

# Ulam’s pathological liar game with one half-lie

Robert B. Ellis\*  
Catherine H. Yan†

Texas A&M University

## Abstract

We introduce a dual game to Ulam’s liar game and consider the case of one half-lie. In the original Ulam’s game, Paul attempts to isolate a distinguished element by disqualifying all but one of  $n$  possibilities with  $q$  Yes-No questions, while the responder Carole is allowed to lie a fixed number  $k$  of times. In the dual game, Paul attempts to prevent disqualification of a distinguished element by “pathological” liar Carole for as long as possible, given that a possibility associated with  $k + 1$  lies is disqualified. We consider the half-lie variant in which Carole may only lie when the true answer is “No”. We prove the equivalence of the dual game to the problem of covering the discrete hypercube with certain asymmetric sets. We define  $A_1^*(q)$  for the case  $k = 1$  to be the minimum number  $n$  such that Paul can prevent Carole from disqualifying all  $n$  elements in  $q$  rounds of questions, and prove that  $A_1^*(q) \sim 2^{q+1}/q$ .

## 1 Introduction

In the original Ulam’s liar game, a seeker attempts to isolate a distinguished element by disqualifying all but one of  $n$  possibilities with  $q$  Yes-No questions. Given that the responder is allowed to lie a fixed number  $k$  of times, there is a maximum  $n$  for which the seeker is guaranteed a strategy to find the distinguished number within  $q$  rounds of questions. In this paper, we present a broadened framework of Ulam-type games which encompasses both the original liar game and a new *pathological liar* game which is the dual to the original liar game. Confidence that the pathological liar game is *the* dual to the liar game arises from the equivalence of the liar game to packing certain sets in the  $q$ -cube, and the equivalence of the pathological liar game to covering the  $q$ -cube with the same type of sets. We clarify the original and dual games by presenting them in equivalent vector formats, making apparent the equivalence of the original and pathological games to packing and covering, respectively, in the  $q$ -cube. The main result of this paper is an asymptotic solution in Theorem 1 to the optimal size of the space for which Paul wins the  $q$ -question pathological liar game with one half-lie.

### 1.1 A framework for Ulam-type games

An Ulam-type game is played by two players, whom we call Paul and Carole. The game begins with a set of  $n$  elements in play, which we may consider to be the set  $[n] := \{1, \dots, n\}$  where  $n \in \mathbb{Z}^+$ , and a *distinguished* element  $x \in [n]$ . The game proceeds by rounds each consisting of

---

\*Research supported in part by NSF grant DMS-9977354.

†Research supported by NSF grant DMS-0070574 and a Sloan Fellowship. This author is also affiliated with the Dalian University of Technology.

a question by Paul and a response by Carole. A question by Paul is of the form “is  $x \in A$ ”, where  $A \subseteq [n]$ . Carole must respond “Yes” or “No”, and is allowed to lie at most  $k$  times. Paul may formulate his questions *adaptively*, taking into consideration Carole’s previous responses. An element is *disqualified* if taking it to be the distinguished element requires Carole to have lied about it  $k + 1$  times. An element which is not yet disqualified is a *surviving* element. Without loss of generality, we assume that  $A$  consists only of surviving elements. The game *terminates* when a specified condition is achieved, such there being at most 1 surviving element. Paul plays a strategy to either achieve or delay the terminating condition, and Carole plays an adversary strategy to Paul. By this we think of neither actually selecting the distinguished element  $x$ ; instead, both play consistently with at least one surviving element possibly being  $x$ .

## 1.2 The original liar game and its vector game equivalent

The original Ulam’s liar game is a searching game in which Paul attempts to “find” the distinguished element  $x$  in a minimum number of rounds by forcing disqualification of all other elements, while Carole answers adversely. Carole does not pick  $x$  at the beginning, but rather answers consistently so that there is an  $x$  she could have picked. The game terminates when there is exactly 1 surviving element, which is “found” by Paul; since Carole responds adversely, the game will always terminate with 1 surviving element (never will *all* elements be disqualified). The largest  $n$  such that Paul can force game termination in  $q$  rounds has been determined exactly in the case of 1 lie [13], 2 lies [11], and 3 lies [7]; and asymptotically in the number of questions in [14]. In the *half-lie* variant of Ulam’s liar game, Carole is only allowed to lie when the truthful answer is “No”, so that every “No” Paul hears is truthful. Thus if Carole answers “Yes”, no lies are assigned to elements of  $A$ , but 1 lie is assigned to each of the surviving elements in  $[n] \setminus A$ . If Carole answers “No”, no lies are assigned to surviving elements of  $[n] \setminus A$ , but all elements of  $A$  are disqualified. The case  $k = 1$  for search spaces of size  $2^m$  is solved in [4], and an asymptotic analysis in  $q$  for fixed  $k$  appears in [8] and is improved in [15].

The original liar game with  $k$  half-lies is equivalent to the following vector game, which arises from an encoding due to Berlekamp [3]. The initial vector of the game on  $n$  elements is the vector  $(n, 0, \dots, 0)$  with  $k + 1$  entries. An intermediate state of the game after some number of rounds is represented by the vector  $\vec{x} = (x_0, x_1, \dots, x_k)$ , where  $x_i$  counts the number of elements of  $[n]$  associated with  $i$  lies. In particular,  $x_i$  is a nonnegative integer for all  $0 \leq i \leq k$ . Disqualified elements are not counted, and so  $\|\vec{x}\|_1 := \sum_{i=0}^k x_i$  counts the total number of surviving elements left in the game. The question “is  $x \in A$ ” by Paul is associated with the vector  $\vec{a} = (a_0, a_1, \dots, a_k)$ , where  $a_i$  counts the number of elements in  $A$  associated with  $i$  lies. Clearly,  $0 \leq a_i \leq x_i$  for all positions  $0 \leq i \leq k$ . If Carole answers “Yes”, all elements in  $A$  stay in the same position in the vector, but all surviving elements in  $[n] \setminus A$  move one position to the right, as they have become associated with an additional lie. An element moving to the right from position  $k$ , the right-most position, is disqualified and removed from the game. If Carole answers “No”, she is not lying, and so all elements in  $A$  are disqualified and removed entirely from the vector. Thus we define the new state vector based on the original state  $\vec{x}$ , Paul’s question  $\vec{a}$ , and Carole’s answer, as one of the following:

$$\begin{aligned} Y(\vec{x}, \vec{a}) &:= (a_0, a_1 + x_0 - a_0, \dots, a_k + x_{k-1} - a_{k-1}) \\ N(\vec{x}, \vec{a}) &:= (x_0 - a_0, x_1 - a_1, \dots, x_k - a_k). \end{aligned} \tag{1}$$

Therefore Paul wins the  $q$ -round game on  $n$  elements with  $k$  half-lies if the vector game with initial state  $(n, 0, \dots, 0)$  terminates in at most  $q$  rounds with  $\|\vec{x}\|_1 \leq 1$ . For a state vector  $\vec{x}$ , define  $r(\vec{x})$  to be the minimum number of rounds in the game starting in state  $\vec{x}$  for which Paul has a winning

strategy. Then  $r(\vec{x})$  satisfies

$$r(\vec{x}) = 1 + \min_{\vec{a}} \{ \max \{ r(Y(\vec{x}, \vec{a})), r(N(\vec{x}, \vec{a})) \} \}, \quad (2)$$

where the minimum is over all possible questions  $\vec{a}$ , and  $r(\vec{x}) = 0$  when  $\|\vec{x}\|_1 \leq 1$ . For all  $q \geq 0$ ,  $A_k(q)$  is defined to be the maximum  $n$  such that  $r((n, 0, \dots, 0)) = q$ ; i.e.,  $A_k(q)$  is the maximum  $n$  such that Paul has a strategy to find the distinguished element  $x$  in  $q$  rounds in the  $k$  half-lie game. For fixed  $k$ ,  $A_k(q)$  was shown in [8] to be asymptotically  $2^{q+k}/\binom{q}{k}$ .

### 1.3 The pathological half-liar game

In the pathological liar game, Paul attempts to preserve the distinguished element  $x$  from disqualification for as long as possible, while Carole lies “pathologically” to disqualify  $x$ . This time, we take the viewpoint dual to that in the original game that Paul does not have to pick  $x$  at the beginning, but rather asks questions consistent with there being an  $x$  he could have picked. We are concerned with the half-lie variant, in which Carole is only allowed to lie when the truthful answer is “No”. We first present the pathological half-lie game in its vector form which is dual to that of the original half-lie game, and then discuss its interpretation. The analysis of the pathological full-lie game is similar to that of the original full-lie game, and will appear in [9].

**The pathological half-liar game:** The pathological liar game on  $n$  elements with  $k$  half-lies has initial state  $(n, 0, \dots, 0)$  and terminating state  $(0, \dots, 0, 1)$ , both  $(k+1)$ -vectors. As before, the game progresses from initial to terminal state by means of question/answer pairs, where a question by Paul is a nonnegative  $(k+1)$ -vector  $\vec{a} \leq \vec{x}$ , where  $\vec{x}$  is the previous state vector, and an answer by Carole is either “Yes” or “No”. As in (1), the next state is  $Y(\vec{x}, \vec{a}) := (a_0, a_1 + x_0 - a_0, \dots, a_k + x_{k-1} - a_{k-1})$  if Carole’s answer is “Yes” and  $N(\vec{x}, \vec{a}) := (x_0 - a_0, x_1 - a_1, \dots, x_k - a_k)$  if Carole’s answer is “No”. Paul plays to delay termination for as long as possible, while Carole can and does answer adversely, always choosing the response which forces earlier game termination.

Note that the game will always reach the terminating state without first reaching the zero vector, because otherwise Paul’s strategy is suboptimal; thus Paul’s goal is to maximize the number of rounds for which the state vector is nonzero. For the state vector  $\vec{x}$ , define  $r^*(\vec{x})$  to be the maximum number of rounds in the game with initial state  $\vec{x}$  for which Paul can force the game to last; dually to (2), we have

$$r^*(\vec{x}) = 1 + \max_{\vec{a}} \{ \min \{ r^*(Y(\vec{x}, \vec{a})), r^*(N(\vec{x}, \vec{a})) \} \}, \quad (3)$$

where the maximum is over all possible questions  $\vec{a}$ , and  $r^*((0, \dots, 0, 1)) = r^*(\vec{0}) = 0$ . For all  $q \geq 0$ ,  $A_k^*(q)$  is defined to be the minimum  $n$  such that  $r^*((n, 0, \dots, 0)) = q$ ; i.e.,  $A_k^*(q)$  is the minimum  $n$  such that Paul has a strategy to delay game termination for  $q$  rounds in the  $k$  half-lie game. Having a larger  $k$  actually helps Paul prolong the game, as it increases the threshold for Carole to disqualify any given element.

Interpreted in the original game form, the pathological game on  $n$  elements corresponds to Paul seeking to preserve a distinguished element  $x \in [n]$  from disqualification for as long as possible, with Carole answering adversely. Answers of “Yes” and “No” by Carole to the question “is  $x$  in  $A$ ” are processed as in the original game, either by assigning an additional lie to surviving elements in  $[n] \setminus A$ , or by immediately disqualifying all elements of  $A$ , respectively. (That is, an answer of “Yes” by Carole may be truthful or mendacious, but an answer of “No” by her is always thought to be truthful.) We remark that under this interpretation, if Paul asks “is  $x$  in  $B$ ”, where  $B$  is

the set of all surviving elements, then a response “No” by Carole would immediately disqualify all elements, and hence end the game.

Under this interpretation, the max and min in (3) correspond to Paul’s choosing the question  $\vec{a}$  to which Carole’s most disruptive answer allows at least one element to survive for as long as possible. The state  $(0, \dots, 0, 1, 0, \dots, 0)$ , in which there is exactly 1 surviving element which is not in position  $k$ , is not terminal, because Paul may ask the question “is  $x$  in  $\emptyset$ ”, to which Carole responds “Yes”, since responding “No” does not change the state of the game. Furthermore, by the remark at the end of previous paragraph, it is suboptimal for Paul to ask the question “is  $x$  in  $B$ ”, where  $B$  is the set of surviving elements, because Carole can immediately end the game by answering “No”.

Paul wins a  $q$ -round game on  $n$  elements with  $k$  half-lies if he can prevent the game from terminating before the completion of  $q$  rounds, or in other words, if he can preserve at least one element for  $q$  rounds. In this paper, we are concerned with the pathological half-lie variant in the case  $k = 1$ . For a fixed  $q \geq 0$ , define  $A_1^*(q)$  to be the minimum  $n$  such that Paul has a strategy to win the  $q$ -round pathological liar game in which Carole is allowed one half-lie. We prove the following asymptotic behavior of  $A_1^*(q)$ .

**Theorem 1.** Let  $q > 0$ . Then

$$A_1^*(q) \sim \frac{2^{q+1}}{q}.$$

The proof of Theorem 1 is divided into the remaining sections of the paper. In Section 2, we show the equivalence of the pathological liar game with one half-lie to a problem of covering the  $q$ -dimensional hypercube with certain asymmetric sets called 1-sets. The asymptotic lower bound for  $A_1^*(q)$  follows from a sphere bound argument in Section 4. An inductive constructive upper bound is given in Section 5, for which we generate a base case in Section 3 by computing  $A_1^*(q)$  for small values of  $q$ .

## 2 Equivalence to covering by 1-sets

In this section, we prove the equivalence of the existence of a strategy for Paul to win the  $q$ -round pathological liar game on  $n$  elements with one half-lie to the existence of a cover of the  $q$ -cube with certain so-called 1-sets. Define  $Q_q := \{N, Y\}^q$ , to be the  $q$ -dimensional hypercube, where we think of 1’s being associated with N’s and 0’s with Y’s. A 1-set consists of

- a stem  $\omega = \omega_1\omega_2 \cdots \omega_q \in Q_q$  with  $r$  N’s at positions  $j_1, \dots, j_r$ , where  $0 \leq r \leq q$ , and
- exactly  $r$  children  $C_1, \dots, C_r$  such that  $C_k \in \omega_1 \cdots \omega_{j_k-1} Y Q_{q-j_k}$ .

Thus a child is obtained from the stem by changing some “N” to “Y” and choosing the remaining sequence to the right arbitrarily. For example,  $\{NNY, YYN, NY Y\}$  is a 1-set with stem  $NNY$ . A collection  $\mathcal{C}$  of 1-sets is a *covering* of  $Q_q$  provided  $\cup_{S \in \mathcal{C}} S = Q_q$ .

The proof of the following theorem is similar to that of Theorem 2.4 in [8], but is adjusted for the pathological half-liar game and specialized to the case of one half-lie.

**Theorem 2.** Let  $q \geq 0$ . Paul has a strategy for winning the  $q$ -round pathological liar game with one half-lie on  $n$  elements if and only if there exists a covering of  $Q_q$  with  $n$  1-sets.

*Proof.* For the forward implication, Paul’s winning strategy corresponds to a decision tree which is a full binary tree of depth  $q$  and has a question at each non-leaf node. The left child of a node

corresponds to a response of “No” by Carole, while the right corresponds to “Yes”. A game played under this strategy is a path from the root to a leaf of the decision tree, passing down  $q$  levels of questions by Paul and answers by Carole. The leaves of the decision tree correspond to the vertices of  $Q_q$ , and a leaf is labeled with the surviving elements in the game determined by the path from the root to that leaf. Because the strategy is a winning strategy, every leaf is associated with at least one element of  $[n]$ . Now let  $P_i$  be the set of paths from the root to any leaf associated with  $i$ , and label each path with Carole’s corresponding response sequence. Exactly one path  $\omega \in P_i$  corresponds to all truthful responses by Carole; this  $\omega$  is the candidate stem for  $P_i$  as a 1-set. Every other  $\omega' \in P_i$  must agree with  $\omega$  up until some position in which  $\omega$  has a “N” and  $\omega'$  has a “Y”, after which there is a unique way to complete  $\omega'$  so that  $i$  is associated with only one lie. Every position in which  $\omega$  has a “N” results in an  $\omega'$  of this form, and so  $P_i$  is a 1-set. Therefore the collection  $\mathcal{C} = \{P_i : i \in [n]\}$  is a cover of  $Q_q$  by  $n$  1-sets.

For the reverse implication, let  $\mathcal{C} = \{P_i : i \in [n]\}$  be a collection of  $n$  1-sets which covers  $Q_q$ . Label the leaf corresponding to the path  $\omega \in \{Y, N\}^q$  with  $i$  if  $\omega \in P_i$ , whence each of the  $2^q$  leaves has at least one label. At each node  $z$  of the decision tree define  $L(z)$  to be the set of labels of leaves of the left subtree of  $z$ , and associate with  $z$  the question “Is  $x \in [n] \setminus L(z)$ ?”. We claim these questions comprise a winning strategy for Paul.

Consider an arbitrary response sequence  $\omega$  by Carole, and assume to the contrary that  $\omega$  associates all elements labeling the leaf corresponding to  $\omega$  with 2 or more lies. The leaf corresponding to  $w$  is labeled by at least one element of  $[n]$ ; let  $i$  be such a label, so that  $w \in P_i$ . By assumption  $i$  is associated with at least two lies; let  $1 \leq s < t \leq q$  be the positions of these lies in Carole’s response sequence. Since Carole may only lie by responding “Y” when the truthful answer is “N”, There must exist  $a, b \in P_i$  of the following form:

$$\begin{aligned} a &= a_1 \cdots a_{s-1} N \cdots \\ b &= a_1 \cdots a_{s-1} Y b_{s+1} \cdots b_{t-1} N \cdots \\ w &= a_1 \cdots a_{s-1} Y b_{s+1} \cdots b_{t-1} Y \cdots \end{aligned}$$

Clearly, it is impossible for  $P_i$  to be a 1-set if any of  $a$ ,  $b$ , or  $w$  is the stem. Neither can there be a distinct stem  $c \in P_i$ , since  $c$  would have  $a_1 \cdots a_{s-1} N$  as a prefix, and could not support the two children  $a$  and  $b$  both obtained by changing position  $s$  from “N” to “Y”. This contradicts the assumption that  $P_i$  is a 1-set, and so  $i$  must be associated with at most one lie. Therefore Paul wins the game with response sequence  $\omega$  because  $i$  is not disqualified.  $\square$

In fact, a more general version of Theorem 2 is true. In [8], an  $i$ -set was defined as the set of all possible response sequences by Carole about an element which contain at most  $i$  lies. It is straightforward to extend the proof of Theorem 2 to show that Paul has a strategy to win the pathological game with  $k$  half-lies in  $q$  rounds starting in position  $(x_0, \dots, x_k)$  if and only if there exists a simultaneous covering of  $Q_q$  with  $x_i$   $(k - i)$ -sets, where  $0 \leq i \leq k$ .

Related to a 1-set with stem  $w$  is the *asymmetric ball* with radius 1 centered at  $w$ , which we denote by  $B^-(w, 1)$  (cf. [6]).  $B^-(w, 1)$  is defined to be the set of all  $w' \in \{Y, N\}^q$  obtained from  $w$  by changing at most one “N” to “Y” and leaving the remaining positions fixed. Thus such an asymmetric ball centered at  $w$  is always a 1-set with stem  $w$ , but not vice versa. We note without proof that a cover of  $Q_q$  by  $n$  asymmetric balls corresponds to a non-adaptive strategy for Paul to win the  $q$ -round pathological liar game on  $n$  elements with one half-lie.

### 3 $A_1^*(q)$ for small values of $q$

Using the vector game formulation and the definition of  $r^*((n, 0))$  in (3), a moderately optimized dynamic programming algorithm computes  $A_1^*(q)$  exactly for small values of  $q$ . The corresponding critical value for minimum coverings of  $Q_q$  by asymmetric balls is  $K^+(q, 1)$ , defined to be the minimum number of asymmetric balls of radius 1 needed to cover  $Q_q$ . Table 1 presents  $A_1^*(q)$  versus  $K^+(q, 1)$  for  $1 \leq q \leq 13$ ; the data on  $K^+(q, 1)$  is due to [6, 2, 10]. The current bound on  $K^+(14, 1)$  is very poor and thus is omitted. The dynamic programming algorithm for  $A_1^*(q)$  is polynomial in  $n$  but exponential in  $q$ ; further optimization or greater computing resources should be able to determine  $A_1^*(q)$  up to at least  $q = 20$ . Determination of more than a few additional values of  $K^+(q, 1)$ , on the other hand, seems very difficult. A key question is whether the ratio  $K^+(q, 1)/A_1^*(q)$  converges to 1 as  $q \rightarrow \infty$ , or if there is an absolute constant  $C > 1$  such that  $\limsup_{q \rightarrow \infty} K^+(q, 1)/A_1^*(q) \geq C$ .

$q$	1	2	3	4	5	6	7
$A_1^*(q)$	1	2	3	5	8	13	22
$K^+(q, 1)$	1	2	3	6	10	18	31

$q$	8	9	10	11	12	13	14
$A_1^*(q)$	39	71	129	239	444	831	1561
$K^+(q, 1)$	58	93–106	162–229	306–433	563–668	1046–1285	—

Table 1: Values of  $A_1^*(q)$ , the minimum number of elements  $n$  for which Paul can win the  $q$ -round pathological liar game with one half-lie, versus  $K^+(q, 1)$ , the minimum number of asymmetric balls of radius 1 needed to cover  $Q_q$

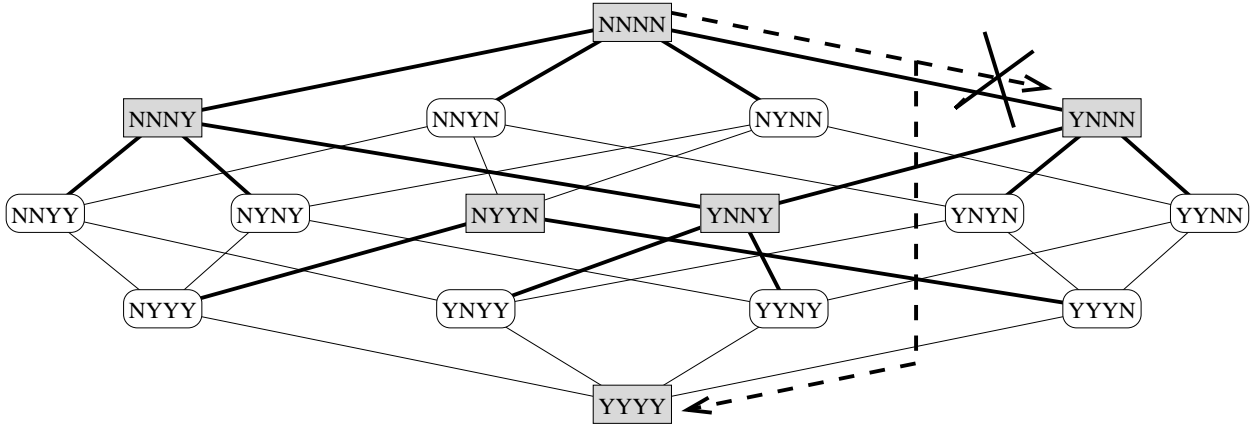


Figure 1: The hypercube  $Q_4$  has an asymmetric cover consisting of 6 asymmetric balls of radius 1, centered at the gray rectangular vertices; children of ball centers are connected upward to their respective parents by bold lines. In relaxing to a cover by 1-sets, the child of NNNN obtained from changing the first “N” to “Y” can be taken as YYYY instead of YNNN, yielding an optimal 1-set cover of size 5.

Because an asymmetric ball of radius 1 in  $Q_q$  is also a 1-set, covering  $Q_q$  with 1-sets is a relax-

ation of covering  $Q_q$  with asymmetric 1-balls. Therefore for all  $q$  we have  $A_1^*(q) \leq K^+(q, 1)$ . Figure 1 illustrates the first disagreement occurring at  $q = 4$  between a minimum cover of asymmetric 1-balls of size 6 centered at NNNN, NNNY, YNNN, NYYN, YNNY, and YYYY; and a minimum 1-set cover of size 5 (listed with stems first):

$$\begin{aligned} & \{\{NNNN, YYYY, NYNN, NNYN, NNNY\}, \{NNNY, YNNY, NYNY, NNYY\}, \\ & \{YNNN, YYNN, YNYN, YNNY\}, \{NYYN, YYYN, NYYY\}, \{YNNY, YYNY, YNYY\}\}. \end{aligned}$$

The two covers are nearly identical, except that in the first 1-set, YYYY is taken to be the child of NNNN with respect to the first coordinate, rather than YNNN in the first asymmetric 1-ball. Thus the extra flexibility in the definition of a 1-set allows a smaller cover. The value  $A_1^*(7)$  will be used in Section 5 as the base case of an inductive construction for an upper bound on  $A_1^*(q)$ .

## 4 An asymptotic sphere lower bound

A *sphere bound* for the number of 1-sets needed to cover  $Q_q$  is obtained by dividing the size of the space to be covered,  $2^q$ , by the size of a typical 1-set. Since asymptotically most vertices of  $Q_q$  are concentrated in the middle with approximately  $q/2$  N's, we expect the stem of a typical 1-set to have about  $q/2$  N's. However, a covering by 1-sets might use stems with a large number of N's more than once; we negotiate this concern by arguing about the maximum number of typical vertices with approximately  $q/2$  N's which can be contained in a 1-set, eventually confirming that an asymptotic sphere bound on  $A_1^*(q)$  of  $\frac{2^{q+1}}{q}(1 + o(1))$  is appropriate.

Define a vertex  $w$  in  $Q_q$  to be *rare* if it contains more than  $(q + \sqrt{q \ln q})/2$  N's. The Chernoff inequality gives that if  $X_1, \dots, X_q$  are independently distributed random variables with  $\Pr[X_i = 0] = \Pr[X_i = 1] = 1/2$ , then for any  $\alpha > 0$ ,  $\Pr[\sum_{i=1}^q X_i > (q + \alpha)/2] < e^{-\alpha^2/2q}$  (see, for example, Theorem A.1.1 of [1]). Choosing  $\alpha = k\sqrt{q \ln q}$  bounds the number of rare vertices in  $Q_q$  from above with  $2^q q^{-k^2/2}$ . By taking  $k = 1$ , we have proved the following.

**Lemma 3.** For  $q \geq 1$ , the number of non-rare vertices of  $Q_q$  is at least  $2^q(1 - q^{-1/2})$ .

We now count the largest number of non-rare vertices in a 1-set to bound  $A_1^*(q)$ .

**Theorem 4.** For  $q \geq 1$ ,

$$A_1^*(q) \geq \frac{2^{q+1}}{q} \left( 1 - O \left( \sqrt{\frac{\ln q}{q}} \right) \right).$$

*Proof.* The stem  $w$  of a 1-set is either non-rare or rare. If  $w$  is non-rare, then it has at most  $(q + \sqrt{q \ln q})/2$  N's, and so the total number of non-rare vertices in the 1-set cannot exceed  $1 + (q + \sqrt{q \ln q})/2$ , which is the size of the 1-set. Now suppose  $w = w_1 \cdots w_q$  is rare, and let  $r$  be the number of N's in  $w$ , so that  $w_{j_1} = \cdots = w_{j_r} = N$ . Each of  $w$ 's children corresponding to positions  $j_{(q+\sqrt{q \ln q})/2+2}, \dots, j_r$  are rare, since the  $1 + (q + \sqrt{q \ln q})/2$  N's to the left remain unchanged. Therefore there are at most  $1 + (q + \sqrt{q \ln q})/2$  non-rare vertices among the remaining children.

In either case, a 1-set has at most  $1 + (q + \sqrt{q \ln q})/2$  non-rare vertices, and so by Lemma 3, the number of 1-sets required for a cover of  $Q_q$  is at least

$$\frac{2^q(1 - q^{-1/2})}{1 + (q + \sqrt{q \ln q})/2} = \frac{2^{q+1}}{q} \left( 1 - O \left( \sqrt{\frac{\ln q}{q}} \right) \right).$$

The result follows by Theorem 2. □

For general coding theory problems, the sphere bound is obtained by dividing the size of the space by the size of a typical set used to pack or cover the space; therefore the sphere bound is an upper bound for the packing problem and a lower bound for the covering problem. For instance, the sphere bound for symmetric binary covering and packing codes of size  $n$  and radius  $R$  is  $2^n / \sum_{i=0}^R \binom{n}{i}$  (see, for example, Theorem 6.1.2 of [5]). A major outstanding question for symmetric codes is whether covering and packing codes meet at the sphere bound for radius  $R > 1$ . A definitive, exact sphere bound unifying asymmetric binary covering and packing codes has not been settled, but an exact sphere bound for the asymmetric covering problem appears in Theorem 2 of [6], and corresponds to a candidate unifying asymptotic sphere bound of  $2^n / \binom{n}{R}$  in [12]. It is completely open for all  $R > 0$  whether optimal asymmetric codes meet at this candidate asymptotic sphere bound. Therefore the asymptotic sphere bound of  $\frac{2^{q+1}}{q}(1 + o(1))$  in Theorem 4 is important because it is met by both packing and covering by 1-sets, which are relaxations of radius  $R = 1$  asymmetric binary packing and covering, respectively. The case of Ulam's pathological liar game and the corresponding asymptotic sphere bound for  $k$  lies will be studied in [9].

## 5 A constructive upper bound

The asymptotic upper bound for  $A_1^*(q)$  is constructive, and arises from taking a known efficient cover for  $Q_q$  of sufficiently small size and building it into an efficient cover for  $Q_{q+1}$ .

**Theorem 5.** For all  $q \geq 1$ ,

$$A_1^*(q+1) \leq 2A_1^*(q) - \left\lfloor \frac{A_1^*(q)^2}{2^q} \right\rfloor$$

*Proof.* Let  $\mathcal{C} = \{S_1, \dots, S_{A_1^*(q)}\}$  be a collection of 1-sets of size  $A_1^*(q)$  which cover  $Q_q$ . Because vertices of  $Q_q$  may lie in overlapping 1-sets, construct the *reduced* collection  $\bar{\mathcal{C}} = \{\bar{S}_1, \dots, \bar{S}_{A_1^*(q)}\}$  of subsets of 1-sets by defining  $\bar{S}_i = S_i \setminus \cup_{j < i} S_j$ . In particular,  $\bar{\mathcal{C}}$  partitions  $Q_q$ .

Define  $N\bar{\mathcal{C}} := \{N\omega : \omega \in \bar{S}\} : \bar{S} \in \bar{\mathcal{C}}\}$  and  $Y\bar{\mathcal{C}} := \{Y\omega : \omega \in \bar{S}\} : \bar{S} \in \bar{\mathcal{C}}\}$ . Certainly,  $N\bar{\mathcal{C}} \cup Y\bar{\mathcal{C}}$  is a size  $2A_1^*(q)$  cover for  $Q_{q+1}$  which is extendible to a 1-set cover. But in extending  $N\bar{\mathcal{C}}$  to be a collection of 1-sets, each member  $N\bar{S} \in N\bar{\mathcal{C}}$  covers an arbitrary vertex in  $YQ_q$  by changing the first "N" to "Y" and arbitrarily choosing the rest;  $A_1^*(q)$  arbitrary vertices in  $YQ_q$  may be covered in this fashion. Therefore we may still extend  $N\bar{\mathcal{C}} \cup Y\bar{\mathcal{C}}$  to a 1-set cover of  $Q_{q+1}$  when some members of  $Y\bar{\mathcal{C}}$  are removed so that no more than  $A_1^*(q)$  vertices of  $YQ_q$  are left uncovered by the remaining members of  $Y\bar{\mathcal{C}}$ .

The average size of a member of  $Y\bar{\mathcal{C}}$  is  $2^q/A_1^*(q)$ , and so there must exist  $\lfloor A_1^*(q)^2/2^q \rfloor$  reduced 1-sets in  $Y\bar{\mathcal{C}}$  which cover at most  $\lfloor A_1^*(q)^2/2^q \rfloor \cdot (2^q/A_1^*(q)) \leq A_1^*(q)$  vertices of  $YQ_q$ . Now construct a 1-set cover for  $Q_{q+1}$  of size  $2A_1^*(q) - \lfloor A_1^*(q)^2/2^q \rfloor$  by removing these members of  $Y\bar{\mathcal{C}}$ , covering the at most  $A_1^*(q)$  exposed vertices in  $YQ_q$  by members of  $N\bar{\mathcal{C}}$ , and extending all sets arbitrarily to be 1-sets.  $\square$

An analysis of the recurrence of Theorem 5 yields the following upper bound on  $A_1^*(q)$ .

**Corollary 6.** For all  $q \geq 7$ ,

$$A_1^*(q) < \frac{2^{q+1}}{q} \left(1 - \frac{2}{q}\right).$$

*Proof.* Define  $f(q)$  by  $f(7) = A_1^*(7)$ , and  $f(q+1) = 2f(q) - f^2(q)/2^q + 1$  for  $q \geq 7$ , so that  $f(q)$  is an upper bound for  $A_1^*(q)$  for all  $q \geq 7$ . Define  $\alpha_q := f(q)/(2^{q+1}/q)$ . The recurrence for  $f$  written



in terms of  $\alpha_q$  is

$$\frac{\alpha_{q+1}}{q+1} = \frac{\alpha_q}{q} - \frac{\alpha_q^2}{q^2} + \frac{1}{2^{q+2}}.$$

The function  $y(x) = \frac{x}{q} - \frac{x^2}{q^2} + \frac{1}{2^{q+1}}$  increases on  $(-\infty, q/2]$ , where  $q/2 \geq 1$  when  $q \geq 2$ . Referring to Table 1, note that  $\alpha_7 = 77/128 < 1 - 2/7$ . For  $q \geq 7$ , if  $\alpha_q < 1 - \frac{2}{q}$ ,

$$\alpha_{q+1} = (q+1)y(\alpha_q) < (q+1)y(1 - 2/q) = (q+1) \left[ \frac{1}{q} \left(1 - \frac{2}{q}\right) - \frac{1}{q^2} \left(1 - \frac{2}{q}\right)^2 + \frac{1}{2^{q+2}} \right].$$

Denoting the right-hand side by  $X$ , we have

$$1 - \frac{2}{q+1} - X = \frac{q^3 - q^2 + 4q + 4}{(q+1)q^4} - \frac{q+1}{2^{q+2}},$$

which is at least 0 for  $q \geq 7$ . Thus by induction  $\alpha_q < 1 - 2/q$  for all  $q \geq 7$ .  $\square$

The closeness of this upper bound to the lower bound is due to flexibility in the choice of children for a given stem of a 1-set. The above analysis depends on knowing the exact value of  $A_1^*(7)$ , but could be rewritten using a near-optimal cover  $\mathcal{C}_q$  for some  $q > 7$ , provided  $|\mathcal{C}_q| \leq 1 - 2/q$ . Furthermore, the choice of  $c = 2$  in the bound  $\alpha_q < 1 - 2/q$  corresponds to the base case  $q = 7$ ; the induction goes through for a smaller but not arbitrary  $c > 0$  provided a sufficiently large base case can be found.

## Acknowledgement

The authors wish to express their appreciation for the careful review by the referee and the several useful comments which resulted.

## References

- [1] Noga Alon and Joel H. Spencer. *The probabilistic method*. John Wiley & Sons, New York, second edition, 2000.
- [2] David Applegate, E. M. Rains, and N. J. A. Sloane. On asymmetric coverings and covering numbers. *J. Combin. Des.*, 11(3):218–228, 2003.
- [3] E. R. Berlekamp. Block coding for the binary symmetric channel with noiseless, delayless feedback. In *Error Correcting Codes (Proc. Sympos. Math. Res. Center, Madison, Wis., 1968)*, pages 61–88. John Wiley, New York, 1968.
- [4] Ferdinando Cicalese and Daniele Mundici. Optimal coding with one asymmetric error: below the sphere packing bound. In *Computing and combinatorics (Sydney, 2000)*, volume 1858 of *Lecture Notes in Comput. Sci.*, pages 159–169. Springer, Berlin, 2000.
- [5] Gérard Cohen, Iiro Honkala, Simon Litsyn, and Antoine Lobstein. *Covering codes*, volume 54 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1997.
- [6] Joshua N. Cooper, Robert B. Ellis, and Andrew B. Kahng. Asymmetric binary covering codes. *J. Combin. Theory Ser. A*, 100(2):232–249, 2002.

- [7] Christian Deppe. Solution of Ulam’s searching game with three lies or an optimal adaptive strategy for binary three-error-correcting codes. *Discrete Math.*, 224(1-3):79–98, 2000.
- [8] Ioana Dumitriu and Joel Spencer. A halfliar’s game. *Theoret. Comput. Sci.* To appear.
- [9] Robert B. Ellis, Vadim Ponomarenko, and Catherine H. Yan. Ulam’s pathological liar game, in preparation.
- [10] Geoffrey Exoo. Private communication.
- [11] Wojciech Guzicki. Ulam’s searching game with two lies. *J. Combin. Theory Ser. A*, 54(1):1–19, 1990.
- [12] Michael Krivelevich, Benny Sudakov, and Van H. Vu. Covering codes with improved density. *IEEE Trans. Inform. Theory*, 49(7):1812–1815, 2003.
- [13] Andrzej Pelc. Solution of Ulam’s problem on searching with a lie. *J. Combin. Theory Ser. A*, 44(1):129–140, 1987.
- [14] Joel Spencer. Ulam’s searching game with a fixed number of lies. *Theoret. Comput. Sci.*, 95(2):307–321, 1992.
- [15] Joel Spencer and Catherine H. Yan. The halfie problem. *J. Combin. Theory Ser. A*, 103(1):69–89, 2003.