Graph Guessing Games and non-Shannon Information Inequalities

Rahil Baber
and Anh N. Dang
School of Electronic Engineering
and Computer Science
Queen Mary, University of London
London, E1 4NS, U.K.
Email: rahilbaber@hotmail.com
and nad30@eecs.qmul.ac.uk

Demetres Christofides
Department of Computing
and Mathematics
UCLan Cyprus, 7080 Pyla, Cyprus
Email: d.christofides@uclancyprus.ac.cy

Søren Riis
and Emil R. Vaughan
School of Electronic Engineering
and Computer Science
Queen Mary, University of London
London, E1 4NS, U.K.
Email: smriis@dcs.qmul.ac.uk
and e.vaughan@qmul.ac.uk

***Abstract*—Guessing games for directed graphs were introduced by Riis [6] for studying multiple-unicast network coding problems. It can be shown that protocols for a multiple-unicast network can be directly converted into a strategy for a guessing game. The performance of the optimal strategy for a graph is measured by the guessing number, and this number can be bounded from above using information inequalities. In [2], a strategy based on the fractional clique cover was introduced and conjectured to be optimal for undirected graphs. In this paper, we disprove this conjecture. We also provide an example of an undirected graph in which non-Shannon inequalities provide a better bound on the guessing number than Shannon inequalities. Finally, we construct a counterexample to a conjecture we raised during our work which we referred to as *the Superman conjecture*.**

I. Introduction

Guessing games [6], [7] emerge naturally from studying network coding problems [1] where the network is a multiple-unicast network, i.e. where each sender has precisely one corresponding receiver who wishes to obtain the sender's message, and an additional constrain that only one message can be sent through each channel at a time. A multiple-unicast can be represented by an acyclic directed graph with n inputs/outputs and m intermediate nodes. By merging the vertices which represent the senders with their corresponding receiver vertices in the acyclic directed graph representing the multiple-unicast network, we can create an auxiliary directed graph which has the nice property that there is no longer any distinction between router, sender, or receiver vertices. Due to the way guessing games are defined, coding functions on the original network can be translated into strategies for the guessing game on the auxiliary graph and vice versa. The performance of the optimal strategy for a guessing game is measured by the guessing number which we will define precisely in Section II.

One of the first applications of guessing games was to disprove two conjectures raised by Valiant in circuit complexity in which he asked about the optimal Boolean circuit for a Boolean function [6].

Our main result is a counterexample to a conjecture of Christofides and Markström given in [2] which states that the optimal strategy for the guessing game of an undirected graph (a directed graph where every edge is bidirected) is based on the fractional clique cover of the graph. Additionally we will show that the guessing number for undirected graphs cannot be determined by considering only the Shannon information inequalities. We will also make and investigate the Superman conjecture which suggests the (asymptotic) guessing number of an undirected graph does not increase when a directed edge is added. Finally we will provide a possible example of a directed graph whose guessing number changes when its edges are reversed.

The outline of our paper is as follows. In Section II we introduce the formal language of guessing games and the asymptotic behaviour of guessing numbers. In Section III we introduce a method for calculating upper bounds of guessing numbers by making use of entropic arguments. Our main results appear in Section IV. Finally, we conclude with some open problems in Section V.

II. Guessing games and its lower bound via fractional clique cover strategy

A *directed graph*, or *digraph* for short, is a pair $G = (V(G), E(G))$, where $V(G)$ is the set of *vertices* of G and $E(G)$ is a set of pairs of vertices of G called the *directed edges* of G. Given a directed edge $e = (u, v)$, which we also denote by uv, we call u the *tail* and v the *head* of e and say that e goes from u to v. For a vertex $v \in V(G)$, we denote its *in-neighbourhood* as $\Gamma^-(v)$ and its *out-neighbourhood* as $\Gamma^+(v)$.

For the purposes of the guessing games we will assume throughout that our graphs are simple. I.e. they contain no edges of the form uu for $u \in V(G)$, and no two edges have the same head and tail.

In this paper our main results will primarily be on *undirected graphs* which are naturally treated as a special type of digraph G where $uv \in E(G)$ if and only if $vu \in E(G)$. We call the pair of directed edges uv and vu, the *undirected edge* uv. A major role in our guessing strategies will be played by *cliques* i.e. subgraphs in which every pair of vertices are joined by an undirected edge.

Given a digraph G and an integer $t \geq 1$, the *t-uniform blowup of* G which we will write as $G(t)$ is a digraph formed by replacing each vertex v in G with a class of t vertices $v_1, \dots, v_t$ with $u_i v_j \in E(G(t))$ if and only if $uv \in E(G)$.

A *guessing game* (G, s) is a game played on a digraph G and the alphabet $A_s = \{0, 1, \dots, s-1\}$. There are $|V(G)|$ players, each corresponding to one of the vertices of the digraph. Throughout the article we will be freely speaking about the player v instead of the player corresponding with the vertex $v \in V(G)$. The players know the digraph G, the natural number s, and are told to which of the vertices they correspond to. They may discuss and agree upon a strategy using this information before the game begins, but no communication between the players is allowed after the game starts.

Once the game begins each player $v \in V(G)$ is assigned a value a_v from A_s uniformly and independently at random. Each player v does not have access to its own values but can see the values assigned to the other players who are represented as in-neighbourhoods $\Gamma^-(v)$. Using just this information each player must guess their own value. If all players guess correctly they will all win, but if just one player guesses incorrectly they will all lose. The objective of the players is to maximise their probability of winning.

Given a guessing game (G, s), for $v \in V(G)$ a *strategy for player* v is formally a function $f_v : A_s^{|\Gamma^-(v)|} \to A_s$. A *strategy* $\mathcal{F}$ for a guessing game is a sequence of such functions $(f_v)_{v \in V(G)}$ where f_v is a strategy for player v. The *guessing number* $\text{gn}(G, s)$ is defined to be:

$$\text{gn}(G, s) = |V(G)| + \log_s \left(\max_{\mathcal{F}} \mathbf{P}[\text{Win}(G, s, \mathcal{F})] \right). \quad (1)$$

where $\mathbf{P}[\text{Win}(G, s, \mathcal{F})]$ is the probability of the event $\text{Win}(G, s, \mathcal{F})$ that all the players guess correctly when playing (G, s) with strategy $\mathcal{F}$. Although this looks like a cumbersome property to work with we can think of it as a measure of how much better the optimal strategy is over the strategy of just making random guesses, as

$$\max_{\mathcal{F}} \mathbf{P}[\text{Win}(G, s, \mathcal{F})] = \frac{s^{\text{gn}(G,s)}}{s^{|V(G)|}}.$$

We note that strategies $\mathcal{F}$ in our definition are *pure strategies* i.e. there is no randomness involved in the guess each player makes given the values it sees. The alternative to a pure strategy is a *mixed strategy* (also called *randomized strategy*) which is a guessing strategy in which each player randomly chooses a strategy to play from a set of pure strategies. The winning probability of the mixed strategy is the average of the winning probabilities of the pure strategies weighted according to the probabilities that they are chosen. This however is at most the maximum of the winning probabilities of the pure strategy and so we gain no advantage by playing a mixed strategy. As such, throughout this paper we will only ever consider pure strategies.

In general $\text{gn}(G, s)$ will depend on s and it is often extremely difficult to determine the guessing number exactly. Consequently we will instead concentrate our efforts on evaluating the *asymptotic guessing number* $\text{gn}(G)$ which we define to be the limit of $\text{gn}(G, s)$ as s tends to infinity. The limit

$$\text{gn}(G) = \lim_{s \to \infty} \text{gn}(G, s)$$

does always exist [2].

From the definition of guessing game, we have the following observation:

Lemma II.1. *Given a digraph G, and integers $s, t \geq 1$, $\text{gn}(G(t), s) = t\,\text{gn}(G, s^t)$. where $G(t)$ is the t-uniformly blowups of G.*

Proof: From Equation (1) it is enough to show that the probability of winning on (G, s^t) is equal to the probability of winning on $(G(t), s)$.

We begin by showing that the optimal probability of winning on (G, s^t) is at least that of $(G(t), s)$. This follows simply from the fact that the members of the alphabet of size s^t, can be represented as t digit numbers in base s. Hence given a strategy on $(G(t), s)$ a corresponding strategy can be played on (G, s^t) by each player pretending to be t players. The value the player was originally assigned now becomes t values from A_s which are assigned to the new t fictitious players respectively. The strategy on $(G(t), s)$ can then be used and the guesses for each vertex class of t players can then be reconstructed into a guess for the original player in (G, s^t).

A similar argument can be used to show that the optimal probability of winning on (G, s^t) is at most that of $(G(t), s)$. We will show that for every strategy on (G, s^t) there is a corresponding strategy on $(G(t), s)$. Every vertex class of t players can simulate playing as one fictitious player by its members agreeing to use the same strategy. The t values assigned to the players in the vertex class can be combined to give an overall value for the vertex class. The strategy on (G, s^t) can then be played allowing the members of the vertex class to make a guess for the overall value assigned to the vertex class. This guess will be the same for each member as they all agreed to use the same strategy and have access to precisely the same information. Once the guess for the vertex class is made its value can be decomposed into t values from A_s which can be used as the individual guesses for each of its members. ■

If G is an undirected graph, Christofides and Markström [2] made the following conjecture:

Conjecture II.2. *Let $K(G)$ be the set of all cliques in undirected graph G, and let $K(G, v)$ be the set of all cliques containing vertex v. A (regular) fractional clique cover of G is a weighting $w : K(G) \to [0, 1]$ such that for all $v \in V(G)$*

$$\sum_{k \in K(G,v)} w(k) \geq 1.$$

The minimum value of $\sum_{k \in K(G)} w(k)$ over all choices of fractional clique covers w is known as the fractional clique cover number *which we will denote by $\kappa_f(G)$. The asymptotic* *guessing number of the undirected graph G is conjectured to be:*

$$\mathrm{gn}(G) = |V(G)| - \kappa_f(G).$$

A useful bound on $\kappa_f(G)$ which we will make use of later is given by the following lemma.

Lemma II.3. *For any undirected graph G*

$$\kappa_f(G) \geq \frac{|V(G)|}{\omega(G)},$$

where $\omega(G)$ is the number of vertices in a maximum clique in G.

Proof: Let w be a optimal fractional clique cover. Since $\sum_{k\in K(G,v)} w(k) = 1$ holds for all $v \in V(G)$, summing both sides over v gives us,

$$\sum_{k\in K(G)} w(k)|V(k)| = |V(G)|,$$

where $|V(k)|$ is the number of vertices in clique k. The result trivially follows from observing

$$\sum_{k\in K(G)} w(k)|V(k)| \leq \sum_{k\in K(G)} w(k)\omega(G) = \kappa_f(G)\omega(G).$$

■

III. Upper bounds using entropy

Recall that it is sufficient to only consider pure strategies on guessing games. Hence given a strategy $\mathcal{F}$ on a guessing game (G, s) we can explicitly determine $\mathcal{S}(\mathcal{F})$ the set of all assignment tuples $(a_v)_{v\in V(G)}$ that result in the players winning given they are playing strategy $\mathcal{F}$. In this context the players' objective is to choose a strategy that maximizes $|\mathcal{S}(\mathcal{F})|$. As it turns out,

$$\mathrm{gn}(G, s) = \max_{\mathcal{F}} \log_s |\mathcal{S}(\mathcal{F})|.$$

Consider the probability space on the set of all assignment tuples $A_s^{|V(G)|}$ with the members in $\mathcal{S}(\mathcal{F})$ occurring with uniform probability and all other assignments occurring with 0 probability. For each $v \in V(G)$ we define the discrete random variable X_v on this probability space to be the value assigned to vertex v. The *s-entropy* of a discrete random variable X with outcomes $x_1, x_2, \ldots, x_n$ is defined as

$$H_s(X) = -\sum_{i=1}^{n} \mathbf{P}[X = x_i] \log_s \mathbf{P}[X = x_i],$$

where we take $0 \log_s 0$ to be 0 for consistency. Note that traditionally entropy is defined using base 2 logarithms, however it will be more convenient for us to work with base s logarithms. We will usually write $H(X)$ instead of $H_s(x)$. We refer the reader to [4].

Given a set of random variables $Y_1, \ldots, Y_n$ with sets of outcomes $\mathrm{Im}(Y_1), \ldots, \mathrm{Im}(Y_n)$ respectively, the *joint entropy* $H(Y_1, \ldots, Y_n)$ is defined as

$$-\sum_{\mathbf{y}\in \mathrm{Im}(\mathbf{Y})} \mathbf{P}[\mathbf{Y} = \mathbf{y}] \log_s \mathbf{P}[\mathbf{Y} = \mathbf{y}].$$

where $\mathbf{Y}$ and $\mathbf{y}$ stand for $(Y_1, \ldots Y_n)$ and $(y_1, \ldots, y_n)$ respectively.

Given a set of random variables $Y = \{Y_1, \ldots, Y_n\}$ we will also use the notation $H(Y)$ to represent the joint entropy $H(Y_1, \ldots, Y_n)$. Furthermore for sets of random variables Y and Z we will use the notation $H(Y, Z)$ as shorthand for $H(Y \cup Z)$. For completeness we also define $H(\emptyset) = 0$.

Observe that under these definitions, the joint entropy of the set of variables $X_G = \{X_v : v \in V(G)\}$ is

$$\begin{aligned} H(X_G) &= -\sum_{(a_v)\in \mathcal{S}(\mathcal{F})} \frac{1}{|\mathcal{S}(\mathcal{F})|} \log_s \left(\frac{1}{|\mathcal{S}(\mathcal{F})|}\right) \\ &= \log_s |\mathcal{S}(\mathcal{F})| \end{aligned}$$

Therefore by upper bounding $H(X_G)$ for all choices of $\mathcal{F}$ we can bound $\mathrm{gn}(G, s)$.

We begin by stating some inequalities that most hold regardless of $\mathcal{F}$.

Theorem III.1. *Given $X, Y, Z \subset X_G$,*

1) $H(X) \geq 0$.
2) $H(X) \leq |X|$.
3) *Shannon's information inequality:*

$$H(X, Z) + H(Y, Z) - H(X, Y, Z) - H(Z) \geq 0.$$

4) $H(X, Y) = H(Y)$ *if there exists disjoint sets $A, B \subset V(G)$ such that $X = \{X_v : v \in A\}$, $Y = \{X_v : v \in B\}$, and $\Gamma^-(u) \subset B$ for all $u \in A$.*

Proof: See our technical report for details. ■

From Theorem III.1 we can form a linear program to upper bound $H(X_G)$. In particular the linear program consists of $2^{|V(G)|}$ variables corresponding to the values of $H(X)$ for each $X \subset X_G$. The variables are constrained by the linear inequalities given in Theorem III.1 and the objective is to maximize the value of the variable corresponding to $H(X_G)$. We call the result of the optimization *the Shannon bound*.

Note that the Shannon bound on $H(X_G)$ can be calculated without making any explicit use of $\mathcal{F}$ or s. Hence it is not only an upper bound on $\mathrm{gn}(G, s)$ but also on $\mathrm{gn}(G)$.

More recently information entropy inequalities that cannot be derived from linear combinations of Shannon's information inequality (Property 3 in Theorem III.1) have been discovered. The first such inequality was found by Zhang and Yeung [8].

The *Zhang-Yeung inequality* states that

$$\begin{aligned} &-2H(A) - 2H(B) - H(C) + 3H(A, B) + \\ &\quad 3H(A, C) + H(A, D) + 3H(B, C) + H(B, D) - \\ &\qquad H(C, D) - 4H(A, B, C) - H(A, B, D) \geq 0 \end{aligned}$$

for sets of random variables A, B, C, D. By setting $A = X \cup Z$, $B = Z$, $C = Y \cup Z$, $D = Z$, the Zhang-Yeung inequality reduces to Shannon's inequality. By replacing the Shannon inequality constraints with those given by the Zhang-Yeung inequality we can potentially get a better upper bound from the linear program. However, we pay for this potentially better bound by a significant increase in the running time of the linear program. We will call the bound on $\mathrm{gn}(G)$ obtained by use of the Zhang-Yeung inequality *the Zhang-Yeung bound*.

In fact there are known to be infinite families of non-Shannon inequalities even on 4 variables. We can not hope to add infinite constraints to the linear program so instead we will consider the 214 inequalities given by Dougherty, Freiling, and Zeger [5, Section VIII]. We will refer to the resulting bound as *the Dougherty-Freiling-Zeger bound*. It is perhaps worth mentioning that the 214 Dougherty-Freiling-Zeger inequalities imply the Zhang-Yeung inequality (simply sum inequalities 56 and 90) and therefore they also imply Shannon's inequality.

The final bound we will consider is *the Ingleton bound* which is obtained when we replace the Shannon inequality constraints with the *Ingleton inequality*

$$-H(A)-H(B)+H(A,B)+H(A,C)+ H(A,D)+H(B,C)+H(B,D)-H(C,D)- H(A,B,C)-H(A,B,D)\geq 0.$$

The Ingleton inequality provides the bound of the inner-cone of linearly representable entropy vectors [3]. It can be seen that if random variables $\{X_1,\dots X_n\}$ satisfy Ingleton inequlities, they also satisfy information inequalities.

If each player's strategy can be expressed as a linear combination of the values it sees then the Ingleton inequality will hold. Therefore the inequality holds for a strategy on (G,s^t) that can be represented as a linear strategy on $(G(t),s)$ (as described in the proof of Lemma II.1). As such the Ingleton bound gives us an upper bound when we restrict ourselves to strategies which are linear on the digits of the values. An important such strategy is the fractional clique cover strategy [2].

In searching for a counterexample to Conjecture II.2 we carried out an exhaustive search on all undirected graphs with at most 9 vertices. We compared the lower bound given by the fractional clique cover with the upper bound given by the Shannon bound and in all cases the two bounds matched.

IV. Main results

In this section we present our new results, which include a counterexample to conjectures II.2 and an example of graph G in which the Shannon-entropy of G and $\mathrm{Reverse}(G)$ are not identical. $\mathrm{Reverse}(G)$ is the digraph formed from G by reversing all the edges, $uv\in E(G)$ if and only if $vu\in E(\mathrm{Reverse}(G))$. Counterexamples were found by searching through all undirected graphs on 10 vertices or less.

For speed purposes, the search was done using floating point arithmetic and as such there may be counterexamples that were missed due to rounding errors. (Although this is highly unlikely, we do not claim it is impossible.) Despite this we feel it is still remarkable that of the roughly 12 million graphs that were checked we only found 2 graphs whose lower and upper bounds (given by the fractional clique cover, and Shannon bound respectively) did not match: the graph R given in Figure 1, and the graph R^- which is identical to R with the undirected edge between vertices 9 and 10 removed. The graph R is particularly extraordinary as we will see that with a few simple modifications we can create graphs which answer a few other open problems.

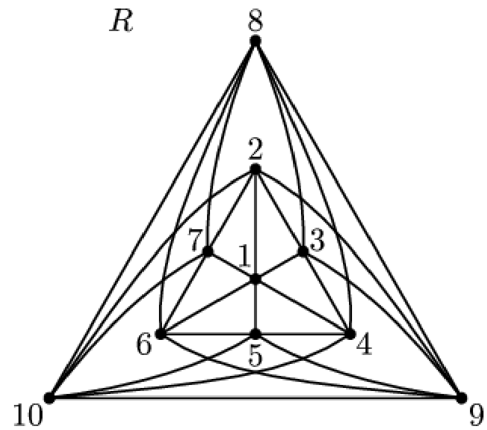

Fig. 1. The undirected graph R.

We begin our analysis of R and R^- by determining their fractional clique cover number.

Lemma IV.1. $\kappa_f(R)=\kappa_f(R^-)=10/3$.

Proof: By Lemma II.3 we know that $\kappa_f(R)$ and $\kappa_f(R^-)$ are bounded below by $10/3$. To show they can actually attain $10/3$ we need to construct explicit fractional clique covers whose weights add up to $10/3$.

For R^- we give a weight of $1/3$ to the cliques $\{1,2,3\}$, $\{1,4,5\}$, $\{1,6,7\}$, $\{2,3,9\}$, $\{2,7,10\}$, $\{3,8,9\}$, $\{4,5,10\}$, $\{4,8,10\}$, $\{5,6,9\}$, $\{6,7,8\}$, and a weight of 0 to all other cliques. Note that this is also an optimal fractional clique cover for R. ■

Theorem IV.2. *For R^-:*

1) *The Shannon bound is* $114/17=6.705\dots$
2) *The Zhang-Yeung bound is* $1212/181=6.696\dots$
3) *The Dougherty-Freiling-Zeger bound is* $59767/8929=6.693\dots$
4) *The Ingleton bound is* $20/3=6.666\dots$

From Lemma IV.1 and Theorem IV.2 we know that

$$20/3\leq \mathrm{gn}(R^-)\leq 59767/8929,$$

and although we could not determine the asymptotic guessing number exactly it does show that it does not equal the Shannon bound. Given that the Shannon bound is not sharp we might be tempted to conjecture that the asymptotic guessing number is the same as the Zhang-Yeung bound, but Theorem IV.2 also shows this to be false. Interestingly the Ingleton bound does match the lower bound, showing that if we restrict ourselves to only considering linear strategies on blowups we can do no better than the fractional clique cover strategy.

It remains an open question as to whether a non-linear strategy on R^- can do better than $20/3$ or whether by considering the right set of entropy inequalities we can push the upper bound down to $20/3$.

Proof of Theorem IV.2.: Calculating the upper bounds involves solving rather large linear programs. Hence the proofs are too long to reproduce here and it is unfeasible for them to be checked by humans. Data files verifying our claims can be provided upon request. We stress that although the results were verified using a computer that no floating point data types were used during the verification. Consequently no rounding errors could occur in the calculations making the results completely rigorous. ■

Although R is a counterexample to Conjecture II.2 its optimal strategy is somewhat complicated. So instead we will disprove the conjecture by showing that a related graph which we will call R_c is a counterexample. The undirected graph R_c is constructed from R by *cloning* 3 of its vertices. (Cloning 3 vertices is equivalent to creating a blowup of R with 2 vertices in 3 of the vertex classes and just 1 vertex in the other classes.) The vertices we clone are $8, 9, 10$, and we label the resulting new vertices $8', 9'$, and $10'$ respectively.

Theorem IV.3. *For R_c we have $\mathrm{gn}(R_c) = 9$ and $\kappa_f(R_c) = \frac{13}{3}$.*

Note that Theorem IV.3 disproves Conjecture II.2 as the conjecture suggests that

$$\mathrm{gn}(R_c) = 13 - \frac{13}{3} = \frac{26}{3} \neq 9.$$

Proof: Lemma II.3 tells us $\kappa_f(R_c) \geq 13/3$. It is also easy to show $\kappa_f(R_c) \leq 13/3$ as it trivially follows from extending the fractional clique cover given in the proof of Lemma IV.1 by giving a weight of 1 to the clique $\{8', 9', 10'\}$.

The Shannon bound of R_c is 9 proving $\mathrm{gn}(R_c) \leq 9$. We do not provide the details of the Shannon bound proof as it is too long to present here, however data files containing the proof are available upon request.

All that remains is to prove $\mathrm{gn}(R_c) \geq 9$. Note that this is enough to disprove Conjecture II.2 and there is no need to check verify that $\mathrm{gn}(R_c) \leq 9$. In [2], the authors proved that the asymptotic guessing number can be lower bounded by considering any strategy on any alphabet size. We will take our alphabet size s to be 3. Our strategy involves all players agreeing to play assuming the following four conditions hold on the assigned values

$$a_1 + a_2 + 2a_3 + a_4 + 2a_5 + a_6 + 2a_7 \equiv 0 \bmod 3, \quad (2)$$

$$a_2 + a_5 + a_8 + a_{8'} + a_9 + a_{10'} \equiv 0 \bmod 3, \quad (3)$$

$$a_3 + a_6 + a_8 + a_{9'} + a_{10} + a_{10'} \equiv 0 \bmod 3, \quad (4)$$

$$a_4 + a_7 + a_{8'} + a_9 + a_{9'} + a_{10} \equiv 0 \bmod 3. \quad (5)$$

Note that the terms in (2) consist of a_1, and values which player 1 can see. Hence (2) naturally gives us a strategy for player 1, i.e. that player 1 should guess $-a_2-2a_3-a_4-2a_5-a_6-2a_7 \bmod 3$. Similarly strategies for players $8, 8', 9, 9', 10$, and $10'$ can be achieved by rearranging conditions (4), (5), (3), (4), (5) and (3) respectively. A strategy for player 2 can be obtained by taking a linear combination of the conditions. In particular if we sum (4), (5), twice (2), and twice (3) we get

$$2a_1 + a_2 + 2a_3 + 2a_7 + 2a_{9'} + 2a_{10} \equiv 0 \bmod 3,$$

which consists of a_2 and values which player 2 can see, allowing us to construct a strategy for player 2. By taking appropriate linear combinations we can similarly produce strategies for players $3, 4, 5, 6$, and 7, though we leave this as an exercise to the reader.

The probability that all players guess correctly under this strategy is 3^{-4} i.e. the probability that (2), (3), (4), (5) all hold. (It is not difficult to check that the conditions are linearly independent.) Consequently

$$\mathrm{gn}(R_c) \geq |V(R_c)| + \log_3 \mathbf{P}[\mathrm{Win}(R_c, 3, \mathcal{F})] = 9$$

as desired. ■

Now that we have shown Conjecture II.2 is not true we turn our attention to other open questions. Due to the limited tools and methods currently available there are many seemingly trivial problems on guessing games which still remain unsolved. One such problem is the following.

Problem IV.4. *Does there exist an undirected graph whose asymptotic guessing number increases when a single directed edge is added?*

Adding a directed edge gives one of the players more information, which cannot lower the probability that the players win. However, surprisingly it seems extremely difficult to make use of the extra directed edge to increase the asymptotic guessing number. An exhaustive (but not completely rigorous) search on undirected graphs with 9 vertices or less did not yield any examples.

As such we significantly weaken the requirements in Problem IV.4 by introducing the concept of a *Superman vertex*. We define a Superman vertex to be one that all other vertices can see, i.e. given a digraph G, $u \in V(G)$ is a Superman vertex if $uv \in E(G)$ for all $v \in V(G) \setminus u$. We similarly define a *Luthor vertex* as one which sees all other vertices. To be precise u is a Luthor vertex if $vu \in E(G)$ for all $v \in V(G) \setminus u$.

Problem IV.5. *Does there exist an undirected graph whose asymptotic guessing number increases when directed edges are added to change one of the vertices into a Superman vertex (or a Luthor vertex)?*

To change one of the vertices into a Superman or Luthor vertex will often involve adding multiple directed edges, meaning the players will have a lot more information at their disposal when making their guesses. We again searched all undirected graphs on 9 vertices or less and remarkably still could not find any examples.

With the discovery of the graph R and in particular the graph R_c we can show the answer is yes to Problem IV.4 and consequently Problem IV.5. We define the undirected graph R_c^- to be the same as the graph R_c but with the undirected edge between vertices 3 and 8 removed. We also define the directed graph R_c^+ to be the same as R_c^- but with the addition of a single directed edge going from vertex 3 to 8.

Theorem IV.6. $\mathrm{gn}(R_c^-) = 53/6$ *and* $\mathrm{gn}(R_c^+) = 9$.

Proof: The Shannon bounds for R_c^- and R_c^+ are $53/6$ and 9 respectively (data files can be provided upon request).

We will prove $\mathrm{gn}(R_c^+) \geq 9$ by showing the strategy for $(R_c, 3)$ (see the proof of Theorem IV.3) is a valid strategy for $(R_c^+, 3)$. With the exception of player 3 all players in $(R_c^+, 3)$ have access to the same information they did in $(R_c, 3)$. Player 3 however now no longer has access to a_8. By studying the strategy player 3 uses in $(R_c, 3)$ we will see that this is of no consequence. Summing conditions (2), (3), (5), and twice (4), gives

$$a_1 + 2a_2 + a_3 + 2a_4 + 2a_{8'} + 2a_9 \equiv 0 \bmod 3,$$

hence player 3 guesses $-a_1 - 2a_2 - 2a_4 - 2a_{8'} - 2a_9 \bmod 3$ in $(R_c, 3)$. Since player 3 makes no use of a_8 this validates our claims.

We complete our proof by showing $\mathrm{gn}(R_c^-) \geq 53/6$. We know $\mathrm{gn}(R_c^-) \geq \mathrm{gn}(R_c^-, 3^6) = \mathrm{gn}(R_c^-(6), 3)/6$ so it is enough to show $\mathrm{gn}(R_c^-(6), 3) \geq 53$ by finding a strategy on $(R_c^-(6), 3)$ that wins with a probability of 3^{-25}. This is indeed enough by Equation (1) as $R_c^-(6)$ has $6 \cdot 13 = 78 = 53 + 25$ vertices. Let us label the vertices of $R_c^-(6)$ such that the six vertices that are constructed from blowing up $v \in V(R_c^-)$ are labelled v_a, v_b, v_c, v_d, v_e, and v_f. Under this labelling our strategy for $R_c^-(6)$ is to play the complete graph strategy on the cliques

$$\{1_a, 2_a, 3_a\}, \quad \{1_b, 2_b, 7_a\}, \quad \{1_c, 3_b, 4_a\}, \quad \{2_c, 3_c, 9'_a\},$$
$$\{4_b, 5_a, 10'_a\}, \quad \{4_c, 5_b, 10'_b\}, \quad \{5_c, 6_a, 9'_b\}, \quad \{6_b, 7_b, 8_a\},$$
$$\{6_c, 7_c, 8_b\}, \quad \{8_c, 9'_c, 10'_c\}, \quad \{8_d, 9'_d, 10'_d\}, \quad \{8_e, 9'_e, 10'_e\},$$
$$\{8_f, 9'_f, 10'_f\},$$

and to play the R_c strategy on the vertices

$$\{1_d, 2_d, 3_d, 4_d, 5_d, 6_d, 7_d, 8'_a, 8'_b, 9_a, 9_b, 10_a, 10_b\},$$
$$\{1_e, 2_e, 3_e, 4_e, 5_e, 6_e, 7_e, 8'_c, 8'_d, 9_c, 9_d, 10_c, 10_d\},$$
$$\{1_f, 2_f, 3_f, 4_f, 5_f, 6_f, 7_f, 8'_e, 8'_f, 9_e, 9_f, 10_e, 10_f\}.$$

■

We finish this section by considering a problem motivated by the reversibility of networks in network coding. Given a digraph G, let $\mathrm{Reverse}(G)$ be the digraph formed from G by reversing all the edges, i.e. $uv \in E(G)$ if and only if $vu \in E(\mathrm{Reverse}(G))$.

Problem IV.7. *Does there exist a digraph G, such that* $\mathrm{gn}(G) \neq \mathrm{gn}(\mathrm{Reverse}(G))$.

We were not able to solve this problem. We did however find a graph R^S for which the Shannon bound of R^S and the Shannon bound of $\mathrm{Reverse}(R^S)$ did not match. R^S is simply the digraph formed by making vertex 1 in R a Superman vertex, in other words we add three directed edges to R: the edge going from 1 to 8, from 1 to 9, and from 1 to 10. Consequently $\mathrm{Reverse}(R^S)$ is the graph formed from making vertex 1 in R a Luthor vertex, as such we will refer to it is R^L.

Theorem IV.8. *The Shannon bound of R^S is $27/4$. For R^L:*

1) *The Shannon bound is $34/5 = 6.8$.*
2) *The Zhang-Yeung bound is $61/9 = 6.777\ldots$*
3) *The Dougherty-Freiling-Zeger bound is $359/53 = 6.773\ldots$*
4) *The Ingleton bound is $27/4 = 6.75$.*

Proof: The proofs are given in data files which can be made available upon request. ■

From the strategy on R we know $\mathrm{gn}(R^S) \geq 27/4$ and $\mathrm{gn}(R^L) \geq 27/4$. Hence we have $\mathrm{gn}(R^S) = 27/4$. We do not however know the value of $\mathrm{gn}(R^L)$ so it is possible that the asymptotic guessing numbers do not match.

V. Open Problems

Even though the fractional clique cover bound is not always the same as the asymptotic guessing number it is interesting to know whether they are equal for specific families of graphs. One such interesting family, where we do not know the answer, is that of the triangle-free graphs. If G is a triangle-free graph then Lemma II.3 tells us that $\kappa_f(G) \geq |V(G)|/2$ and so we conjecture that $\mathrm{gn}(G) \leq |V(G)|/2$ always holds for a triangle-free undirected graph G.

Problem IV.7 asks whether there exists an irreversible guessing game, i.e. a guessing game G such that $\mathrm{gn}(G) \neq \mathrm{gn}(\mathrm{Reverse}(G))$.

References

[1] R. Ahlswede, N. Cai, S-Y. R. Li and R. W. Yeung, Network information flow, IEEE Trans. Inform. Theory **46** (2000), 1204–1216.

[2] D. Christofides and K. Markström, The guessing number of undirected graphs, Electron. J. Combin. **18** (2011), Paper 192, 19 pp.

[3] A. Cohen, M. Effros, S. Avestimehr and R. Koetter, Linearly representable entropy vectors and their relation to network coding solutions, IEEE Information Theory Workshop (2009)

[4] T. M. Cover and J. A. Thomas, *Elements of information theory*, second edition, Wiley-Interscience, Hoboken, NJ, 2006.

[5] R. Dougherty, C. Freiling and K. Zeger, "Non-Shannon Information Inequalities in Four Random Variables", arXiv:1104.3602.

[6] S. Riis, Information flows, graphs and their guessing numbers, Electron. J. Combin. **14** (2007), Research Paper 44, 17 pp.

[7] S. Riis, Reversible and irreversible information networks, IEEE Trans. Inform. Theory **53** (2007), 4339–4349.

[8] Z. Zhang and R. W. Yeung, On characterization of entropy function via information inequalities, IEEE Trans. Inform. Theory **44** (1998), 1440–1452.