Secret sharing on the d-dimensional cube

László Csirmaz
Central European University

June 10, 2005

Abstract

We prove that for $d > 1$ the information rate of the perfect secret sharing scheme based on the edge set of the d-dimensional cube is exactly $2/d$. Using the technique developed, we also prove that the information rate of the infinite d-dimensional lattice is $1/d$.

Key words. Secret sharing scheme, polymatroid, information theory.

1 Introduction

Secret sharing scheme is a method of distributing a secret data among a set of participants so that only qualified subsets are able to recover the secret. If, in addition, unqualified subsets have no extra information, i.e. their joint shares is statistically independent of the secret, then the scheme is called *perfect*. The description of the qualified subsets among all possible subsets of participants is called *access structure*. The main question is the efficiency of the system: how many bits of information the participants must remember for each bit of the secret, either in the average, or in the worst case. The particular case when the access structure is defined by a graph with vertices as participants, and edges as minimal qualified subsets attracted considerable interest, see, e.g., [1, 2]. When the graph is the edge set of the d dimensional cube was considered in [2], giving upper and lower bounds for the number of bits to be remembered. Determining the exact value remained an open problem. Using a new technique we solve it. We extend the definition of information rate for infinite graphs, and investigate the case when the graph is the infinite d dimensional lattice.

The paper is organized as follows. In the next session we give the definitions necessary to state and prove our theorems. Section 3 deals with the case of the d-dimensional cube, section 4 with the lattice. Finally section 5 concludes the paper, and list some related problems. For undefined notions and for more introduction to secret sharing schemes see [1, 2], and for those in information theory consult [3].

2 Definitions

In this section we recall the notions we shall use later. First we give a formal definition of a perfect secret sharing scheme, then connect it to submodular functions.

Let $G = \langle V, E\rangle$ be a graph with vertex set V and edge set E. A subset A of V is *independent* if there is no edge between vertices in A. A *covering* of the graph G is a collection of subgraphs of G such that every edge is contained on one of the (not necessarily spanned) subgraphs in the collection. The collection is k*-covering* if every edge of G is covered exactly k times. For subsets of vertices we usually omit the $\cup$ sign, and write AB for $A \cup B$. Also, it $v \in V$ is a vertex then Av denotes $A \cup \{v\}$.

A *perfect secret sharing scheme* $\mathcal{S}$ for a graph G is a collection of random variables ξ_v for each $v \in V$ and a ξ (the secret) with a joint distribution so that

- if vw is an edge in G, then ξ_v and ξ_w together determine the value of ξ;
- if A is an independent set, then ξ and the collection $\{\xi_v : v \in A\}$ are statistically independent.

The *size* of the random variable ξ is measured by its entropy, or information content, and is denoted by $\mathbf{H}(\xi)$, see [3]. The *information ratio* for a vertex v (sometimes called participant) of the graph G is $\mathbf{H}(\xi_v)/\,\mathbf{H}(\xi)$. This value tells how many bits of information v must remember for each bit in the secret. The *worst case* and *average* information ratio of $\mathcal{S}$ are the highest and average information ratio among all participants, respectively.

Given a graph G its (worst case or average) information ratio is the infimum of the corresponding values for all perfect secret sharing schemes $\mathcal{S}$ defined on G. The widely used *information rate* and *average information rate* is the inverse of this value. While "information rate" is the customary measure in the literature, we found "information ratio" more intuitive and use it throughout this paper.

Let $\mathcal{S}$ be a perfect secret sharing scheme based on the graph G with the random variable ξ as secret, and ξ_v for $v \in V$ as shares. For each subset A of the vertices let us define

$$f(A) \stackrel{\text{def}}{=} \frac{\mathbf{H}(\{\xi_v : v \in A\})}{\mathbf{H}(\xi)}.$$

Clearly, the average information ratio of $\mathcal{S}$ is the average of $\{f(v) : v \in V\}$, and the worst case information ratio is the maximal value in this set. Using standard properties of the entropy function, cf. [3], we have

(a) $f(\emptyset) = 0$, and in general $f(A) \ge 0$ (positivity);

(b) if $A \subseteq B \subseteq V$ then $f(A) \le f(B)$ (monotonicity);

(c) $f(A) + f(B) \ge f(A \cap B) + f(A \cup B)$ (submodularity).

It is well known that for two random variables η and ξ, the value of η determines the value of ξ iff $\mathbf{H}(\eta\xi) = \mathbf{H}(\eta)$, and η and ξ are (statistically) independent iff $\mathbf{H}(\eta\xi) = \mathbf{H}(\eta) + \mathbf{H}(\xi)$. Using these facts and the definition of the perfect secret sharing scheme, we also have

(d) if $A \subseteq B$, A is an independent set and B is not, then $f(A) + 1 \le f(B)$ (strong monotonicity);

(e) if neither A nor B is independent but $A \cap B$ is so, then $f(A) + f(B) \ge 1 + f(A \cap B) + f(A \cup B)$ (strong submodularity).

The celebrated *entropy method*, see, e.g., [1], can be rephrased as follows. Prove that for *any* real-valued function f satisfying properties (a)–(e) the average (or largest) value of f on the vertices is at least ρ. Then, as functions coming from secret sharing schemes also satisfies these properties, conclude, that the average (or worst case) information ratio of G is also at least ρ. We note that this method is not necessarily universal, as properties (a)–(c) are too weak to capture exactly the functions coming from entropy. However, all existing lower bound proofs use the entropy method, and no example is known where the entropy method would not work.

We frequently use the submodular (c) and the strong submodular (e) properties in the following rearranged form whenever A, X, and Y are disjoint subsets of the vertex set V:

(c$'$) $f(AX) - f(A) \ge f(AXY) - f(AY)$;

moreover, if A is independent (i.e. empty), AX and AY are not, then

(e$'$) $f(AX) - f(A) \ge f(AXY) - f(AY) + 1$.

In particular, if both X and Y contain an edge (and are disjoint), then $f(X) \ge f(XY) - f(Y) + 1$.

The proof of the following easy fact is omitted:

Fact 2.1 *Suppose G_2 is a spanned subgraph of G_1. The worst case (average) information ratio of G_1 is at least as large as the worst case (average) information ratio of G_2.* ∎

3 The case of the cube

The *d-dimensional cube*, denoted here by C^d, is the following graph. Its vertices are 0–1 sequences of length d. Two vertices are connected by an edge if the sequences differ in exactly one place. This cube can be embedded into the d-dimensional Euclidean space. Points with all coordinates in the set $\{0, 1\}$ are the vertices, and two vertices are connected if their distance is 1.

The d-dimensional cube has 2^d vertices, $d \cdot 2^{d-1}$ edges (they correspond to one-dimensional affine subspaces in the embedding), and each vertex has degree d. The two-dimensional subspaces are squares, i.e. cycles of length four, we call them 2-*faces*. Each vertex v is adjacent to $\binom{d}{2}$ such 2-face, as any pair of edges starting from v spans a 2-face. Consequently the number of 2-faces is $2^{d-2}\binom{d}{2}$. For any edge there are exactly $(d-1)$ many 2-faces adjacent to that edge. It means that 2-faces, as subgraphs, constitute a $(d-1)$-cover of C^d.

Theorem 3.1 *The information ratio of the $d \ge 2$ dimensional cube is $d/2$.*

We note that this statement is not true for $d = 1$. The 1-dimensional "cube" is the graph with two vertices and an edge between them. In this graph both worst case and average information ratio is equal to 1, and not to 1/2. The 2-dimensional "cube" is the square, i.e. a cycle on four vertices, which is a complete bipartite graph. Thus both worst case and average information ratio of C^2 is 1, in full agreement with the statement of the theorem.

Proof First we prove that this ratio is at most $d/2$. To this end we construct a perfect secret sharing scheme witnessing this value. The construction uses Stinson's idea from [4].

Let F be a sufficiently large finite field, and X be the $(d-1)$-dimensional vector space over F. For every 2-face of the cube choose a vector $\mathbf{x}_i \in X$ in such a way that any $d-1$ of these vectors span the whole vector space X. (This is the point where we use the fact that F is sufficiently large.) The vectors $\mathbf{x}_i$ are public information, and the secret is a random element $\mathbf{s} \in X$. For each vector $\mathbf{x}_i$ take the inner product $a_i = \mathbf{s} \cdot \mathbf{x}_i$. Clearly, given any $(d-1)$ of these inner products, one can recover the secret $\mathbf{s}$. Now suppose the i-th 2-face has vertices v_1, v_2, v_3, v_4 in this order. Distribute a_i among these vertices as follows. Choose a random element $r \in F$ and give it to v_1 and v_3, and give $r + a_i$ (computed in the field F) to v_2 and v_4. Any edge of this 2-face can recover a_i, thus any edge of the d-dimensional cube can recover $d-1$ of the a_i's, and therefore can recover the secret $\mathbf{s}$ as well. Now consider the values an independent set of the vertices possess. All different values in this set can be chosen independently and randomly from F, thus they are (statistically) independent of the secret $\mathbf{s}$.

We have verified that this is a perfect secret sharing system. The secret is a $(d-1)$-tuple from the field F. Each vertex is given as many elements from F as many 2-faces it is in, namely $\binom{d}{2}$ elements. Therefore both worst case and average information ratio for this scheme is $\binom{d}{2}/(d-1) = d/2$, which proves the upper bound.

Before handling the lower bound, observe that the worst case and the average case information ratio for cubes coincide. The reason is that C^d is highly symmetrical. Let H be the automorphism group of the graph C^d, this group has $2^d \cdot d!$ elements. If v_1 and v_2 are two (not necessarily different) vertices of C^d, then the number of automorphisms $\pi \in H$ with $\pi(v_1) = v_2$ is exactly $|H|/|C^d| = d!$. Now let $\mathcal{S}$ be any perfect secret sharing scheme on C^d, and apply $\mathcal{S}$ for πC^d independently for each $\pi \in H$. The size of the secret in this compound scheme increases $|H|$-fold, and each participant will get a share which has size $|H|/|C^d|$-times the sum of all share sizes in $\mathcal{S}$. Therefore in this "symmetrized" scheme all participants have the same amount of information to remember, consequently all have the same ratio which equals to the average ratio of the scheme $\mathcal{S}$.

Thus to prove that $d/2$ is also a lower bound for both the worst case and average information ratio of C^d it is enough to show that for any real valued function f satisfying properties (a)–(e) enlisted in section 2 we have

$$\sum\{f(v) : v \in V\} \geq \frac{d}{2}.$$

This is exactly what we will do.

Split the vertex set of the d-dimensional cube C^d into two equal parts in a "chessboard-like" fashion: $C^d = A_d \cup B_d$, where A_d and B_d are disjoint, independent, and $|A_d| = |B_d| = 2^{d-1}$. Vertices in A_d have neighbors in B_d only, and vertices in B_d have neighbors in A_d only. The $(d+1)$-dimensional cube consist of two disjoint copy of the d dimensional cube at two levels, and there is a perfect matching between the corresponding vertices. Each edge of C^{d+1} is either a vertex of one of the lower dimensional cubes, or is a member of the perfect matching. Suppose the vertices on these two smaller cubes are split as $A_d \cup B_d$ and $A'_d \cup B'_d$, respectively, such that the perfect matching is between A_d and B'_d, and between B_d and A'_d. Then the splitting of the vertices of the $(d+1)$-dimensional cube can be done as

$$A_{d+1} = A_d \cup A'_d \quad \text{and} \quad B_{d+1} = B_d \cup B'_d.$$

Using this decomposition, we can use induction on the dimension d. In the inductive statement we shall use the following notation:

$$[\![A, B]\!] \stackrel{\text{def}}{=} \sum_{b\in B} f(bA) - \sum_{a\in A} f(A - \{a\}).$$

When using this notation we implicitly assume that A and B have the same cardinality.

Lemma 3.2 *For the d-dimensional cube with the split $C^d = A_d \cup B_d$ we have*

$$\sum_{v\in C^d} f(v) \geq [\![A_d, B_d]\!] + (d-1)2^{d-1}. \tag{1}$$

Proof First check this inequality for $d = 1$. The 1-cube has two connected vertices a and b. Then, say, $A_1 = \{a\}$, $B_1 = \{b\}$, and equation (1) becomes

$$f(a) + f(b) \geq f(ab) - f(\emptyset) + 0,$$

which holds by the submodular property (c) of the function f.

Now suppose (1) holds for both d-dimensional subcubes of the $(d+1)$-dimensional cube with split $A_{d+1} = A_d \cup A'_d$, and $B_{d+1} = B_d \cup B'_d$ as discussed above. Then by the inductive hypothesis,

$$\begin{aligned} \sum_{v\in V_{d+1}} f(v) &= \sum_{v\in V_d} f(v) + \sum_{v'\in V'_d} f(v') \\ &\geq [\![A_d, B_d]\!] + [\![A'_d, B'_d]\!] + (d-1)2^d. \end{aligned} \tag{2}$$

Each $b \in B_d$ is connected to a unique $a' \in A'_d$, let (a', b) be such a pair. Then

$$f(bA_d) - f(A_d) \geq f(bA_dA'_d - \{a'\}) - f(A_dA'_d - \{a'\}) \tag{3}$$

by submodularity. Now let $a \in A_d$ any vertex which is connected to $b \in B_d$. As b is connected to both a and a', both bA'_d and $abA'_d - \{a'\}$ are qualified (i.e. not independent) subsets, while their intersection, $bA'_d - \{a'\}$, is independent. Therefore the strong submodularity yields

$$f(bA'_d) - f(bA'_d - \{a'\}) \geq 1 + f(baA'_d) - f(baA'_d - \{a'\}).$$

Using this inequality and the submodularity twice we get

$$\begin{aligned} f(A'_d) - f(A'_d - \{a'\}) &\geq f(bA'_d) - f(bA'_d - \{a'\}) \\ &\geq 1 + f(baA'_d) - f(baA'_d - \{a'\}) \\ &\geq 1 + f(bA_dA'_d) - f(bA_dA'_d - \{a'\}). \end{aligned}$$

Adding (3) to this inequality, for each connected pair (a', b) from $a' \in A'_d$ and $b \in B_d$ we have

$$f(bA_d) - f(A_d) + f(A'_d) - f(A'_d - \{a'\}) \geq 1 + f(bA_dA'_d) - f(A_dA'_d - \{a'\}).$$

By analogy we can swap (A_d, B_d) and (A'_d, B'_d) yielding

$$f(b'A'_d) - f(A'_d) + f(A_d) - f(A_d - \{a\}) \geq 1 + f(b'A_dA'_d) - f(A_dA'_d - \{a\})$$

for each connected pair (a, b') from $a \in A_d$ and $b' \in B'_d$. There are 2^{d-1} edges between A'_d and B_d, and also 2^{d-1} edges between A_d and B'_d. Thus adding up all of these 2^d inequalities, on the left hand side all $f(A_d)$ and $f(A'_d)$ cancel out, and the remaining terms give

$$[\![A_d, B_d]\!] + [\![A'_d, B'_d]\!] \geq [\![A_dA'_d, B_dB'_d]\!] + 2^d.$$

Combining this with (2) we arrive at

$$\sum_{v \in V_{d+1}} f(v) \geq [\![A_dA'_d, B_dB'_d]\!] + (d-1)2^d + 2^d,$$

which is exactly inequality (1) for $d+1$, which was to be proved. ∎

We continue with the proof of theorem 3.1. Let $C^d = A_d \cup B_d$ be the disjoint "chessboard" splitting of the vertices. As there are exactly 2^{d-1} vertices in both A_d and B_d, we can match them. If (a, b) is such a matched pair, then by strong monotonicity

$$f(bA_d) - f(A_d - \{a\}) \geq 1,$$

as $A_d - \{a\}$ is independent, while bA_d is not. Adding up these inequalities we get

$$[\![A_d, B_d]\!] = \sum_{b \in B_d} f(bA_d) - \sum_{a \in A_d} f(A_d - \{a\}) \geq 2^{d-1}.$$

This, together with the claim of Lemma 3.2 gives

$$\sum_{v \in V_d} f(v) \geq (d-1)2^{d-1} + 2^{d-1} = d2^{d-1}.$$

There are 2^d vertices in V_d, thus the average value of f on the vertices of V_d is at least $d/2$. This shows that the average information ratio of the d-dimensional cube is at least $d/2$. From this it follows, as has been explained before, that the worst case information ratio is also at least $d/2$. ∎

4 The case of the lattice

The vertices of the *d-dimensional lattice* L^d are the integer points of the d-dimensional Euclidean space, i.e. points having integer coordinates only. Two vertices are connected if their distance is exactly 1, i.e. if they differ in a single coordinate, and the difference in that coordinate is exactly 1. Of course, L^d is an infinite graph.

Each vertex in L^d has degree $2d$, and the whole graph is *edge transitive.* Namely, given any two edges v_1v_2 and w_1w_2 from L^d, there is an automorphism of L^d which maps v_1 to w_1 and v_2 to w_2.

Defining information ratio for an infinite graph is not a straightforward task. As each non-isolated vertex must remember at least as much information as the secret contains (cf. [4]), the total amount of information to be distributed is infinite. We need to circumvent this infinity some way.

Definition 4.1 The *information ratio* of an infinite graph G is the supremum of the information ratio of all finite spanned subgraphs of G.

By Fact 2.1 the information ratio of a finite graph G is at least as large as that of its spanned subgraphs. Thus for finite graphs the above definition gives back the original value. This would not be the case if not only spanned subgraphs were considered.

Several well-known theorems generalize easily to the infinite case. Here we mention only two:

Claim 4.2 *The information ratio of a complete multipartite graph is* 1. ∎

Claim 4.3 *If all vertices of the graph have degree $\leq d$, then the worst case information ratio is at most* $(d+1)/2$. ∎

As each vertex of the d dimensional lattice has degree $2d$, by Claim 4.3 the information ratio of L^d is at most $(2d+1)/2$. When $d=1$ then actually this is the truth, namely the information rate of L^1 is $3/2$. The 1-dimensional lattice is the infinite path. The upper bound $3/2$ comes from Claim 4.3, and the lower bound comes from the following

Fact 4.4 ([1, 2, 4]) *If the path abcd is a spanned subgraph, then* $f(b)+f(c)\geq 3$. ∎

Any two consecutive vertices along the (infinite) path have average (and worst case) information ratio $\geq 3/2$, and then the same bound applies to the whole L^1.

For $d\geq 2$ the information ratio for L^d is smaller than the bound $(2d+1)/2$ given by Claim 4.3. In fact,

Theorem 4.5 *For $d\geq 2$ the information ratio of the d dimensional lattice L^d is d.*

Proof First we show that d is an upper bound. This requires a construction of a perfect secret sharing scheme in which every vertex should remember at most d times as much information as there is in the secret. Let v be a vertex of L^d whose all coordinates have the same parity – i.e. either all are odd or all are even integers. Increase each coordinate of v either by 0 or 1. The resulting 2^d points form a d-dimensional cube. Consider all of these cubes. They fill out the whole space in a chessboard-like fashion. Each vertex of L^d belongs to exactly two such cubes: one starting form a point with even coordinates only, and one starting from a point with odd coordinates only. Furthemore each edge of L^d belongs to exactly one of these cubes.

Distribute the secret in each of these (infinitely many) cubes independently. By Theorem 3.1 this can be done so that each vertex of the cube gets exactly $d/2$ bits for each bit in the secret. As each vertex in L^d is in exactly two cubes, each vertex gets two times $d/2$ bits. And as each vertex of L^d is a vertex in some cube, endpoints of a vertex can recover the secret.

The distribution of the shares in each cube was made by a perfect system, and random values were chosen independently for each cube. Therefore independent subsets of L^d have no information on the secret. This proves that d is an upper bound for both the average and worst case information ratio.

Proving that d is also a lower bound first we prove a generalizetion of Lemma 3.2. To describe the setting, suppose we have a graph with vertices split into six disjoint sets $(A \cup A^*) \cup (B \cup B^*) \cup (A' \cup B')$. Subsets $A \cup A^* \cup A'$ and $B \cup B^* \cup B'$ are independent, cardinalities of the subsets A, A', B, and B' are equal, furthermore $|A^*| = |B^*|$. Edges of the graph go between $A \cup A^*$ and $B \cup B^*$, between A' and B', moreover there is a perfect matching between A' and B, and there is a perfect matching between A and B'. This means, for example, that each $a' \in A'$ is connected to exactly one member of B, and there is no edge, for example, between B' and A^*.

Lemma 4.6 *With the notations above, let $|A| = |B| = |A'| = |B'| = k$. Suppose moreover that each $b \in B$ is connected to some $a \in A \cup A^*$, and each $b' \in B'$ is connected to some $a' \in A'$. Then*

$$[\![AA^*, BB^*]\!] + [\![A', B']\!] \geq 2k + [\![A'AA^*, B'BB^*]\!].$$

Proof As in the proof of Lemma 3.2, for $b \in B$ let $a' \in A'$ be the only vertex it is connected to in A', and let $a \in A \cup A^*$ which b is connected to as well. Then using submodularity and strong submodularity,

$$f(bAA^*) - f(AA^*) \geq f(bAA^*A' - \{a'\}) - f(AA^*A' - \{a'\}),$$

and

$$\begin{aligned} f(A') - f(A' - \{a'\}) &\geq f(bA') - f(bA' - \{a'\}) \\ &\geq 1 + f(baA') - f(baA' - \{a'\}) \\ &\geq 1 + f(bAA^*A') - f(bAA^*A' - \{a'\}) \end{aligned}$$

On the other hand, if $b' \in B'$ is connected to $a \in A$, and $a' \in A'$, then

$$f(b'A') - f(A') \geq f(b'A'A^*A - \{a\}) - f(A'A^*A - \{a\}),$$

and

$$\begin{aligned} f(AA^*) - f(AA^* - \{a\}) &\geq f(b'AA^*) - (b'AA^* - \{a\}) \\ &\geq 1 + f(b'a'AA^*) - f(b'a'AA^* - \{a\}) \\ &\geq 1 + f(b'A'AA^*) - f(b'A'AA^* - \{a\}) \end{aligned}$$

Summing up all of these inequalities, $2k$ in total, $f(AA^*)$ and $f(A')$ are canceled out, and we get

$$\begin{aligned} &\Big(\sum_{b \in B} f(bAA^*) - \sum_{a \in A} f(AA^* - \{a\}) \Big) + \Big(\sum_{b' \in B'} f(b'A') - \sum_{a' \in A'} f(A' - \{a'\}) \Big) \\ &\quad \geq 2k + \sum_{b \in B \cup B'} f(bAA^*A') - \sum_{a \in A \cup A'} f(AA^*A' - \{a\}). \end{aligned}$$

The missing part, namely that

$$\sum_{b \in B^*} f(bAA^*) - \sum_{a \in A^*} f(AA^* - \{a\}) \geq \sum_{b \in B^*} f(bAA^*A') - \sum_{a \in A^*} f(AA^*A' - \{a\})$$

follows immediately from submodularity and from $|A^*| = |B^*|$. ∎

As we will use Lemma 4.6 inductively, we need to consider the base case first, namely the case when the dimension is 1. The 1-dimensional lattice is an infinite path; we handle its finite counterparts. Thus let $k \geq 2$ be an even number, and let a_1, b_1, ..., $a_{k/2}$, $b_{k/2}$ be the vertices, in this order, of a path of length k. Let A be the set of odd vertices, and B be the set of even vertices.

Lemma 4.7 *For each path P of even length $k \geq 2$,*

$$\sum_{v \in P} f(v) \geq [\![A, B]\!] + \frac{k}{2} - 1. \tag{4}$$

Proof By induction on the length of the path. When $k = 2$, i.e. the graph consists of two connected vertices a and b only, then by submodularity

$$f(a) + f(b) \geq f(ab) = [\![\{a\}, \{b\}]\!],$$

which is just the statement of the lemma.

Now let the first two vertices on the path be a' and b' (in this order), and let A^* be the set of odd vertices except for a', and B^* be the set of even vertices except for b'. Add two extra vertices, a'', and b'' to beginning of the path. The lemma follows by induction on the length of the path if we show that

$$f(a'') + f(b'') + [\![A^*a', B^*b']\!] \geq 1 + [\![A^*a'a'', B^*b'b'']\!].$$

Now $f(a'') + f(b'') \geq f(a''b'')$, and by submodularity

$$\sum_{b \in B^*} f(ba'A^*) - \sum_{a \in A^*} f(a'A^* - \{a\}) \geq \sum_{b \in B^*} f(ba'a''A^*) - \sum_{a \in A^*} f(a'a''A^* - \{a\}),$$

thus it is enough to show that

$$f(a''b'')+f(b'a'A^*)-f(A^*) \ge 1+f(b'a'a''A^*)+f(b''a'a''A^*)-f(a'A^*)-f(a''A^*).$$

But this is just the sum of the following three submodular inequalities:

$$\begin{aligned} f(a''b'')-f(b'') &\ge 1+f(b''a'a''A^*)-f(b''a'A^*) \\ f(b'') &\ge f(b''a'A^*)-f(a'A^*) \\ f(b'a'A^*)-f(A^*) &\ge f(b'a'a''A^*)-f(a''A^*); \end{aligned}$$

the first inequality holds as both $a''b''$ and $b''a'$ are edges in the graph. ∎

Now let k be an even number, and let L_k^d be the spanned subgraph of the the d-dimensional lattice L^d where only vertices with all coordinates between 0 and k inclusive are considered. Thus, for example L_2^d is just the d-dimensional cube with two vertices along each dimension. As L_k^d is a spanned subgraph of L_ℓ^d whenever $k \le \ell$, the average information ratio of L_k^d (not necessarily strictly) increases with k. Observe also that every finite spanned subgraph of L^d is isomorphic to a spanned subgraph of L_k^d for every large enough k. Thus the average information ratio of L^d is the limit of the average information ratio of L_k^d as k tends to infinity. In the sequel we estimate this latter value.

As in the proof of Theorem 3.1, split the vertices of L_k^d into two disjoint sets A_k^d and B_k^d in a "chessboard-like" fashion so that both sets are independent, and contain just half of the vertices: $|A_k^d| = |B_k^d| = k^d/2$.

Lemma 4.8 *With the notation as above,*

$$\sum_{v\in L_k^d} f(v) \ge [\![A_k^d, B_k^d]\!] + d(k^d - k^{d-1}) - \frac{k^d}{2}.$$

Proof For $d=1$ this is the claim of lemma 4.7. For larger dimensions we use induction on d. The $(d+1)$-dimensional lattice L_k^{d+1} consist of just k levels of L_k^d with a perfect matching between the levels. Thus we can apply lemma 4.6 $(k-1)$ times, each application increases the constant by the number of vertices on the new level, i.e. by k^d. Thus the constant for $(d+1)$ is k times the constant for d, plus $(k-1)$ times k^d. From here an easy calculation finishes the proof. ∎

Theorem 4.9 *The average information ratio of the d dimensional lattice of edge length k is at least $d(1-1/k)$.*

Proof Using the notations of lemma 4.8, observe that $[\![A_k^d, B_k^d]\!]$ can be written as the sum of $k^d/2$ differences. Each of these differences have value ≥ 1 by the strong monotonicity, since the first subset contains an edge, while the second one is independent. Thus $[\![A_k^d, B_k^d]\!] \ge k^d/2$. Using this, lemma 4.8 gives

$$\sum_{v\in L_k^d} f(v) \ge d(k^d - k^{d-1}).$$

As there are k^d vertices in L_k^d, the claim of the theorem follows. ∎

Setting $k = 2$ here, we get, as a special case, that the average information ratio of the d-dimensional cube is at least $d/2$. This was the hard part of Theorem 3.1.

Now we can finish the proof of Theorem 4.5. We have seen that d is an upper bound for the worst case information ratio of the d-dimensional lattice L^d. In Theorem 4.9 we gave the lower bound $(d - d/k)$ for the graph L_k^d, which can be embedded as a spanned subgraph into L^d. Thus the average information ratio of L^d is larger than, or equal to, the supremum of $(d - d/k)$ as k runs over the even integers. Thus $d \leq$ average information ratio of $L^d \leq$ worst case information ratio $\leq d$, which proves the theorem. ∎

5 Conclusion

Determining the exact amount of information a participant must remember in a perfect secret sharing scheme is an important problem both from theoretical and practical point of view. Access structures based on graphs pose special challenges. They are easier to define, have a transparent, and sometimes trivial, structure. Thus showing that certain graph-based structures require large shares might have serious consequences. Research along this line was initiated in [2], where several questions were asked about the information rate of the access structure based on the d-dimensional cube. Developing a new technique, we determined the *exact* information rate for all dimensions $d \geq 2$, which is $2/d$. Previously this value was known to be between $2/(d + 1)$ and $4/d$.

We extended the definition of information rate to cover infinite graphs. We consider this extension to be an important contribution, and hope to see further applications. As a non-trivial example, we determined the information rate of the (infinite) d-dimensional lattice, which is $1/d$.

During the proof we estimated the information rate of the "finite" lattice L_k^d which has exactly k vertices along each dimensions. While the estimate was enough to get the information rate of the infinite lattice L^d, the exact information rate for the finite graph L_k^d (both worst case and average) remains an open problem. To get a better upper bound, consider the following secret sharing sceme. Use the construction of Theorem 4.5 only inside L_k^d, and for the missing edges on the surface use similar construction but with one dimension less. In this scheme inner vertices will receive a total of d bits, while vertices on the surface will receive $1/2$ bit less. Thus the sum the size of all shares is

$$dk^d - \frac{1}{2}\bigl(k^d - (k-2)^d\bigr) \approx dk^d - dk^{d-1},$$

as there are $(k-2)^d$ inside vertices in L_k^d. Comparing this to the bound in Theorem 4.9, the two values are approximately equal, but remains some discrepancy.

Determining the worst case information rate of L_k^d seems to be a harder problem. We conjecture that for $d \geq 2$, $k \geq 4$ this value equals to $1/d$, i.e. the average information rate for the whole infinite lattice. This conjecture was verified for $d = 2$: we showed that there is a graph G which can be embedded into L_k^2, $k \geq 4$ as a spanned subgraph such that for a subset A of the vertices of G,

$$\sum_{v \in A} f(v) \geq 2|A|$$

for each feasible function f. Consequently in any perfect secret sharing scheme at least one element of A will receive a share which is at least twice as large as the secret is, i.e. the worst case information ratio of G, and thus of L_k^2, is at least 2.

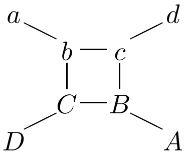

Figure 1: the graph G

The graph G is depicted on figure 1. Clearly G is a spanned subgraph of L_k^2 when $k \geq 4$. The subset in question is formed by the four vertices b, c, B, and C. We need to check that the sum of $f(b)$, $f(c)$, $f(B)$ and $f(C)$ is at least 8. As $f(b) + f(c) \geq f(bc)$ and $f(B) + f(C) \geq f(BC)$ by submodularity, it is enough to prove to following

Claim 5.1 *In the graph G, $f(bc) + f(BC) \geq 8$.*

Proof Each of the inequalities below are instances of one of the properties (a)–(e) of the function f, and the one below the line is their sum:

$$\begin{array}{rl}
f(a) + f(b) & \geq f(ab) \\
f(ab) + f(bc) & \geq 1 + f(b) + f(abc) \\
f(acAC) - f(acA) & \geq f(acACD) - f(acAD) \geq 1 \\
f(acABC) - f(acAC) & \geq 1 \\
f(ac) - f(a) & \geq f(acB) - f(aB) \\
f(acB) - f(aB) & \geq 1 + f(acABC) - f(aABC) \\
f(abc) - f(ac) & \geq f(abcA) - f(acA) \\
\hline
f(bc) & \geq 4 + f(abcA) - f(aABC)
\end{array}$$

Swapping lower case and upper case letters leaves the graph unchanged, thus we also have the "swapped" instance:

$$f(BC) \geq 4 + f(aABC) - f(abcA).$$

Adding these latter two inequalities up, we get the required result. ∎

References

[1] C. Blundo, A. De Santis, D. R. Stinson, U. Vaccaro: Graph Decomposition and Secret Sharing Schemes *Journal of Cryptology*, Vol 8(1995) pp. 39–64.

[2] L. Csirmaz: Secret sharing schemes on graphs, *Studia Mathematica Hungarica*, submitted – available as IACR preprint `http://eprint.iacr.org/2005/059`

[3] I. Csiszár and J. Körner: *Information Theory. Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981.

[4] D. R. Stinson: Decomposition construction for secret sharing schemes, *IEEE Trans. Inform. Theory* Vol 40(1994) pp. 118-125.