

ARTICLE

On the possible values of the entropy of undirected graphs

Maximilien Gadouleau 

Department of Computer Science, Durham University, Durham DH1 3LE, UK

Correspondence

Maximilien Gadouleau, Department of Computer Science, Durham University, Durham DH1 3LE, UK.
Email: m.r.gadouleau@durham.ac.uk

Abstract

The entropy of a digraph is a fundamental measure that relates network coding, information theory, and fixed points of finite dynamical systems. In this article, we focus on the entropy of undirected graphs. We prove any bounded interval only contains finitely many possible values of the entropy of an undirected graph. We also determine all the possible values for the entropy of an undirected graph up to the value of four.

KEYWORDS

clique cover, graph entropy, information inequalities, network coding, vertex cover

1 | INTRODUCTION

1.1 | Finite dynamical systems and their fixed points

Finite Dynamical Systems (FDSs) have been used to represent a network of interacting entities as follows. A network of n entities has a state $x = (x_1, \dots, x_n) \in [q]^n$, represented by a q -ary variable $x_v \in [q] = \{0, 1, \dots, q-1\}$ on each entity v , which evolves according to a deterministic function $f = (f_1, \dots, f_n) : [q]^n \rightarrow [q]^n$, where $f_v : [q]^n \rightarrow [q]$ represents the update of the local state x_v . FDSs have been used to model gene networks [18,19,26], neural networks [17], social interactions [16,20], and more (see [15,27]).

The architecture of an FDS $f : [q]^n \rightarrow [q]^n$ can be represented via its **interaction graph** $\text{IG}(f)$, which indicates which update functions depend on which variables. More formally, $\text{IG}(f)$ has $\{1, \dots, n\}$ as vertex set and there is an arc from u to v if $f_v(x)$ depends on x_u . In different contexts, the interaction graph is known—or at least well approximated—, while the actual update functions are not. One main problem of research on FDSs is then to predict their dynamics according to their interaction graphs.

Among the many dynamical properties that can be studied, **fixed points** are crucial because they represent stable states; for instance, in the context of gene networks, they correspond to stable pat-

terns of gene expression at the basis of particular biological processes. As such, they are arguably the property that has been the most thoroughly studied. The study of the number of fixed points and its maximization in particular is the subject of a stream of work, e.g. in [2–4, 13, 21, 22].

1.2 | Network coding and entropy of digraphs

Network coding is a technique to transmit information through networks, which can significantly improve upon routing in theory [1,28]. At each intermediate node v , the received messages x_{u_1}, \dots, x_{u_k} are combined, and the combined message $f_v(x_{u_1}, \dots, x_{u_k})$ is then forwarded toward its destinations. The main problem is to determine which functions f_v can transmit the most information. In particular, the **network coding solvability problem** tries to determine whether a certain network situation, with a given set of sources, destinations, and messages, is solvable, i.e. whether all messages can be transmitted to their destinations. This problem being very difficult, different techniques have been used to tackle it, including matroids [8], Shannon and non-Shannon inequalities for the entropy function [9,25], error-correcting codes [14], and closure operators [10,11].

The network coding solvability problem can be recast in terms of fixed points of FDSs as follows [24, 25]. The so-called **q -guessing number** [25] of a digraph D is the logarithm of the maximum number of fixed points over all FDSs f whose interaction graph is a subgraph of D : $\text{IG}(f) \subseteq D$. The guessing number is always upper bounded by the size of a minimum feedback vertex set of D ; if equality holds, we say that D is **solvable** over an alphabet of size q and the FDS f reaching this bound is called a solution. Then, a network coding instance N is solvable if and only if some digraph D_N related to the instance N is solvable (see [10, 12, 14] for further illustration of the relation between network coding and the guessing number).

For a given digraph D , the supremum over all q of the q -ary guessing numbers of D is referred to as the **entropy** of D . The entropy then represents the maximum amount of information that can be transmitted over a network; it is the capacity of network coding. Although the entropy can be determined for large classes of graphs (perfect graphs, for instance), the entropy of digraphs is not so well understood in general. For instance, very little is known about the set of all possible values of the entropy. Some graphs, notably the odd cycles and their complements, have noninteger entropies [7]; moreover, [5] exhibits a small graph (10 vertices) whose entropy cannot be expressed as a rational number with a small denominator. It is also worth noting that some famous simple triangle-free graphs, such as the Clebsch graph of the Higman-Sims graph, have an entropy above $n/2$ [6].

1.3 | Outline

The rest of the article is organized as follows. Section 2 reviews some important background on the guessing number and the entropy of graphs. Section 3 then gives our two main results. First, we show that for any integer k the number of possible values of the entropy of an undirected graph up to k is finite. We then determine all the possible values for the entropy of an undirected graph up to the value of four.

2 | BACKGROUND

We consider digraphs $D = (V, E)$, where $E \subseteq V^2$; we shall usually set $V = \{v_1, \dots, v_n\}$. An undirected graph is a digraph where E is a symmetric set, i.e. $(u, v) \in E$ if and only if $(v, u) \in E$; we view the pair of arcs $(u, v), (v, u)$ as an edge and denote it by uv . A simple graph is an undirected graph

without loops. For any digraph $D = (V, E)$ and any $S \subseteq V$, we denote the subgraph of D induced by S as $D[S]$:

$$D[S] = (S, \{(u, v) \in E : u, v \in S\}).$$

For any $S \subseteq V$, we also denote $D - S = D[V \setminus S]$.

We denote a bipartite graph with bipartition A and B and edge set E' by the triple (A, B, E') . Then, if G is a simple graph and $S, T \subseteq V$ with $T \cap S = \emptyset$, we denote the bipartite subgraph of G induced by S and T as $G[S, T]$:

$$G[S, T] = (S, T, \{uv \in E : u \in S, v \in T\}).$$

The neighborhood of a vertex u in G is $N(u; G) = \{v : uv \in E\}$; we shall usually omit the dependence in G . The neighborhood of a set S of vertices is $N(S) = \bigcup_{u \in S} N(u)$.

The guessing number and the entropy are formally defined as follows. Let $f : [q]^n \rightarrow [q]^n$. The **interaction graph** of f is the digraph on n vertices where $(u, v) \in \text{IG}(f)$ if and only if f_v depends essentially on x_u , i.e. there exist $a, b \in [q]^n$ that only differ by $a_u \neq b_u$ such that $f_v(a) \neq f_v(b)$. For any digraph D and any $q \geq 2$, we denote the set of all functions $f : [q]^n \rightarrow [q]^n$ interaction graph contained in D as

$$F(D, q) = \{f : [q]^n \rightarrow [q]^n : \text{IG}(f) \subseteq D\}.$$

The **guessing number** of $f : [q]^n \rightarrow [q]^n$ is $g(f) := \log_q |\text{Fix}(f)|$. The **q -guessing number** [25] of D is

$$g(D, q) := \max\{g(f) : f \in F(D, q)\}.$$

For any D , the q -guessing number tends to a limit when q tends to infinity, which we shall refer to as the **entropy** of D [7,23]:

$$H(D) := \sup_{q \geq 2} g(D, q) = \lim_{q \rightarrow \infty} g(D, q).$$

We now review some general properties of the guessing number of digraphs found in [7, 12, 14, 23, 25]. By definition, if $D' \subseteq D$, then $g(D', q) \leq g(D, q)$. For any D_1 and D_2 on disjoint vertex sets, we have

$$H(D_1 \cup D_2) = H(D_1) + H(D_2).$$

For any digraph D and any two induced subgraphs $D[V_1]$ and $D[V_2]$ of D such that $V_1 \cap V_2 = \emptyset$ and $V_1 \cup V_2 = V$, we have

$$g(D, q) \leq g(D[V_1], q) + |V_2|.$$

Equality is reached in an important special case. Let L be the set of vertices with a loop in D , then

$$g(D, q) = |L| + g(D - L, q).$$

Indeed, for any $f \in F(D - L, q)$, let $f' \in F(D, q)$ defined as $f'_l(x) = x_l$ for every $l \in L$ and $f'_v(x) = f_v(x)$ for every $v \notin L$. We then have $g(f') = |L| + g(f)$.

Let I be an acyclic set of D , then the I -reduction of D , denoted as $D^{-I} = (V^{-I}, E^{-I})$, is given as follows [12]. Let us denote $u \rightarrow_I v$ if there are vertices $i_1, \dots, i_k \in I$ such that u, i_1, \dots, i_k, v is a directed path in D . Then

$$V^{-I} = V \setminus I, \quad E^{-I} = \{(u, v) : u, v \in V^{-I}, (u, v) \in E \text{ or } u \rightarrow_I v\}.$$

By [12], for any acyclic set I of D and any $q \geq 2$,

$$g(D, q) \leq g(D^{-I}, q).$$

A **clique cover** of D is a set of cliques H_1, \dots, H_t such that every vertex belongs to at least one clique. The **clique cover number** of D , denoted by $\pi(D)$, is the minimum number of cliques in a clique cover of D . Since $g(K_n, q) = n - 1$ for all $n \geq 1$ and $q \geq 2$, we obtain that for any D and any $q \geq 2$,

$$g(D, q) \geq n - \pi(D) \geq \mu(D), \quad (1)$$

where $\mu(D)$ is the maximum matching number of D . In particular, if G is simple, then the clique cover number of G is equal to the chromatic number of its complement, thus $g(G, q) \geq n - \chi(\bar{G})$ for all $q \geq 2$. More specifically, if G is a perfect graph, then $g(G, q) = n - \pi(G)$ for all $q \geq 2$.

The lower bound on the guessing number in (1) is refined as follows. A **fractional clique cover** of a digraph D is a family of cliques H_1, \dots, H_t of D together with nonnegative weights w_1, \dots, w_t such that

$$\sum_{i: v \in H_i} w_i \geq 1 \quad \forall v \in V.$$

A clique cover is a fractional clique cover where all the weights belong to $\{0, 1\}$. The minimum value of $\sum_{i=1}^t w_i$ over all clique covers is the **fractional clique cover number** and is denoted by $\pi^*(D)$. Then there exists $k \geq 1$ such that for any $q \geq 2$

$$g(D, q^k) \geq n - \pi^*(D).$$

Let $\tau(D)$ denote the **transversal number** of D , i.e. the minimum number of vertices that have to be removed from D in order to obtain an acyclic digraph. If G is a simple graph, then $\tau(G)$ is the minimum vertex cover number of G . Then for any $q \geq 2$,

$$g(D, q) \leq \tau(D) \leq 2\mu(D).$$

A finer upper bound on the entropy of a digraph is based on the submodularity of the entropy of random variables [25]. The **Shannon entropy** of D is defined as

$$\eta(D) := \sup h(V),$$

where the supremum is taken over all functions $h : 2^V \rightarrow \mathbb{R}$ satisfying

$$\begin{aligned} h(v) &\leq 1 & \forall v \in V, \\ h(S) &\leq h(T) & \forall S \subseteq T \subseteq V, \\ h(S \cup T) + h(S \cap T) &\leq h(S) + h(T) & \forall S, T \subseteq V, \end{aligned}$$

$$h(N(v) \cup \{v\}) = h(N(v)) \quad \forall v \in V.$$

We then have for any $q \geq 2$

$$g(D, q) \leq \eta(D).$$

Based on the Shannon entropy, we obtain that for odd cycles, $H(C_{2l+1}) = l + 1/2$ and for their complements, $H(\overline{C_{2l+1}}) = 2l - 1 - 1/l$ for any $l \geq 2$ [7].

3 | MAIN RESULTS

Let the set of all entropies of simple graphs be

$$\mathcal{H} = \{H(G) : G \text{ simple, finite}\}.$$

We remark that this is the same as considering undirected graphs.

By the results above, the set \mathcal{H} is closed under addition and contains all natural numbers. The only noninteger values known so far are based on odd cycles $H(C_{2l+1}) = l + 1/2$ and their complements $H(\overline{C_{2l+1}}) = 2l - 1 - 1/l$ for $l \geq 2$ [7] and some additional results in [5]. Our first main result is that $\mathcal{H} \cap [0, k]$ is finite for every integer k ; more precisely, it is $O(2^{6k^2})$.

Theorem 3.1. *For any positive integer k ,*

$$|\mathcal{H} \cap [k-1, k]| \leq \sum_{h=1}^k \sum_{m=\lfloor h/2 \rfloor}^{h-1} 2^{2m(3m-2)}.$$

The proof is based on the following lemma. We say that a matching saturates a set S of vertices if every vertex in S belongs to an edge of the matching.

Lemma 3.2. *Let $G = (A, B, E)$ be a nonempty bipartite graph with $|A| \geq |B|$. Then there exists a nonempty $A' \subseteq A$ such that $G[A', N(A')]$ has a matching saturating $N(A')$.*

Proof. This is clearly true for $|B| = 1$. Let G be a counterexample with minimizes first the value of $|B| \geq 2$, and then the value of $|A|$.

Secondly, suppose that for all $A' \subseteq A$, $|N(A')| \geq |A'|$. Then by Hall's marriage theorem there exists a matching in G saturating A . Thus $|A| = n$ and we obtain a matching in $G[A, N(A)] = G$ saturating $B = N(A)$.

Thirdly, suppose that there exists a nonempty $A' \subset A$ such that $|N(A')| \leq |A'|$. Denote $D = G[A' \cup N(A')] = G[A', N(A')]$. We remark that for any $A'' \subseteq A'$, we have $N(A''; D) = N(A''; G)$. By minimality hypothesis there exists $A'' \subseteq A'$ such that $D[A'', N(A''; D)] = G[A'', N(A''; G)]$ has a matching saturating $N(A''; D) = N(A''; G)$. ■

We say G is **entropy-minimal** if for any G' with $|G'| < |G|$, $H(G') - \lfloor H(G') \rfloor \neq H(G) - \lfloor H(G) \rfloor$. For any nonempty $S \subseteq V$, we denote

$$c(S) := \{v \in V \setminus S : N(v) \subseteq S\}.$$

We remark that $c(S)$ is an independent set.

Lemma 3.3. *Let G be entropy-minimal. Then for any nonempty $S \subseteq V$, $G[c(S), S]$ does not have a matching saturating S .*

Proof. Suppose there exists S such that $G[c(S), S]$ has a matching saturating S . Denoting $d(S) = c(S) \cup S$, we prove that $H(G) = |S| + H(G - d(S))$, which contradicts entropy-minimality of G (for $G' = G - d(S)$). By reducing $c(S)$, we obtain $H(G) \leq H(G^{-c(S)})$. Since $G^{-c(S)}$ has a loop on every vertex in S and $G^{-c(S)} - S = G - d(S)$, we obtain

$$H(G) \leq H(G^{-c(S)}) = |S| + H(G^{-c(S)} - S) = |S| + H(G - d(S)).$$

For the reverse inequality, let $f \in F(G - d(S), q)$ achieve the guessing number for $G - d(S)$. We extend it to $f' \in F(G, q)$ by considering the matching in $G[c(S), S]$ saturating S . Let $c_1 s_1, \dots, c_k s_k$ be the matching saturating S where $c_i \in c(S)$ and $s_i \in S$, then

$$f'_v(x) = \begin{cases} f_v(x_{V \setminus d(S)}) & \text{if } v \in V \setminus d(S) \\ x_{c_i} & \text{if } v = s_i \quad 1 \leq i \leq k \\ x_{s_i} & \text{if } v = c_i \quad 1 \leq i \leq k \\ 0 & \text{otherwise.} \end{cases}$$

Then $g(f') = g(f) + k$ and we obtain $H(G) \geq |S| + H(G - d(S))$. ■

Proof of Theorem 3.1. Let G be entropy-minimal with $H(G) \in (h - 1, h)$ for some $1 \leq h \leq k$. Let M be the vertices in a maximum matching of G (with m edges). We then have $h \leq 2m = |M| \leq 2(h - 1)$. The induced subgraph $G - M$ is an independent set, hence $c(M) = V \setminus M$ and the edges in G either belong to $G[M]$ or to $G[c(M), M]$. Moreover, for every $c' \subseteq c(M)$, $N(c'; G) = N(c'; G[c(M), M])$. If $|c(M)| \geq |M|$, then by Lemma 3.2, there exists $c' \subseteq c(M)$ such that $G[c', N(c'; G)]$ has a matching saturating $N(c'; G)$ and by Lemma 3.3, G is not entropy-minimal. Therefore, $|c(M)| < 2m$.

Altogether, there are at most $2^{2m(m-1)}$ choices for $G[M]$, at most $2^{2m(2m-1)}$ choices for $G[c(M), M]$ and hence at most $2^{2m(3m-2)}$ choices for G . It is clear that choosing G empty, apart from the matching in M , yields an entropy of $h - 1$. Also, choosing G to be empty, apart from the matching in M with $m = h - 1$ edges, and two more edges cm_1, cm_2 where $c \in c(M)$ and $m_1 m_2$ is an edge of the matching yields an entropy of h . Thus, adding up for all possible values of h and m , we obtain the result. ■

In Question 3.4 below, we then ask the corresponding question for the set of entropies of digraphs. The closure solvability problem [10] generalizes the network coding solvability problem. Then for any closure operator on V we can associate its entropy; the entropy of a digraph is then the entropy of the corresponding closure operator. The analog of Question 3.4 for the set $\mathcal{H}'' \supseteq \mathcal{H}$ of entropies of closure operators was answered negatively in [11]: $\mathcal{H}'' \cap [0, k]$ contains all rational numbers between 1 and k .

Question 3.4. Let $\mathcal{H}' = \{H(D) : G \text{ digraph, finite}\}$. Is $\mathcal{H}' \cap [0, k]$ finite for all integers k ?

We are now interested in determining the first values in \mathcal{H} . The first three values in \mathcal{H} are already known. We have $\mathcal{H} \cap [0, 1] = \{0, 1\}$, since a graph is either empty (and its entropy is equal to zero) or it contains an edge (and its entropy is at least one). The entropy of any directed graph greater than one is at least equal to two [10]. We shall give an elementary and short proof of this result for undirected graphs (namely $\mathcal{H} \cap [1, 2] = \{1, 2\}$). We then determine the values in \mathcal{H} up to the value four. We also show that if G is a connected undirected graph with entropy $5/2$, then G is the pentagon; similarly, if G is a connected undirected graph with entropy $11/3$, then G is the graph G_1 displayed on Figure 1.

Theorem 3.5. $\mathcal{H} \cap [0, 4] = \{0, 1, 2, 5/2, 3, 7/2, 11/3, 4\}$.

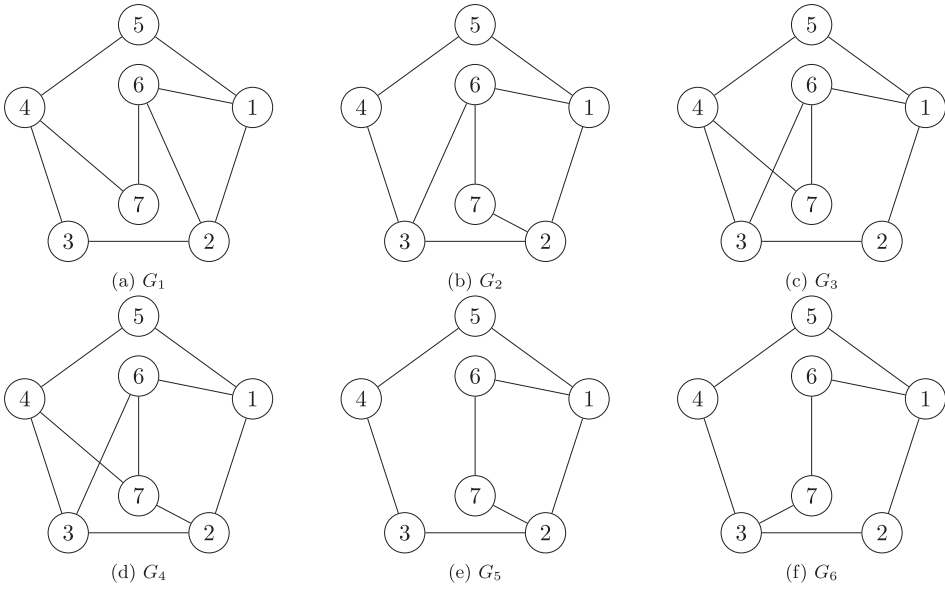


FIGURE 1 The graphs G_1, \dots, G_6

Lemma 3.6. *Let G be a graph on six vertices $V = \{v_1, \dots, v_6\}$ where $G[\{v_1, \dots, v_5\}] = C_5$. Then $H(G) = 2.5$ if v_6 is isolated, $H(G) = 3.5$ if $N(v_6)$ contains three consecutive vertices of C_5 , and $H(G) = 3$ otherwise.*

Proof. The result is clear when v_6 is isolated. Suppose that $N(v_6)$ contains three consecutive vertices, say v_1, v_2 , and v_3 . We have $H(G) \leq H(C_5) + 1 = 3.5$. Conversely, by fractional clique packing, we obtain $H(G) \geq 3.5$. Otherwise, $N(v_6) \subseteq S$, where S is a minimum feedback vertex set of $G[\{v_1, \dots, v_5\}]$. Therefore, S is a minimum feedback vertex set of G of size 3, thus $H(G) \leq 3$. Conversely, say $v_5v_6 \in E$, then $\{v_1v_2, v_3v_4, v_5v_6\}$ is a matching of size 3, thus $H(G) \geq 3$. ■

Proof of Theorem 3.5. We consider a graph G such that $H(G') \neq H(G)$ for all $|G'| < |G|$. By minimality, G has no isolated vertices. By construction, we have $\{0, 1, 2, 2.5, 3, 3.5, 4\} \subseteq \mathcal{H} \cap [0, 4]$. We now prove that the only other value in $\mathcal{H} \cap [0, 4]$ is $11/3$.

Suppose G has entropy between 1 and 2. Let $ab \in E$, then all the other vertices must be adjacent to a (or by symmetry, to b), i.e. $V = N(a) \cup a$. The neighborhood of a must form an independent set, for if $b, c \in N(a)$ are adjacent, then G contains the triangle a, b, c and $H(G) \geq 2$. But then, $\{a\}$ is a vertex cover and $H(G) \leq 1$.

Suppose G has entropy between 2 and 3. Since G is not perfect, it contains C_5 as an induced subgraph. If $G = C_5$, then $H(G) = 2.5$. Otherwise, Lemma 3.6 shows that $H(G) \geq 3$.

Suppose that the entropy of G is between 3 and 4 and is not equal to 3.5. Then G is entropy-minimal. If G contains C_7 as an induced subgraph, then either $G = C_7$ and $H(G) = 3.5$ or G has eight or more vertices, and an argument similar to that used in Lemma 3.6 indicates that $H(G) \geq 4$. Thus, G contains C_5 as an induced subgraph, say by vertices v_1, \dots, v_5 . The rest of the proof will show that the entropy of G is $11/3$. The first step, culminating in Claim 3 characterizes $G[\{v_6, \dots, v_n\}]$. ■

Claim 1. $G[\{v_6, \dots, v_n\}]$ has at most one edge.

Proof. Let $D = G[\{v_6, \dots, v_n\}]$ have at least two edges (hence $n \geq 8$). If D has a triangle or a matching of size two, then $H(G) \geq H(C_5) + 2 \geq 4$. Hence all the edges in D are adjacent to one vertex, say v_6 ,

and let $v_6v_7, v_6v_8 \in E$. If $N(v_7; G) = \{v_6\}$, then denoting $S = \{v_6\}$, we have $v_7 \in c(S)$ and hence the edge v_6v_7 is a matching in $G[c(S), S]$ saturating S , which contradicts Lemma 3.3. Therefore, v_7 is adjacent to some vertex in $\{v_1, \dots, v_5\}$, say $v_1v_7 \in E$. Then $\{v_2v_3, v_4v_5, v_1v_7, v_6v_8\}$ is a matching in G and $H(G) \geq 4$. ■

Claim 2. $G[\{v_6, \dots, v_n\}]$ has exactly one edge.

Proof. If G has six vertices, then its entropy is either 3 or 3.5 by Lemma 3.6.

Let G contain at least seven vertices, such that $\{v_6, \dots, v_n\}$ is an independent set. We first remark that for any $i \geq 6$, $|N(v_i; G)| \geq 2$. Indeed, if v_i is only adjacent to say v_j ($1 \leq j < i \leq n$), then $S = \{v_j\}$ violates Lemma 3.3. Moreover, $N(\{v_6, \dots, v_n\}; G)$ is not contained in a minimum vertex cover of $G[\{v_1, \dots, v_5\}]$, since otherwise $H(G) \leq \tau(G) = \tau(C_5) = 3$. Hence $N(\{v_6, \dots, v_n\})$ contains three consecutive vertices in the pentagon, say $v_1, v_2, v_3 \in N(\{v_6, \dots, v_n\})$. Say v_6 has the most neighbors in $\{v_1, v_2, v_3\}$ among v_6, \dots, v_n ; we finish the proof by a case analysis on $N(v_6)$.

1. $\{v_1, v_2, v_3\} \subseteq N(v_6)$. Then $v_1v_2v_6, v_2v_3v_6$, and all the other edges of the pentagon are cliques. Since v_7 has at least two neighbors on the pentagon, $\{v_1, v_3, v_4, v_5\} \cap N(v_7) \neq \emptyset$. We then consider all possible vertices for the intersection. If $v_1v_7 \in E$, then $\{v_2v_3v_6, v_3v_4, v_5v_7\}$ are disjoint cliques and hence $H(G) \geq 4$; the case $v_3v_7 \in E$ is similar. If $v_4v_7 \in E$, then $\{v_2v_3v_6, v_1v_2, v_4v_7\}$ are disjoint cliques and hence $H(G) \geq 4$; the case $v_5v_7 \in E$ is similar.
2. $N(v_6) = \{v_1, v_2\}$. Then v_3 must be contained in the neighborhood of a vertex in $\{v_7, \dots, v_n\}$, say $v_3 \in N(v_7)$. Then $\{v_1v_2v_6, v_3v_7, v_4v_5\}$ are disjoint cliques and hence $H(G) \geq 4$. The case where $N(v_6) = \{v_2, v_3\}$ is similar.
3. $N(v_6) = \{v_1, v_3\}$. Denoting $S = \{v_1, v_3\}$, we have $c(S) \subseteq \{v_2, v_6\}$ and $G[S, c(S)]$ contains the matching $\{v_1v_2, v_3v_6\}$ saturating S , which violates Lemma 3.3.
4. $|N(v_6)| = 1$. Then without loss of generality, $N(v_6) = \{v_1\}$, $N(v_7) = \{v_2\}$, $N(v_8) = \{v_3\}$ and $\{v_1v_6, v_2v_7, v_3v_8, v_4v_5\}$ is a matching and hence $H(G) \geq 4$. ■

Claim 3. $n = 7$ and $G[\{v_6, v_7\}] = K_2$.

Proof. We only have to prove that G has no more than seven vertices. Suppose G has eight vertices or more, and that the only edge in $G[\{v_6, \dots, v_n\}]$ is v_6v_7 . Say $v_1v_8 \in E$, then $\{v_1v_8, v_2v_3, v_4v_5, v_6v_7\}$ is a matching in G and hence $H(G) \geq 4$. ■

The second step reduces the search to just six possible graphs. We define the graphs G_1, \dots, G_6 displayed in Figure 1.

Claim 4. $G \in \{G_1, \dots, G_6\}$.

Proof. First, $N(v_6) \cap N(v_7) = \emptyset$. Otherwise, say $v_1v_6, v_1v_7 \in E$, then $\{v_1v_6v_7, v_2v_3, v_4v_5\}$ are disjoint cliques and $H(G) \geq 4$. Suppose $|N(v_6)| \geq |N(v_7)|$, we then have $|N(v_7)| \geq 2$.

1. $\{v_1, v_2, v_3\} \subseteq N(v_6)$. Then since $v_2v_7 \notin E$, we can use Case 1 of the proof of Claim 2 to show that $H(G) \geq 4$.
2. $N(v_6) = \{v_1, v_2, v_4, v_7\}$. Then $v_3v_7 \in E$ (or $v_5v_7 \in E$, but this is equivalent by symmetry) and $\{v_1v_2v_6, v_3v_7, v_4v_5\}$ are disjoint cliques and $H(G) \geq 4$.
3. $N(v_6) = \{v_1, v_2, v_7\}$. Again, if $v_3v_7 \in E$ (or $v_5v_7 \in E$) then $\{v_1v_2v_6, v_3v_7, v_4v_5\}$ are disjoint cliques and $H(G) \geq 4$. Otherwise, $G = G_1$.
4. $N(v_6) = \{v_1, v_3, v_7\}$. Then since $v_6 \in N(v_7)$, $v_1, v_3 \notin N(v_7)$, and $2 \leq |N(v_7)| \leq 3$, we have $G \in \{G_2, G_3, G_4\}$.

5. $N(v_6) = \{v_1, v_7\}$. Then $G \in \{G_5, G_6\}$. ■

The third and final step determines the entropy of G_1, \dots, G_6 .

Claim 5. $H(G_1) = 11/3$ and $H(G_2) = \dots = H(G_6) = 7/2$.

Proof. Direct calculations show that the Shannon entropy of G_1 is $\eta(G_1) = 11/3$. Conversely, the fractional clique covering number is $10/3$, which is achieved by the following weights:

$$w(v_1v_2v_6) = w(v_3v_4) = w(v_4v_5) = w(v_4v_7) = 1/3, \quad w(v_1v_5) = w(v_2v_3) = w(v_6v_7) = 2/3.$$

Therefore, $11/3 = n - \pi^*(G_1) \leq H(G_1) \leq \eta(G_1) = 11/3$. For $G_i, i \geq 2$, direct calculations show that $\eta(G_i) = 7/2$. Conversely, $H(G_i) \geq H(C_5) + 1 = 7/2$. ■

ACKNOWLEDGMENT

This work was supported by the Engineering and Physical Sciences Research Council [grant number EP/K033956/1].

ORCID

Maximilien Gadouleau  <http://orcid.org/0000-0003-4701-738X>

REFERENCES

- [1] R. Ahlswede et al., *Network information flow*, IEEE Trans. Inform. Theory **46** (4)(2000), 1204–1216.
- [2] J. Aracena, J. Demongeot, and E. Goles, *Fixed points and maximal independent sets in AND-OR networks*, Discrete Appl. Math. **138** (2004), 277–288.
- [3] J. Aracena, A. Richard, and L. Salinas, *Maximum number of fixed points in AND-OR-NOT networks*, J. Comp. Syst. Sci. **80** (2014), 1175–1190.
- [4] J. Aracena, *Maximum number of fixed points in regulatory boolean networks*, Bull. Math. Biol. **70** (2008), 1398–1409.
- [5] R. Baber et al., *Multiple unicasts, graph guessing games, and non-Shannon inequalities*, Proc. NetCod (2013), 1–6.
- [6] P. J. Cameron, A. N. Dang, and S. Riis, *Guessing games on triangle-free graphs*, Electronic J. Comb. **23** (1)(2016), 1–48.
- [7] D. Christofides and K. Markström, *The guessing number of undirected graphs*, Electronic J. Combin. **18** (1)(2011), 1–19.
- [8] R. Dougherty, C. Freiling, and K. Zeger, *Unachievability of network coding capacity*, IEEE Trans. Inform. Theory **52** (6)(2006), 2365–2372.
- [9] R. Dougherty, C. Freiling, and K. Zeger, *Networks, matroids, and non-Shannon information inequalities*, IEEE Trans. Inform. Theory **53** (6)(2007), 1949–1969.
- [10] M. Gadouleau, *Closure solvability for network coding and secret sharing*, IEEE Trans. Inform. Theory, **59** (12)(2013), 7858–7869.
- [11] M. Gadouleau, *Entropy of closure operators and network coding solvability*, Entropy **16** (9)(2014), 5122–5143.
- [12] M. Gadouleau, A. Richard, and E. Fanchon, *Reduction and fixed points of boolean networks and linear network coding solvability*, IEEE Transactions on Information Theory, **62** (5)(2016), 2504–2519.
- [13] M. Gadouleau, A. Richard, and S. Riis, *Fixed points of boolean networks, guessing graphs, and coding theory*, SIAM Journal on Discrete Mathematics **29** (4)(2015) 2312–2335.
- [14] M. Gadouleau and S. Riis, *Graph-theoretical constructions for graph entropy and network coding based communications*, IEEE Trans. Inform. Theory **57** (10)(2011), 6703–6717.

- [15] E. Goles and S. Martínez, *Neural and automata networks: Dynamical behavior and applications*, Kluwer Academic Publishers, Norwell, MA, USA, 1990.
- [16] E. Goles and M. Tchuente, *Iterative behaviour of generalized majority functions*, Math. Soc. Sci. **4** (1983), 197–204.
- [17] J. Hopfield, *Neural networks and physical systems with emergent collective computational abilities*, Proc. Nat. Acad. Sci. USA **79** (1982), 2554–2558.
- [18] G. Karlebach and R. Shamir, *Modelling and analysis of gene regulatory networks*, Nature **9** (2008), 770–780.
- [19] S. A. Kauffman, *Metabolic stability and epigenesis in randomly connected nets*, J. Theoret. Biol. **22** (1969), 437–467.
- [20] S. Poljak and M. Sura, *On periodical behaviour in societies with symmetric influences*, Combinatorica **3** (1983), 119–121.
- [21] E. Remy, P. Ruet, and D. Thieffry, *Graphic requirements for multistability and attractive cycles in a Boolean dynamical framework*, Adv. Appl. Math. **41** (2008), 335–350.
- [22] A. Richard, *Positive circuits and maximal number of fixed points in discrete dynamical systems*, Discrete Appl. Math. **157** (2009), 3281–3288.
- [23] S. Riis, *Utilising public information in network coding*, General Theory of Information Transfer and Combinatorics, Lecture Notes in Computer Science, vol. 4123/2006, Springer, Berlin, Heidelberg, 2006, pp. 866–897.
- [24] S. Riis, *Graph entropy, network coding and guessing games*, available at <http://arxiv.org/abs/0711.4175>, November 2007.
- [25] S. Riis, *Information flows, graphs and their guessing numbers*, Electronic J. Combin. **14** (2007), 1–17.
- [26] R. Thomas, *Boolean formalization of genetic control circuits*, J. Theoret. Biol. **42** (1973), 563–585.
- [27] R. Thomas and R. D'Ari, *Biological feedback*, CRC Press, Boca Raton, Florida, USA, 1990.
- [28] R. W. Yeung et al., Network coding theory, vol. 2, *Foundation and Trends in Communications and Information Theory*, no. 4–5, now Publishers, Hanover, MA, 2006.

How to cite this article: Gadouleau M. On the possible values of the entropy of undirected graphs. *J Graph Theory*. 2018;88:302–311. <https://doi.org/10.1002/jgt.22213>