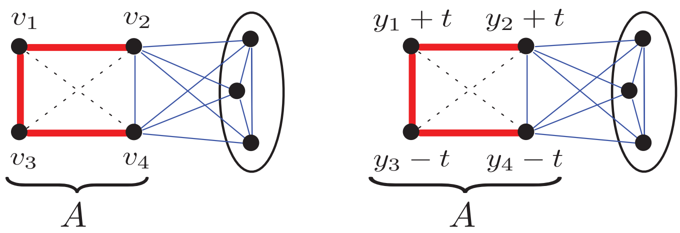
$$\frac{\partial}{\partial y_2} f(G_2,\mathbf{y}) = \frac{\partial}{\partial y_3} f(G_2,\mathbf{y}) = y_1.$$

Since v_2 and v_3 are non-adjacent, the conditions of (3.1) hold, a contradiction.

Case 1b. Assume that there is a G_2 edge, say v_3v_4, connecting $\{v_1,v_2,v_3\}$ to $K \setminus \{v_1,v_2,v_3\}$ such that $v_2v_4 \notin E(H)$.

According to (3.2) the set $A := \{v_1, \ldots, v_4\}$ spans only these three G_2 edges, v_1 and v_3 are degree 2 vertices, and $(K \setminus A) \cup \{v_i\}$ are complete graphs for $i \in \{2,4\}$. Since H must contain triangles we have $|K \setminus A| \geqslant 2$, and H does not contain a further G_2 edge (see Figure 2). Suppose

https://doi.org/10.1017/S0963548316000407 Published online by Cambridge University Press



(a) the structure of H in Case 1c (b) the change of the weights

Figure 3. Case 1c.

that $y_1 \geqslant y_3$. We obtain that

$$\frac{\partial}{\partial y_2} f(G_2,\mathbf{y}) = y_1 \geqslant \frac{\partial}{\partial y_4} f(G_2,\mathbf{y}) = y_3$$

and

$$\frac{\partial}{\partial y_2} f(G_1,\mathbf{y}) = y_1 + \sum_{\ell>4} y_\ell \geqslant \frac{\partial}{\partial y_4} f(G_1,\mathbf{y}) = y_3 + \sum_{\ell>4} y_\ell.$$

Since v_2 and v_4 are non-adjacent, this contradicts condition (3.1).

Case 1c. Assume that there is a G_2 edge, say v_3v_4, connecting $\{v_1,v_2,v_3\}$ to $K \setminus \{v_1,v_2,v_3\}$ such that $v_2v_4 \in E(H)$.

According to (3.2) the set $A := \{v_1,\dots,v_4\}$ spans only these four edges, v_1 and v_3 have degree 2, and $K \setminus \{v_1,v_3\}$ is a complete graph of size at least 3 (see Figure 3). H does not contain other G_2 edges. We have

$$f(G_1,\mathbf{y}) = (y_1+y_4)(y_2+y_3) + (y_2+y_4)\left(\sum_{\ell>4} y_\ell\right) + \sum_{i>j>4}\sum y_i y_j$$

and

$$f(G_2,\mathbf{y}) = y_1y_2 + y_1y_3 + y_3y_4.$$

Substitute $\mathbf{y}' := \mathbf{y}'(t) = \mathbf{y} + t(\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 - \mathbf{e}_4)$ into the above equations (Figure 3) where $\mathbf{e}_i$ is the unit vector with 1 at the ith coordinate and 0 elsewhere. Note that $\mathbf{y}' \in S_n$ if

$$t \in I := [\max\{-y_1,-y_2\}, \min\{y_3,y_4\}].$$

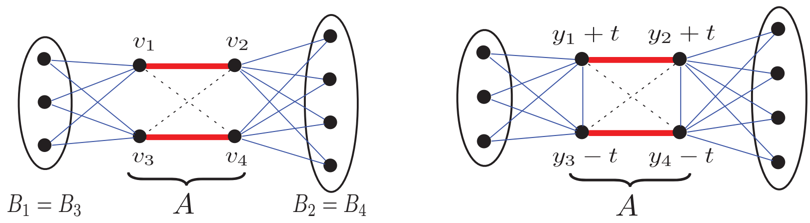
We get $f(G_1,\mathbf{y}') = f(G_1,\mathbf{y})$ and

$$f(G_2,\mathbf{y}') - f(G_2,\mathbf{y}) = t^2 + t(y_2 - y_4).$$

The right-hand side is a convex polynomial in t and it takes its maximum on I in one of the endpoints. Taking this optimal t we obtain that $\max_{t\in I} f(G_2,\mathbf{y}') > f(G_2,\mathbf{y})$ and $|\text{support}(\mathbf{y}')| < |\text{support}(\mathbf{y})|$, a contradiction. This completes the proof of Claim 3.2 that H has no adjacent G_2 edges. □

Claim 3.3. *$G_2[K]$ does not have two independent edges.*

https://doi.org/10.1017/S0963548316000407 Published online by Cambridge University Press



(a) the structure of H in Case 2 (b) the change of weights in Case 2b

Figure 4. Case 2.

Proof of Claim 3.3. According to Claim 3.2, $G_2[K]$ is a matching, $\{v_1v_2, v_3v_4, \ldots, v_{2k-1}v_{2k}\}$. We will show $k = 1$. Assume to the contrary that v_1v_2 and v_3v_4 are two disjoint G_2 edges of H.

Define $A := \{v_1, \ldots, v_4\}$. Since v_1v_2 and v_3v_4 are two non-triangular edges, the set A can contain at most two more edges of H, and these should be disjoint. So without losing generality, we may assume that $v_1v_4, v_2v_3 \notin E(G_1)$ (see Figure 4).

Let $B_i := \{v \in K \setminus A : vv_i \in E(H)\}$ for $1 \leqslant i \leqslant 4$. We claim that $B_1 = B_3$. Indeed, if $v_5 \in B_1$ then $v_2v_5 \notin E(G_1)$, otherwise $\{v_1, v_2, v_5\}$ forms a triangle. Then $v_5v_3 \in E(H)$ otherwise $\{v_2, v_3, v_5\}$ forms an independent set. Hence $v_5 \in B_3$, implying $B_1 \subseteq B_3$. By symmetry $B_3 \subseteq B_1$, and we obtain $B_1 = B_3$ and similarly $B_2 = B_4$.

Since v_1v_2 is a G_2 edge we have $B_1 \cap B_2 = \emptyset$ (in fact $\{A, B_1, B_2\}$ is a partition of K). We distinguish two cases.

Case 2a. Assume first that $v_1v_3 \notin E(H)$.

Suppose that $y_2 \geqslant y_4$. Since no G_2-edge joins A to $K \setminus A$ and $B_1 = B_3$, we obtain that

$$\frac{\partial}{\partial y_1} f(G_2, \mathbf{y}) = y_2 \geqslant y_4 = \frac{\partial}{\partial y_3} f(G_2, \mathbf{y})$$

and

$$\frac{\partial}{\partial y_1} f(G_1, \mathbf{y}) = y_2 + \sum_{y_\ell \in B_1} y_\ell \geqslant y_4 + \sum_{y_\ell \in B_1} y_\ell = \frac{\partial}{\partial y_3} f(G_1, \mathbf{y}).$$

Since v_1 and v_3 are non-adjacent, this contradicts (3.1).

So we may assume that A contains the edge v_1v_3. By symmetry, we may assume that A contains the edge v_2v_4 too.

Case 2b. Finally, A contains the edges v_1v_3 and v_2v_4 (see Figure 4).

We have

$$f(G_1, \mathbf{y}) = (y_1 + y_4)(y_2 + y_3) + (y_1 + y_3)\left(\sum_{y_\ell \in B_1} y_\ell\right) + (y_2 + y_4)\left(\sum_{y_\ell \in B_2} y_\ell\right) + \sum_{v_i, v_j \notin A,} \sum_{v_iv_j \in E(H)} y_iy_j$$

and

$$f(G_2, \mathbf{y}) = y_1y_2 + y_3y_4 + \cdots + y_{2k-1}y_{2k}.$$

https://doi.org/10.1017/S0963548316000407 Published online by Cambridge University Press

Substitute $\mathbf{y}' := \mathbf{y}'(t) = \mathbf{y} + t(\mathbf{e}_1 + \mathbf{e}_2 - \mathbf{e}_3 - \mathbf{e}_4)$ into the above equations. Note that $\mathbf{y}' \in S_n$ if

$$t \in I := [\max\{-y_1, -y_2\}, \min\{y_3, y_4\}].$$

We get $f(G_1, \mathbf{y}') = f(G_1, \mathbf{y})$ and

$$f(G_2, \mathbf{y}') - f(G_2, \mathbf{y}) = 2t^2 + t(y_1 + y_2 - y_3 - y_4).$$

The right-hand side is convex; it takes its maximum on I in one of the endpoints. Taking this optimal t we obtain that $\max_{t \in I} f(G_2, \mathbf{y}') > f(G_2, \mathbf{y})$ and $|\text{support}(\mathbf{y}')| < |\text{support}(\mathbf{y})|$, a contradiction. This completes the proof of Claim 3.3. □

Completion of the proof of Lemma 3.1. Claims 3.2 and 3.3 imply that H has a unique G_2 edge. We claim that the vertices in H which are not adjacent to any G_2 edge of H induce a clique. To see this, consider two such vertices v_i and v_j. We have

$$\frac{\partial}{\partial y_i} f(G_2, \mathbf{y}) = 0 = \frac{\partial}{\partial y_j} f(G_2, \mathbf{y}),$$

so the inequalities of (3.1) hold. Therefore v_i and v_j must be adjacent to avoid a contradiction. □

4. A continuous lower bound for the number of triangular edges

In this section, by using Lemma 3.1, we will prove the main result of this paper, Theorem 1.2. Recall that

$$g(n,e) := \min\left\{ e - bc : a + b + c = n,\, a,b,c \in \mathbb{N} \cup \{0\},\, \binom{a}{2} + ab + bc \geqslant e \right\}$$

(see (1.2)). We define $t(n,e)$ to be a real-valued version of $g(n,e)$ as follows:

$$t(n,e) := \min\left\{ e - bc : a + b + c = n,\, a,b,c \in \mathbb{R}_+,\, \frac{1}{2}a^2 + ab + bc \geqslant e \right\}. \tag{4.1}$$

Obviously, $t(n,e) \leqslant g(n,e)$ for $n^2/4 \leqslant e \leqslant \binom{n}{2}$. Furthermore,

$$g(n,e) - (3/2)n \leqslant t(n,e). \tag{4.2}$$

Indeed, suppose that $(a,b,c) \in \mathbb{R}_+^3$ yields the optimal value, $t(n,e) = e - bc$. It is a straightforward calculation to show that the choice of $(a',b',c') := (\lceil a+1 \rceil, \lceil b \rceil, n - a' - b')$ satisfies (1.2) and the difference between $(e - b'c')$ and $(e - bc)$ is at most $(3/2)n$.

We cannot prove Conjecture 1.1 that $g(n,e) \leqslant \text{Tr}(n,e)$ (*i.e.* $g(n,e) = \text{Tr}(n,e)$), but as an application of Lemma 3.1 we will show that $t(n,e)$ is a lower bound for $\text{Tr}(n,e)$.

Theorem 4.1. *For $e > n^2/4$ we have $t(n,e) \leqslant \text{Tr}(n,e)$.*

Proof. Suppose that G_1 is a graph with n vertices, e edges and minimum number of edges in triangles, that is, G_1 has $\text{Tr}(n,e)$ triangle edges. Let G_2 be the subgraph of G_1 consisting of the edges not in any triangle of G_1. Consider the vector $(1/n)\mathbf{j} = (1/n, 1/n, \ldots, 1/n) \in \mathbb{R}^n$. By Lemma 3.1 there exists a $\mathbf{y} = (y_1, \ldots, y_n) \in S_n$ with support K such that $G_2[K]$ consists of a single edge, say $v_1 v_2$. Moreover,

$$\frac{e}{n^2} = f(G_1, (1/n)\mathbf{j}) \leqslant f(G_1, \mathbf{y}) \tag{4.3}$$

and

$$\frac{e - \mathrm{Tr}(n,e)}{n^2} = f(G_2, (1/n)\mathbf{j}) \leqslant f(G_2, \mathbf{y}) = y_1 y_2. \tag{4.4}$$

Assume that $y_1 \geqslant y_2$ and define

$$a := \left(\sum_{k \neq 1,2} y_k\right) n, \quad b := y_1 n, \quad c := y_2 n$$

Then (4.4) yields that $\mathrm{Tr}(n,e) \geqslant e - bc$. We claim that the reals a, b, and c satisfy the constraints in (4.1), hence $e - bc \geqslant t(n,e)$, completing the proof.

Indeed, since $v_1 v_2$ is not in any triangle, $N(v_1) \cap N(v_2) = \emptyset$, we get from (4.3) that

$$\begin{aligned} \frac{e}{n^2} &\leqslant f(G_1, \mathbf{y}) \\ &= y_1 y_2 + y_1 \left(\sum_{y_k \in N(v_1), k \neq 2} y_k\right) + y_2 \left(\sum_{y_k \in N(v_2), k \neq 1} y_k\right) + \sum_{i<j,\, i,j \neq 1,2} \sum y_i y_j \\ &\leqslant \frac{bc}{n^2} + \frac{b}{n} \times \frac{a}{n} + \frac{1}{2}\left(\frac{a}{n}\right)^2. \end{aligned}$$

□

5. Further problems, minimizing C_{2k+1} edges

In addition to the question of minimizing the number of triangular edges, Erdős, Faudree and Rousseau [3] also considered a conjecture of Erdős [2] regarding pentagonal edges asserting that a graph on n vertices and at least $\lfloor n^2/4 \rfloor + 1$ edges has at most $n^2/36 + O(n)$ non-pentagonal edges. This value can be obtained by considering a graph having two components: a complete graph on $\lceil 2n/3 \rceil + 1$ vertices and a complete bipartite graph on the rest. This conjecture was mentioned in the papers of Erdős [2] and also in the problem book by Fan Chung and Graham [1].

Erdős, Faudree and Rousseau [3] proved that if G is a graph with n vertices and at least $\lfloor n^2/4 \rfloor + 1$ edges, then for any fixed $k \geqslant 2$ at least $\frac{11}{144} n^2 - O(n)$ edges of G are in cycles of length $2k+1$. So there is a jump of $\Omega(n^2)$ in the number of C_5-edges, while the construction of $\mathcal{G}(n,e)$ shows that for K_3-edges the change is smoother, $\mathrm{Tr}(n, n^2/4 + x) = O(n\sqrt{x})$.

In a forthcoming paper [5] we give an example of graphs with $\lfloor n^2/4 \rfloor + 1$ edges and

$$n^2/8(2+\sqrt{2}) + O(n) = n^2/27.31\ldots$$

non-pentagonal edges, disproving Erdős's conjecture. By using the weighted symmetrization method we show that this coefficient is asymptotically the best possible for $e > (n^2/4) + o(n^2)$. On the other hand, we asymptotically establish the conjecture of Erdős that for every $k \geqslant 3$ the maximum number of non-C_{2k+1} edges in a graph of size exceeding $(n^2/4) + o(n^2)$ is at most $n^2/36 + o(n^2)$, as in the graph of two components described above.

More generally, given a graph F, one can define $h(n,e,F)$ as the minimum number of F-edges among all graphs of n vertices and e edges. In our forthcoming paper [5] we asymptotically determine $h(n,\lambda n^2,F)$ for any fixed λ, when $1/4 < \lambda < 1/2$ and F is 3-chromatic. Many problems, for example an F with a higher chromatic number, or natural generalizations for hypergraphs, remain open.

A remark on very dense graphs. One can verify Conjecture 1.1 for $n \leqslant 8$ and in general for $e \geqslant \binom{n}{2} - (3n-13)$. This and (1.1) yield the exact value of $\mathrm{Tr}(n,e)$ for all pairs with $n \leqslant 10$ except $\mathrm{Tr}(10,27)$. More details can be found in the arXiv version of this paper [4].

More remarks on keeping equalities. The sharper version of Lemma 2.3 (mentioned after its proof) yields a sharper version of Theorem 2.2. Namely, there exists an appropriate vector $\mathbf{y} \in S_n$ such that $f(G_\ell,\mathbf{x}) \leqslant f(G_\ell,\mathbf{y})$ and $f(G_i,\mathbf{x}) = f(G_i,\mathbf{y})$ for every $1 \leqslant i \leqslant d$, $i \neq \ell$.

Then the proof of Lemma 3.1 can be adjusted so that given $\ell \in \{1,2\}$ one can find an appropriate vector $\mathbf{y} \in S_n$ such that $f(G_\ell,\mathbf{x}) = f(G_\ell,\mathbf{y})$ and $f(G_{3-\ell},\mathbf{x}) \leqslant f(G_{3-\ell},\mathbf{y})$.

5.1. New developments (as of May 2016)

Since the first public presentations of our results (for example in the Combinatorics seminar of the Department of Mathematics and Computer Science at Emory University, 6 December 2013, and in the Oberwolfach Combinatorics Workshop, 5–11 January 2014) and posting the present manuscript on arXiv [4] on 4 November 2014, there have been (at least) two remarkable achievements.

Gruslys and Letzter [6], using a refined version of the symmetrization method, proved that there exists an n_0 such that $\mathrm{Tr}(n,e) = g(n,e)$ for all $n > n_0$. The second part of our Conjecture 1.1, namely that the extremal graph should be from a $\mathcal{G}(a,b,c)$, is still open.

Grzesik, P. Hu and Volec [7], using Razborov's flag algebra method, showed that every n-vertex graph with $\lfloor n^2/4 \rfloor + 1$ edges has at least $n^2/8(2+\sqrt{2}) - \varepsilon n^2$ pentagonal edges for $n > n_0(\varepsilon)$ for every $\varepsilon > 0$. They also proved that these graphs have at least $n^2/36 - \varepsilon n^2$ C_{2k+1}-edges for $n > n_k(\varepsilon)$ for every $\varepsilon > 0$ and $k \geqslant 3$. In [5] we were able to prove the same results only for graphs with $\lfloor n^2/4 \rfloor + \varepsilon n^2$ edges (for $n > n_0(k,\varepsilon)$, $k \geqslant 2$). Let us close with a slightly corrected version of Erdős's conjecture.

Conjecture 5.1. *Suppose that G is an n-vertex graph with e edges, such that $e > n^2/4$ and it has the minimum number of C_{2k+1}-edges, $k \geqslant 3$, $n > n_k$. Then G is connected and has two blocks, of which one is a complete bipartite graph and the other is almost complete.*

Acknowledgement

The authors are very grateful to the referee for helpful comments.

References

[1] Chung, F. and Graham, R. (1998) *Erdős on Graphs: His Legacy of Unsolved Problems*, A. K. Peters.

[2] Erdős, P. (1997) Some recent problems and results in graph theory. *Discrete Math.* **164** 81–85.

[3] Erdős, P., Faudree, R. J. and Rousseau, C. C. (1992) Extremal problems involving vertices and edges on odd cycles. *Discrete Math.* **101** 23–31.

[4] Füredi, Z. and Maleki, Z. (2014) The minimum number of triangular edges and a symmetrization for multiple graphs. arXiv:1411.0771

[5] Füredi, Z. and Maleki, Z. A proof and a counterexample for a conjecture of Erdős concerning the minimum number of edges in odd cycles. Manuscript.

[6] Gruslys, V. and Letzter, S. (2016) Minimising the number of triangular edges. arXiv:1605.00528

[7] Grzesik, A., Hu, P. and Volec, J. (2016) Minimum number of edges that occur in odd cycles. Manuscript.

[8] Motzkin, T. S. and Straus, E. G. (1965) Maxima for graphs and a new proof of a theorem of Turán. *Canad. J. Math.* **17** 533–540.

[9] Turán, P. (1941) On an extremal problem in graph theory (in Hungarian). *Matematikai és Fizikai Lapok* **48** 436–452.

[10] Zykov, A. A. (1949) On some properties of linear complexes (in Russian). *Mat. Sbornik* (NS) **24** 163–188.