ON A PROBLEM OF AHLSWEDE AND KATONA

STEPHAN WAGNER AND HUA WANG

Abstract. Let $p(G)$ denote the number of pairs of adjacent edges in a graph G. Ahlswede and Katona considered the problem of maximizing $p(G)$ over all simple graphs with a given number n of vertices and a given number N of edges. They showed that $p(G)$ is either maximized by a quasi-complete graph or by a quasi-star. They also studied the range of N (depending on n) for which the quasi-complete graph is superior to the quasi-star (and vice versa) and formulated two questions on distributions in this context. This paper is devoted to the solution of these problems.

1. Introduction

For a simple undirected graph G, we consider the number

$$p(G) = \sum_{v \in V(G)} \binom{\deg v}{2} \tag{1}$$

of pairs of adjacent edges. A question stemming originally from information theory that was considered by Ahlswede and Katona in their paper [1] is to maximize $p(G)$, given the number n of vertices and the number N of edges of G. The problem had also been dealt with earlier (and solved for "nice" values of N) in an equivalent formulation for matrices by Katz [2].

In order to formulate the main result given by Ahlswede and Katona, we need the notions of quasi-complete graphs and quasi-stars. Let n and $0 \leq N \leq \binom{n}{2}$ be nonnegative integers. Then, the quasi-complete graph C_n^N with n vertices and N edges is constructed as follows:

- Write $N = \binom{a}{2} + b$, where $0 \leq b < a$.
- Let a vertices form a complete graph.
- Add another vertex and attach it to b of the previous vertices.
- Finally, add $n - a - 1$ isolated vertices.

Furthermore, let the quasi-star S_n^N be the complement of $C_n^{\binom{n}{2}-N}$. Then, the following theorem holds:

Theorem 1 (Ahlswede and Katona [1]). *Among all graphs G with a given number n of vertices and a given number N of edges, $p(G)$ is either maximized by the quasi-complete graph C_n^N or the quasi-star S_n^N.*

The obvious problem that arises is to characterize the values of N for which C_n^N resp. S_n^N is optimal. Indeed, Ahlswede and Katona showed more precisely that the following holds:

Theorem 2 (Ahlswede and Katona [1]). *There is a nonnegative integer R (depending on n) such that*

$$\begin{aligned}
p(C_n^N) &\leq p(S_n^N) \quad \textit{for } 0 \leq N \leq \frac{1}{2}\binom{n}{2} - R, \\
p(C_n^N) &\geq p(S_n^N) \quad \textit{for } \frac{1}{2}\binom{n}{2} - R < N \leq \frac{1}{2}\binom{n}{2}, \\
p(C_n^N) &\leq p(S_n^N) \quad \textit{for } \frac{1}{2}\binom{n}{2} \leq N < \frac{1}{2}\binom{n}{2} + R, \\
p(C_n^N) &\geq p(S_n^N) \quad \textit{for } \frac{1}{2}\binom{n}{2} + R \leq N \leq \binom{n}{2}.
\end{aligned}$$

For the sake of completeness, we will prove the existence of R *en route* to our main result. As Ahlswede and Katona noticed, R shows some rather unexpected behavior, and so they posed the following questions:

- What is the relative density of the numbers n for which $R = 0$?
- How is the normalized value $\frac{R}{n}$ distributed?

The aim of this paper is to answer these two questions. It will turn out that the distribution is as follows:

Theorem 3. *The relative density of the numbers n for which $R = 0$ is $2 - \sqrt{2}$. The distribution function of $\frac{R}{n}$ is given by*

$$D(x) = \begin{cases} 0 & x < 0, \\ \frac{2(\sqrt{2}-1-x)}{\sqrt{2}-4x} & 0 \le x \le 1 - \frac{1}{\sqrt{2}}, \\ 1 & x \ge 1 - \frac{1}{\sqrt{2}}. \end{cases}$$

Figure 1 shows a graph of the distribution function.

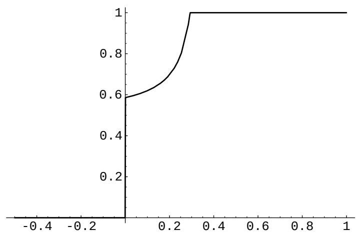

FIGURE 1. The distribution of $\frac{R}{n}$

2. PRELIMINARIES

From the definition of the quasi-complete graph and the quasi-star and equation (1), we immediately get the following formulas:

Lemma 4. *If $N = \binom{a}{2} + b$ with $0 \le b \le a$, then*

$$p(C_n^N) = b\binom{a}{2} + (a-b)\binom{a-1}{2} + \binom{b}{2} = \frac{a(a-1)(a-2)}{2} + ab + \frac{b(b-3)}{2}$$

and

$$p(S_n^N) = 2(n-2)N - \frac{n(n-1)(n-2)}{2} + p(C_n^{\binom{n}{2}-N}).$$

Note that the latter statement actually holds for any graph and its complement. This follows easily from the identity

$$\binom{n-1-d}{2} = \binom{d}{2} + \binom{n-1}{2} - (n-2)d$$

and equation (1) together with the well-known fact that

$$\sum_{v \in V(G)} \deg v = 2|E(G)|.$$

From now on, we are going to use the abbreviation $Q(N) = p(C_n^N)$ (note that this does not depend on n). Then the second formula also shows that $p(C_n^N) \ge p(S_n^N)$ is equivalent to

$$Q\left(\binom{n}{2} - N\right) - Q(N) \le \frac{(n-2)(n(n-1)-4N)}{2}. \tag{2}$$

If we set $N = \frac{1}{2}\binom{n}{2} - r$, this simplifies to

$$Q\left(\frac{1}{2}\binom{n}{2} + r\right) - Q\left(\frac{1}{2}\binom{n}{2} - r\right) \le 2r(n-2). \tag{3}$$

First we need an estimate for $Q(N)$:

Lemma 5. *For all nonnegative integers N, we have*

$$\frac{N}{2}\left(\sqrt{8N+1}-3\right) \geq Q(N) \geq \frac{N}{2}\left(\sqrt{8N+1}-\frac{7}{2}\right).$$

Proof. We perform the substitution $\sqrt{8N+1}=2a+c$ (with a as in Lemma 4), so that $-1 \leq c \leq 1$. Then the left hand side is equivalent to

$$\frac{1-c^2}{128}\left(16a^2-24a-1+8c(a-1)+c^2\right) \geq 0.$$

The first factor is obviously nonnegative, and the second factor can be estimated by

$$16a^2-24a-1+8c(a-1)+c^2 \geq 16a^2-24a-1,$$

which is positive for $a>1$. Similarly, the right hand side is equivalent to

$$\frac{1}{128}\left(24a-3+8c(a+1)+c^2\left(16a^2-24a+2+8c(a-1)+c^2\right)\right) \geq 0,$$

and

$$24a-3+8c(a+1)+c^2\left(16a^2-24a+2+8c(a-1)+c^2\right) \geq 16a-11+c^2(16a^2-24a+2),$$

which is also positive for $a>1$. The case $a=1$ is trivial. ∎

Lemma 6. *If $n \geq 50$ and (3) is satisfied for some $r \geq 0$, then either $r \geq \frac{1}{2}\binom{n}{2}-3$ or $r \leq \frac{2n}{3}$.*

Proof. Making use of the previous lemma, we obtain

$$\frac{n(n-1)+4r}{8}\left(\sqrt{2n(n-1)+8r+1}-\frac{7}{2}\right)-\frac{n(n-1)-4r}{8}\left(\sqrt{2n(n-1)-8r+1}-3\right)$$
$$\leq Q\left(\frac{1}{2}\binom{n}{2}+r\right)-Q\left(\frac{1}{2}\binom{n}{2}-r\right) \leq 2r(n-2)$$

or

$$(n(n-1)+4r)\sqrt{2n(n-1)+8r+1}-(n(n-1)-4r)\sqrt{2n(n-1)-8r+1}$$
$$\leq \frac{n(n-1)}{2}+2r(8n-3).$$

Note that the function $f(x)=x\sqrt{2x+1}-\sqrt{2x^3}$ is increasing, so that the left hand side can be estimated below by

$$\sqrt{2}(n(n-1)+4r)^{3/2}-\sqrt{2}(n(n-1)-4r)^{3/2}.$$

Now, we are left with

$$\sqrt{2}(n(n-1)+4r)^{3/2}-\sqrt{2}(n(n-1)-4r)^{3/2} \leq \frac{n(n-1)}{2}+2r(8n-3).$$

Substituting $r=\frac{n(n-1)}{4}s$ (so that $0 \leq s \leq 1$), we find that this is equivalent to

$$\sqrt{n(n-1)}\left((2+2s)^{3/2}-(2-2s)^{3/2}\right) \leq 1+s(8n-3).$$

Let $f(s)$ be the function $\left((2+2s)^{3/2}-(2-2s)^{3/2}\right)$. It is easily seen to be concave, and so we can estimate it by its secants. For $s \leq \frac{1}{2}$, we have

$$f(s) \geq 2s(3\sqrt{3}-1),$$

and together with $\sqrt{n(n-1)} \geq n-\frac{101}{200}$, which holds for $n \geq 50$, we get

$$\left(n-\frac{101}{200}\right)\cdot 2s(3\sqrt{3}-1) \leq 1+s(8n-3).$$

Solving for s, we have

$$s \leq \frac{100}{200(3\sqrt{3}-5)n-(303\sqrt{3}-401)},$$

which is easily seen to be $\leq \frac{8}{3(n-1)}$ for $n \geq 50$, implying $r \leq \frac{2n}{3}$. On the other hand, if $s \geq \frac{1}{2}$, we estimate $f(s)$ by

$$f(s) \geq 8+6(3-\sqrt{3})(s-1)$$

and get

$$\left(n - \frac{101}{200}\right) \cdot \left(8 + 6(3-\sqrt{3})(s-1)\right) \leq 1 + s(8n-3).$$

Solving for s again, we obtain

$$s \geq 1 - \frac{204}{200(3\sqrt{3}-5)n + (609 - 303\sqrt{3})} \geq 1 - \frac{6}{n},$$

which implies that $\binom{n}{2} - N = \frac{1}{2}\binom{n}{2} + r \geq \frac{(n-1)(n-3)}{2} \geq \binom{n-2}{2}$. But this allows us to refine our estimate:

- If $\binom{n}{2} - N < \binom{n-1}{2}$ (which implies $n \leq N \leq 2n-3$), we can use Lemma 4 with $a = n-2$ and $b = 2n-3-N$, and (2) becomes, upon simplification,
$$\left(n - \frac{N+4}{2}\right)^2 + \frac{N^2+2N-4}{4} \leq Q(N).$$
Using Lemma 5, we get
$$\frac{N^2+2N-4}{4} \leq \frac{N}{2}\left(\sqrt{8N+1} - 3\right),$$
which is easily shown to be wrong for $N \geq 13$, contradicting $N \geq n \geq 50$.
- If $\binom{n}{2} - N \geq \binom{n-1}{2}$, we use Lemma 4 with $a = n-1$ and $b = n-1-N$ to find
$$\frac{N(N-1)}{2} \leq Q(N) \leq \frac{N}{2}\left(\sqrt{8N+1} - 3\right),$$
which simplifies further to $N = 0$ or
$$N + 2 \leq \sqrt{8N+1}$$
and finally to $N \leq 3$ (indeed, if $N \leq 3$, we have equality in (2)).

∎

This basically leaves us with the case $r \leq \frac{2n}{3}$. In the following, we write $\frac{1}{2}\binom{n}{2} = \binom{e}{2} + f$, where $0 \leq f < e$. The following estimate will be important for the proof of our main theorem:

Lemma 7. *If $n \geq 50$, then $e > \frac{2n}{3} + 1$.*

Proof. Assume the contrary. Then

$$\frac{1}{2}\binom{n}{2} < \binom{e+1}{2} \leq \frac{\left(\frac{2n}{3}+1\right)\left(\frac{2n}{3}+2\right)}{2} = \frac{(2n+3)(n+3)}{9},$$

which is equivalent to $n^2 - 45n - 36 < 0$. This inequality is false for $n \geq 50$. ∎

3. Proof of the main result

Now, we are ready to formulate and prove our main theorem:

Theorem 8. *Let n be a positive integer and $0 \leq N \leq \frac{1}{2}\binom{n}{2}$. The inequality (2) holds with equality if any of the following is true:*

- $N \leq 3$,
- $N = \frac{1}{2}\binom{n}{2}$,
- $N = \binom{e}{2}$ *and* $n \in \{2f, 2f+1\}$,
- $N = \frac{1}{2}\binom{n}{2} - \frac{f(e-2)}{3e+2f-2n-1}$ *and* $2e + 2f + 1 < 2n$,
- $\binom{e}{2} \leq N < \frac{1}{2}\binom{n}{2}$ *and* $2e + 2f + 1 = 2n$.

Furthermore, (2) holds with strict inequality for

$$\frac{1}{2}\binom{n}{2} - \frac{f(e-2)}{3e+2f-2n-1} < N < \frac{1}{2}\binom{n}{2}$$

if $2e + 2f + 1 < 2n$. For all other values of N, the inequality is not satisfied.

Proof. Let first n be ≥ 50, and note that equality indeed holds for $N \leq 3$ and $N = \frac{1}{2}\binom{n}{2}$, which is easy to check. Otherwise, (2) can only be satisfied if or $N \geq \frac{1}{2}\binom{n}{2} - \frac{2n}{3}$ by Lemma 6. We write $N = \frac{1}{2}\binom{n}{2} - r$ (so that $0 \leq r \leq \frac{2n}{3}$ now) again and work with (3) rather than (2). Furthermore, we have to distinguish two cases: $2f \geq e$ and $2f < e$.

Throughout the rest of the proof we will make frequent use of the simple fact that a linear inequality holds on an entire interval if it does at both endpoints.

Case 1: $2f \geq e$. In this case, there are the following possibilities for r:

- $0 \leq r < e - f$, which implies that
$$\binom{e}{2} \leq \binom{e}{2} + f - (e-f) < \frac{1}{2}\binom{n}{2} - r \leq \frac{1}{2}\binom{n}{2} + r < \binom{e}{2} + f + e - f = \binom{e+1}{2}.$$
Now, application of Lemma 4 to (3) yields $(2e+2f-3)r \leq 2r(n-2)$, which holds if either $r = 0$ (one of our cases of equality) or $2e + 2f + 1 \leq 2n$. However, the latter is impossible since
$$2e + 2f + 1 \geq 3e + 1 > 2n + 4 > 2n$$
by our assumptions and Lemma 7.
- $e - f \leq r \leq f$, from which it follows that
$$\binom{e}{2} \leq \frac{1}{2}\binom{n}{2} - r < \binom{e+1}{2} \leq \frac{1}{2}\binom{n}{2} + r < \binom{e+2}{2}.$$
We apply Lemma 4 again and find that (3) is equivalent to
$$(e-f)(e-1) \leq r(2n - e - 2f - 2).$$
For $r = e - f$, this is equivalent to $2e + 2f + 1 \leq 2n$, which has already been excluded above. For $r = f$, on the other hand, the inequality reduces to
$$e^2 - e \leq f(2n - 2f - 3),$$
and since $\binom{e}{2} = \frac{1}{2}\binom{n}{2} - f$ or $e^2 - e = \binom{n}{2} - 2f$, this is in turn equivalent to
$$(2f - n)(2f - n + 1) \leq 0,$$
which is only satisfied for $n = 2f$ or $n = 2f + 1$, with equality in both cases. In any case, the inequality does not hold on the entire interval $e - f \leq r < f$.
- Finally, we are left with $f < r < 2e+1-f$ (the right hand side being true since $2e+1-f > e > \frac{2n}{3}$ by Lemma 7), when
$$\binom{e-1}{2} \leq \frac{1}{2}\binom{n}{2} - r < \binom{e}{2} < \binom{e+1}{2} \leq \frac{1}{2}\binom{n}{2} + r < \binom{e+2}{2}.$$
In this case, we obtain that (3) is equivalent to
$$e^2 - e - (2e-3)f \leq 2r(n - e - f).$$
Now, $r = f$ leads to the same inequality as before, which holds only (with equality) in two cases. We already know that $e + f > n$, which implies that the right hand side decreases as r increases, and so the inequality cannot hold for any other value of r.

Summarizing, we found two cases of equality ($r = 0$ and $r = f$ if $n \in \{2f, 2f+1\}$). In all other cases, the opposite of (3) holds.

Case 2: $2f < e$. Again, there are three possibilities for r:

- For $0 \leq r \leq f$,
$$\binom{e}{2} \leq \frac{1}{2}\binom{n}{2} - r \leq \frac{1}{2}\binom{n}{2} + r < \binom{e+1}{2}.$$
As before, we obtain the inequality $(2e + 2f - 3)r \leq 2r(n-2)$ from (3), which holds if either $r = 0$ or $2e + 2f + 1 \leq 2n$ (in the latter case for all $0 < r \leq f$, with equality if $2e + 2f + 1 = 2n$ and with strict inequality otherwise).

- For $f < r < e - f$,
$$\binom{e-1}{2} \leq \frac{1}{2}\binom{n}{2} - r < \binom{e}{2} \leq \frac{1}{2}\binom{n}{2} + r < \binom{e+1}{2},$$
and application of Lemma 4 to (3) yields
$$(2n - 3e - 2f + 1)r + (e-2)f \geq 0.$$
Now, $r = f$ leads to the inequality $2e + 2f + 1 \leq 2n$ again, while $r = e - f$ gives us
$$2f^2 - f(2n - 2e + 3) - 3e^2 + 2en + e \geq 0.$$
The left hand side is convex in f, and it is equal to
$$-e(3e - 2n - 1) < 0$$
for $f = 0$ and to
$$-\frac{e(3e-2n+1)}{2} < 0$$
for $f = e/2$. Both estimates are consequences of Lemma 7. Therefore, the inequality (3) cannot be satisfied for $r = e - f$, which shows that (3) either holds for no r within our interval (if $2e + 2f + 1 > 2n$) or for all r which satisfy the inequality
$$r \leq \frac{f(e-2)}{3e + 2f - 2n - 1},$$
where equality also implies equality in (3).
- Finally, for $e - f \leq r \leq e + f$ (the right hand side being true since $e + f \geq e > \frac{2n}{3}$ by Lemma 7),
$$\binom{e-1}{2} \leq \frac{1}{2}\binom{n}{2} - r < \binom{e}{2} < \binom{e+1}{2} \leq \frac{1}{2}\binom{n}{2} + r < \binom{e+2}{2},$$
and we obtain from Lemma 4 that (3) is equivalent to
$$e^2 - e - (2e-3)f \leq 2r(n - e - f).$$
$r = e - f$ leads to the same inequality as before, and for $r = e + f > \frac{2n}{3}$, Lemma 6 shows immediately that the inequality (3) cannot be satisfied. Therefore, (3) does not hold on the entire interval.

Summarizing this second case, we found that the inequality is only satisfied for $r = 0$ if $2f + 2e + 1 > 2n$, and that it is satisfied with equality for all $r \in [0, f]$ if $2f + 2e + 1 = 2n$. Finally, in the case $2f + 2e + 1 < 2n$, (3) holds with strict inequality for $0 < r < \frac{f(e-2)}{3e+2f-2n-1}$ and with equality for $r = \frac{f(e-2)}{3e+2f-2n-1}$ (if this is an integer).

Finally, let us just mention that the case $n < 50$ can be checked directly by means of a computer. ■

This allows us to formulate Theorem 2 in a more precise way:

Theorem 9. *There is a nonnegative integer R such that*
$$\begin{aligned}
p(C_n^N) &\leq p(S_n^N) \quad \textit{for } 0 \leq N \leq \frac{1}{2}\binom{n}{2} - R,\\
p(C_n^N) &\geq p(S_n^N) \quad \textit{for } \frac{1}{2}\binom{n}{2} - R < N \leq \frac{1}{2}\binom{n}{2},\\
p(C_n^N) &\leq p(S_n^N) \quad \textit{for } \frac{1}{2}\binom{n}{2} \leq N < \frac{1}{2}\binom{n}{2} + R,\\
p(C_n^N) &\geq p(S_n^N) \quad \textit{for } \frac{1}{2}\binom{n}{2} + R \leq N \leq \binom{n}{2}.
\end{aligned}$$

Specifically, we can take
$$R = \begin{cases} 0 & 2e + 2f + 1 \geq 2n, \\ \left\lceil \frac{f(e-2)}{3e+2f-2n-1} \right\rceil & 2e + 2f + 1 < 2n. \end{cases}$$

Note also that R is unique in almost all cases: the only exception is the case $2e+2f+1=2n$, when R could be chosen arbitrarily from the interval $[0,f]$. However, substituting $\frac{1}{2}\binom{n}{2}-\binom{e}{2}$ for f, we find that this equation is equivalent to

$$(2n-5)^2 = 2(2e-3)^2 - 1.$$

The solution of this Pell-type equation is given by

$$n = \frac{1}{4}\left((1+\sqrt{2})^{2k-1} + (1-\sqrt{2})^{2k-1}\right) + \frac{5}{2}$$

for some nonnegative integer k, and the set of these numbers has density zero.

Now, we are also finally able to solve the problem posed by Ahlwede and Katona:

Proof of Theorem 3. First, we have to determine the probability that $2e+2f+1 \geq 2n$. For this purpose, let us first note that our definition of e and f implies that

$$e = \left\lfloor \frac{1}{2}\left(1+\sqrt{2n^2-2n+1}\right)\right\rfloor = \frac{1}{2}\left(1+\sqrt{2n^2-2n+1}\right) - \alpha,$$

where $\alpha = \left\{\frac{1}{2}\left(1+\sqrt{2n^2-2n+1}\right)\right\} \in [0,1)$ ($\{x\}$ denotes the fractional part of x). Note now that

$$\frac{1}{2}\left(1+\sqrt{2n^2-2n+1}\right) = \frac{n}{\sqrt{2}} + \frac{2-\sqrt{2}}{4} + O(n^{-1}),$$

and since $\frac{1}{\sqrt{2}}$ is irrational, this implies that α is equidistributed in the interval $[0,1)$ (see for instance [3]).

Furthermore, since $f = \frac{1}{2}\binom{n}{2} - \binom{e}{2}$, the inequality $2e+2f+1 \geq 2n$ is equivalent to

$$\frac{n^2-5n}{2} \geq e^2 - 3e - 1,$$

which is in turn equivalent to $e \leq \frac{1}{2}\left(3+\sqrt{2n^2-10n+13}\right)$ or

$$\frac{1}{2}\left(3+\sqrt{2n^2-10n+13}\right) \geq \frac{1}{2}\left(1+\sqrt{2n^2-2n+1}\right) - \alpha.$$

This can be rewritten as

$$\alpha \geq \frac{1}{2}\left(1+\sqrt{2n^2-2n+1}\right) - \frac{1}{2}\left(3+\sqrt{2n^2-10n+13}\right) = \sqrt{2} - 1 - \frac{1}{4\sqrt{2}n^2} + O(n^{-3}),$$

which means that $2e+2f+1 \geq 2n$ is satisfied for an asymptotic fraction of $1-(\sqrt{2}-1) = 2-\sqrt{2}$ of the values of n. This proves the first part of the statement. For the second part, note that

$$\begin{aligned}\frac{R}{n} &= \frac{(e-2)f}{n(2f+3e-2n-1)} + O(n^{-1}) = \frac{(e-2)(n^2-n-2e^2+2e)}{2n(n^2-5n-2e^2+8e-2)} + O(n^{-1}) \\ &= \frac{\alpha(\sqrt{2n^2-2n+1}-\alpha)(\sqrt{2n^2-2n+1}-2\alpha-3)}{2n((2\alpha+3)\sqrt{2n^2-2n+1}-4n-2\alpha^2-6\alpha+1)} + O(n^{-1}) = \frac{\alpha}{3\sqrt{2}-4+2\sqrt{2}\alpha} + O(n^{-1}),\end{aligned}$$

from which we deduce that

$$\frac{R}{n} \sim \frac{\alpha}{3\sqrt{2}-4+2\sqrt{2}\alpha}$$

follows the distribution function given in the theorem, which is calculated by solving the equation

$$\frac{\alpha}{3\sqrt{2}-4+2\sqrt{2}\alpha} = x$$

for α. ■

References

[1] R. Ahlswede and G. O. H. Katona. Graphs with maximal number of adjacent pairs of edges. *Acta Math. Acad. Sci. Hungar.*, 32(1-2):97–120, 1978.

[2] M. Katz. Rearrangements of (0, −1) matrices. *Israel J. Math.*, 9:53–72, 1971.

[3] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences.* Wiley-Interscience [John Wiley & Sons], New York, 1974. Pure and Applied Mathematics.

Stephan Wagner, Department of Mathematical Sciences, Stellenbosch University, 7602 Stellenbosch, South Africa

E-mail address: swagner@sun.ac.za

Hua Wang, Department of Mathematics, University of Florida, Gainesville, FL, 32611, United States of America

E-mail address: hua@math.ufl.edu