ISSN 0032-9460, Problems of Information Transmission, 2012, Vol. 48, No. 1, pp. 31–46. © Pleiades Publishing, Inc., 2012.
Original Russian Text © R. Ahlswede, V.S. Lebedev, 2012, published in Problemy Peredachi Informatsii, 2012, Vol. 48, No. 1, pp. 37–53.

Shadows under the Word-Subword Relation

R. Ahlswede† **and V. S. Lebedev**[a1]

[a] *Kharkevich Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow*
lebedev37@mail.ru

Received May 24, 2011; in final form, October 6, 2011

Abstract—We introduce a minimal shadow problem for a word-subword relation. We obtain upper and lower bounds for the minimal shadow cardinality.

DOI: 10.1134/S0032946012010048

1. INTRODUCTION

Quite surprisingly, it seems that the minimal shadow problem for the word-subword relation introduced here has not been studied before, whereas its analogs for sets [1–4], sequences [5], and vector spaces over finite fields [6] are well known.

For an alphabet $\mathcal{X} = \{0, 1, \ldots, q-1\}$, we consider the set $\mathcal{X}^k$ of words $x^k = x_1 x_2 \ldots x_k$ of length k. For a word $a^k = a_1 a_2 \ldots a_k \in \mathcal{X}^k$, we define its left shadow

$$\mathrm{shad}^L(a^k) = a_2 \ldots a_k, \tag{1}$$

i.e., the subword resulting from deleting the first letter a_1 in a^k, and its right shadow

$$\mathrm{shad}^R(a^k) = a_1 \ldots a_{k-1}, \tag{2}$$

i.e., the subword resulting from deleting the last letter a_k in a^k. Note that $\mathrm{shad}^L(a^k) = \mathrm{shad}^R(a^k)$ if and only if $a^k = aa\ldots a$, $a \in \mathcal{X}$, because $a_2 a_3 \ldots a_k = a_1 a_2 \ldots a_{k-1}$ implies $a_1 = a_2 = a_3 = \ldots = a_k$.

We define the shadow of a^k by

$$\mathrm{shad}(a^k) = \mathrm{shad}^L(a^k) \cup \mathrm{shad}^R(a^k). \tag{3}$$

Unless a^k has identical letters, $\mathrm{shad}(a^k)$ consists of two elements.

Now for any subset $A \subset \mathcal{X}^k$ we define its left shadow

$$\mathrm{shad}^L(A) = \bigcup_{a^k \in A} \mathrm{shad}^L(a^k), \tag{4}$$

right shadow

$$\mathrm{shad}^R(A) = \bigcup_{a^k \in A} \mathrm{shad}^R(a^k), \tag{5}$$

and shadow

$$\mathrm{shad}(A) = \mathrm{shad}^L(A) \cup \mathrm{shad}^R(A). \tag{6}$$

We are interested in finding the minimal shadow of N-sets $A \subset \mathcal{X}^k$, i.e., the function

$$\triangle_k(q, N) = \min\{|\mathrm{shad}(A)| : \ A \subset \mathcal{X}^k, \ |A| = N\}. \tag{7}$$

We write for short $\triangle_k(N)$ if q is fixed, and $\triangle(N)$ if k is also fixed. We also use the functions $\triangle_k^L(N)$ and $\triangle_k^R(N)$ (respectively, $\triangle^L(N)$ and $\triangle^R(N)$), where the minimization is over left and right shadows, respectively.

[1] Supported in part by the Russian Foundation for Basic Research, project no. 09-01-00536.

2. PRELIMINARY RESULTS

We denote by ab the concatenation of words a and b (the length of this word is the sum of lengths of a and b). Denote by AB the set of all words ab where $a \in A$ and $b \in B$. For example, the set $\mathcal{X}b\mathcal{X}$ consists of q^2 words that have any symbols in the first and last positions and have the word b in the middle.

Consider the following configurations:

(i) Words $xxx\ldots x$, $x \in \mathcal{X}$, whose number is $q = |\mathcal{X}|$. Their shadow has cardinality 1.

(ii) Words

$$\begin{aligned} a^k &= cdcd\ldots cd, \\ b^k &= dcdc\ldots dc \end{aligned} \quad \text{if } k \text{ is even,}$$

and analogously

$$\begin{aligned} a^k &= cd\ldots c, \\ b^k &= dc\ldots d \end{aligned} \quad \text{if } k \text{ is odd.}$$

Shadows of these words have cardinality 2.

(iii) In the set $\mathcal{X}B\mathcal{X}$, all the q words of the form xby, where x is a fixed element, $b \in \mathcal{B}$, and $y \in \mathcal{X}$, have identical right shadows. Similarly for left shadows.

Note that for all these configurations we have $\triangle(N) \le N$; let us prove this in general.

First consider the binary case.

Lemma 1. *For $q = 2$ and $k \ge 3$ we have*

$$\triangle(N) \le N, \quad \textit{for all} \quad N \le 2^k.$$

Proof. Write N in the form $N = 4M + p$, where $0 \le p < 4$.

Case $p = 0$. Choose any $B \subset \mathcal{X}^{k-2}$ with $|B| = M$; then $A = \mathcal{X}B\mathcal{X}$ is of cardinality N. It is easily seen that

$$|\text{shad}(A)| = |\mathcal{X}B \cup B\mathcal{X}| \le |B\mathcal{X}| + |\mathcal{X}B| = 4M = N.$$

Case $3 \ge p \ge 1$. Choose $\mathcal{B} \subset \mathcal{X}^{k-2} \setminus \{0^{k-2}\}$ with $|\mathcal{B}| = M$ and $A_p = \mathcal{X}\mathcal{B}\mathcal{X} \cup C_p$, where $C_1 = \{00^{k-2}0\}$, $C_2 = \{00^{k-2}0, 00^{k-2}1\}$, and $C_3 = \{00^{k-2}0, 00^{k-2}1, 10^{k-2}0\}$. It is clear that $|\text{shad}(A_p)| \le 4M + p$. $\triangle$

For a q-ary case, we have the following fact.

Lemma 2. *Consider $\mathcal{X} = \{0, 1, \ldots, q-1\}$, $k \ge 3$, and $N \le q^k$. Write $N = q^2M + p$, $0 \le p < q^2$; then*

$$\triangle(N) \le 2qM + \begin{cases} 0 & \textit{if } p = 0, \\ \lceil\sqrt{p}\rceil + \lfloor\sqrt{p}\rfloor - 1 & \textit{if } \lceil\sqrt{p}\rceil\lfloor\sqrt{p}\rfloor \ge p > 0, \\ 2\lceil\sqrt{p}\rceil - 1 & \textit{otherwise} \end{cases} \tag{8}$$

and

$$\triangle(N) \le \frac{2}{q}N - \frac{2}{q}p + \begin{cases} 0 & \textit{if } p = 0, \\ \lceil\sqrt{p}\rceil + \lfloor\sqrt{p}\rfloor - 1 & \textit{if } \lceil\sqrt{p}\rceil\lfloor\sqrt{p}\rfloor \ge p > 0, \\ 2\lceil\sqrt{p}\rceil - 1 & \textit{otherwise.} \end{cases} \tag{9}$$

Proof. *Case* $p = 0$. Choose any $\mathcal{B} \subset \mathcal{X}^{k-2}$ with $|\mathcal{B}| = M$ and $A = \mathcal{X}\mathcal{B}\mathcal{X}$. Then $|\text{shad}(A)| \le 2qM$, and we obtain (8).

Case $q^2 - 1 \ge p \ge 1$. Choose $\mathcal{B} \subset \mathcal{X}^{k-2} \setminus \{0^{k-2}\}$ with $|\mathcal{B}| = M$ and $A_p = \mathcal{X}\mathcal{B}\mathcal{X} \cup D_p$, where D_p is a balanced subset of $\mathcal{X}0^{k-2}\mathcal{X}$ with p elements. This means that we take $D_p = \mathcal{Y}0^{k-2}\mathcal{Y}'$, where the difference $|\{\mathcal{Y} \setminus \mathcal{Y}'\} \cup \{\mathcal{Y}' \setminus \mathcal{Y}\}|$ between $|\mathcal{Y}|$ and $|\mathcal{Y}'|$ is the minimum possible. Then

$$|\text{shad}(A_p)| \le 2qM + 2\lceil\sqrt{p}\rceil - 1,$$

and (8) is proved. From this, an easy computation yields (9). $\triangle$

Remark 1. For $q = 2$ bound (8) is equal to N. Hence, Lemma 2 implies Lemma 1.

Remark 2. For $N = q^\ell < q^k$ we may choose $A = \mathcal{X}^{\ell-1}0^{k-\ell}\mathcal{X}$ to obtain $|\mathrm{shad}(A)| = \Big(\frac{2}{q} - \frac{1}{q^2}\Big)q^\ell = \Big(\frac{2}{q} - \frac{1}{q^2}\Big)N$, which is better than (8). For $q = 2$ we get $\triangle_k(2^\ell) \le \frac{3}{4}2^\ell$.

3. CONCEPT OF BASIC SETS

In Section 2 we have obtained our first upper bounds on minimal shadows for sets with the structure $A = \mathcal{X}B\mathcal{X}$. We generalize this structure by taking unions of such sets. Consider the sets

$$\mathcal{X}^\ell 0^m \mathcal{X}^r. \tag{10}$$

Now we define our main concept.

Definition 1. For nonnegative integers ℓ, m, and r satisfying

$$\ell \ge r \tag{11}$$

and

$$k = \ell + m + r, \tag{12}$$

we define a *basic set* $\mathcal{B}(k, \ell, r)$ in $\mathcal{X}^k$ as the following union:

$$\mathcal{B}(k, \ell, r) = \bigcup_{s=0}^{\ell-r} \mathcal{X}^{\ell-s}0^m\mathcal{X}^{r+s}. \tag{13}$$

For instance, $\mathcal{B}(7, 3, 1)$ is the union of rows of the matrix

$$\begin{matrix} \mathcal{X} & \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & \mathcal{X} \\ \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & \mathcal{X} & \mathcal{X} \\ \mathcal{X} & 0 & 0 & 0 & \mathcal{X} & \mathcal{X} & \mathcal{X}, \end{matrix}$$

and $\mathcal{B}(8, 3, 2)$ is the union of rows of the matrix

$$\begin{matrix} \mathcal{X} & \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & \mathcal{X} & \mathcal{X} \\ \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & \mathcal{X} & \mathcal{X} & \mathcal{X}. \end{matrix}$$

We denote these matrices by $[\mathcal{B}(7, 3, 1)]$ and $[\mathcal{B}(8, 3, 2)]$, and in the general case, by $[\mathcal{B}(k, \ell, r)]$.

Here are key properties of such sets.

Lemma 3. *For all* $\ell \ge r \ge 1$, $m + r > \ell$ *(i.e.,* $k = \ell + m + r > 2\ell$*), and* $q = 2$*, we have*

(i) $|\mathcal{B}(k, \ell, r)| = 2^{\ell+r} + 2^{\ell+r-1}(\ell - r) = 2^{\ell+r-1}(\ell - r + 2)$,
(ii) $\mathrm{shad}\,\mathcal{B}(k, \ell, r) = \mathcal{B}(k - 1, \ell, r - 1)$,
(iii) $\mathcal{B}(k, \ell, r) \subset \mathcal{B}(k, \ell + 1, r - 1)$,
(iv) $|\mathrm{shad}\,\mathcal{B}(k, \ell, r)| = |\mathcal{B}(k - 1, \ell, r - 1)| = \dfrac{|\mathcal{B}(k, \ell, r)|}{2} + 2^{\ell+r-2}$,
(v) $|\mathrm{shad}\,\mathcal{B}(k, \ell, r)| = 2^{\ell+r-2}(\ell - r + 3)$.

Example. Let $k = 9$, $\ell = 4$, and $r = 1$. Then

$$|\mathcal{B}(9, 4, 1)| = 2^5 + 2^4 3 = 32 + 48 = 80,$$
$$\triangle_9(80) \le 2^{4+1-2}(4 - 1 + 3) = 48.$$

This is clearly better than the bound in Lemma 1.

An important consequence is as follows.

Corollary 1. *For* $N = 2^{\ell+r-1}(\ell - r + 2)$ *and* $k = \ell + m + r > 2\ell \geq 2r \geq 2$, *we have*

$$\triangle_k(N) \leq \frac{1}{2}\frac{\ell - r + 3}{\ell - r + 2}N. \tag{14}$$

Proof of Lemma 3. (i) First, as an example of a basic set $\mathcal{B}(k, \ell, r)$, consider $(k, \ell, r) = (9, 4, 2)$:

$$\begin{matrix} \mathcal{X} & \mathcal{X} & \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & \mathcal{X} & \mathcal{X} \\ \mathcal{X} & \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & \mathcal{X} & \mathcal{X} & \mathcal{X} \\ \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & \mathcal{X} & \mathcal{X} & \mathcal{X} & \mathcal{X}. \end{matrix}$$

Note that $\mathcal{B}(9, 4, 2)$ equals the union of the following sets:

$$\begin{matrix} \mathcal{X} & \mathcal{X} & \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & \mathcal{X} & \mathcal{X} \\ \mathcal{X} & \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & 1 & \mathcal{X} & \mathcal{X} \\ \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & 1 & \mathcal{X} & \mathcal{X} & \mathcal{X}. \end{matrix}$$

These row sets have the total cardinality of $2^6 + 2^5 + 2^5$.

For the general case of $\ell \leq m + r$, we find that the first set has cardinality $2^{\ell+r}$, and the other $\ell - r$ sets have cardinality $2^{\ell+r-1}$. Hence,

$$|\mathcal{B}(k, \ell, r)| = 2^{\ell+r} + 2^{\ell+r-1}(\ell - r).$$

(ii) We illustrate the claim by the following example:

$$\begin{matrix} & & & & \text{shad}^L\,\mathcal{B}(9,4,2) & & & & & & & & & \text{shad}^R\,\mathcal{B}(9,4,2) & & & \\ & & & & & & & & & \mathcal{X} & \mathcal{X} & \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & \mathcal{X} \\ \mathcal{X} & \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & \mathcal{X} & \mathcal{X} & = & \mathcal{X} & \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & \mathcal{X} & \mathcal{X} \\ \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & \mathcal{X} & \mathcal{X} & \mathcal{X} & = & \mathcal{X} & \mathcal{X} & 0 & 0 & 0 & \mathcal{X} & \mathcal{X} & \mathcal{X} \\ \mathcal{X} & 0 & 0 & 0 & \mathcal{X} & \mathcal{X} & \mathcal{X} & \mathcal{X}. & & & & & & & & & \end{matrix} \tag{15}$$

If we add the first row of the second matrix to the first matrix (respectively, the last row of the first matrix to the second matrix), then $\text{shad}\,\mathcal{B}(9, 4, 2) = \mathcal{B}(8, 4, 1)$, so k and r are reduced by 1.

In the general case, right shadow deletes from the basic set one $\mathcal{X}$ from the right, and left shadow, from the left. Hence, in the general case k and r are reduced by 1 too.

(iii) Simply note that for $\ell > r$ the matrix $[\mathcal{B}(k, \ell, r)]$ is obtained from $[\mathcal{B}(k - 1, \ell, r - 1)]$ by deleting the first and last row.

(iv) Note that in $\text{shad}^L\,\mathcal{B}(k, \ell, r)$ we have one $\mathcal{X}$ less than in $\mathcal{B}(k, \ell, r)$ in each row. Also, we have an extra row; this row $\mathcal{X}^\ell 0^m \mathcal{X}^{r-1}$ in $\text{shad}^R\,\mathcal{B}(k, \ell, r)$ corresponds to $\mathcal{X}^{\ell-1} 1\, 0^m \mathcal{X}^{r-1}$ of cardinality $2^{\ell+r-2}$.

Formally, (iv) follows from the equality

$$2^{\ell+r-2}(\ell - r + 2) + 2^{\ell+r-2} = 2^{\ell+(r-1)-1}(\ell - (r - 1) + 2).$$

(v) Follows from (i) and (ii). △

Generalization to the *q*-ary case. It is easily seen that (i) and (iv) in Lemma 3 can be extended to (i′) and (iv′) in Lemma 4. In (i) one should take any nonzero element instead of 1, so the first row has cardinality $q^{\ell+r}$ and the other $\ell - r$ rows have cardinality $q^{\ell+r-1}(q - 1)$. Hence, we have the following result.

Lemma 4. *For all $\ell \ge r \ge 1$, $m + r > \ell$ (i.e., $k = \ell + m + r > 2\ell$), and $q \ge 2$, we have*

$$\text{(i}'\text{)}\qquad |\mathcal{B}(k,\ell,r)| = q^{\ell+r} + q^{\ell+r-1}(\ell - r)(q-1),$$

$$\text{(iv}'\text{)}\qquad |\operatorname{shad}\mathcal{B}(k,\ell,r)| = |\mathcal{B}(k-1,\ell,r-1)| = \frac{|\mathcal{B}(k,\ell,r)|}{q} + q^{\ell+r-2}(q-1)$$
$$= q^{\ell+r-2}((\ell - r + 2)(q-1) + 1). \tag{16}$$

For $N = |\mathcal{B}(k,\ell,r)| = q^{\ell+r} + q^{\ell+r-1}(\ell - r)(q-1)$, from $|\operatorname{shad}\mathcal{B}(k,\ell,r)| = \dfrac{N}{q} + q^{\ell+r-2}(q-1)$ we obtain

$$\frac{\triangle_k(q,N)}{N} \le \frac{1}{q} + \frac{1}{q}\frac{(q-1)}{q + (\ell - r)(q-1)} = \frac{1}{q}\left(1 + \frac{q-1}{(\ell - r + 1)(q-1) + 1}\right)$$
$$\le \frac{1}{q}\left(1 + \frac{1}{\ell - r + 1}\right). \tag{17}$$

Hence follows an important consequence.

Corollary 2. *For $N = q^{\ell+r} + q^{\ell+r-1}(\ell - r)(q-1)$ and $k = \ell + m + r > 2\ell \ge 2r \ge 2$, we have*

$$\triangle_k(q,N) \le \frac{1}{q}\left(1 + \frac{1}{\ell - r + 1}\right)N. \tag{18}$$

Remark 3. For $q = 2$ we had a smaller factor $1 + \dfrac{1}{\ell - r + 2}$ in Corollary 1.

4. LOWER BOUND

For any $A \subset \mathcal{X}^k$ and $\mathcal{Y} \subset \mathcal{X}$, define

$$A^1_{\mathcal{Y}} = \{x_2 \ldots x_k \in \mathcal{X}^{k-1} \colon\ \mathcal{Y}x_2 \ldots x_k \subset A \text{ and } xx_2 \ldots x_k \notin A \text{ for all } x \in \mathcal{X} \setminus \mathcal{Y}\}. \tag{19}$$

Clearly, these sets are contained in $\mathcal{X}^{k-1}$ and are disjoint. Moreover,

$$\operatorname{shad}(A) \supset \operatorname{shad}^L(A) = \bigcup_{\mathcal{Y} \subset \mathcal{X}} A^1_{\mathcal{Y}}, \tag{20}$$
$$A = \bigcup_{\mathcal{Y} \subset \mathcal{X}} \mathcal{Y}A^1_{\mathcal{Y}}, \tag{21}$$

and since $|\mathcal{Y}| \le q$, we get

$$|\operatorname{shad}(A)| \ge \frac{1}{q}|A|. \tag{22}$$

Hence, with the use of Corollary 2, we obtain the following result.

Theorem 1. *For $N = q^{\ell+r} + q^{\ell+r-1}(\ell - r)(q-1)$ and $k = \ell + m + r > 2\ell \ge 2r \ge 2$, we have*

$$\frac{1}{q}N \le \triangle_k(q,N) \le \frac{1}{q}\left(1 + \frac{1}{\ell - r + 1}\right)N. \tag{23}$$

Moreover, the lower bound holds for all N.

5. CARDINALITY OF BASIC SETS FOR $\ell > m$

Note that $|\mathcal{B}(k,\ell,r)| = |\mathcal{B}(k-2r,\ell-r,0)|q^{2r}$. Hence, we are interested in the cardinality of $\mathcal{B}(k,\ell,0)$ for an arbitrary ℓ and $m = k-\ell-r$, $\ell > m$ (the case of $\ell \le m$ was considered in Lemma 4).

Theorem 2. *For any ℓ and m such that $\ell > m$, we have*

$$|\mathcal{B}(k,\ell,0)| = q^{\ell-1}(\ell(q-1)+q) - (q-1)\sum_{i=1}^{\ell-m} q^{\ell-m-i}|\mathcal{B}(m+i-1,i-1,0)|,$$

and for $N = |\mathcal{B}(k,\ell,1)| = q^2|\mathcal{B}(k-2,\ell-1,0)|$,

$$\frac{\triangle(N)}{N} \le \frac{1}{q}\left(1+\frac{1}{\ell}\right).$$

Proof. Denote by $H(\ell,m,a)$ the number of sequences from $\mathcal{X}^{\ell+m}$ that are not covered by the first a rows of the matrix $[\mathcal{B}(k,\ell,0)]$. Consider the jth row $\mathcal{X}^{\ell-j+1}0^m\mathcal{X}^{j-1}$ in $[\mathcal{B}(k,\ell,0)]$. How many new sequences does it add? Using our notation, we obtain

$$q^{\ell-j+1}(q-1)H(\ell,m,j-m-1)$$

such sequences.

We have

$$H(\ell,m,a) = q^{m+a-1} - |\mathcal{B}(m+a-1,a-1,0)|. \tag{24}$$

Let $i = j-m-1$; then for $i = 1,2,\ldots,\ell-m$ we add

$$q^{\ell-i-m}(q-1)\left(q^{m+i-1} - |\mathcal{B}(m+i-1,i-1,0)|\right)$$

sequences, and this proves that

$$|\mathcal{B}(k,\ell,0)| = q^{\ell} + q^{\ell-1}m(q-1) + \sum_{i=1}^{\ell-m} q^{\ell-m-i}(q-1)\left(q^{m+i-1} - |\mathcal{B}(m+i-1,i-1,0)|\right).$$

Hence,

$$|\mathcal{B}(k,\ell,0)| = q^{\ell} + q^{\ell-1}m(q-1) + q^{\ell-1}(\ell-m)(q-1) - \sum_{i=1}^{\ell-m} q^{\ell-m-i}(q-1)|\mathcal{B}(m+i-1,i-1,0)|.$$

We obtain

$$\triangle(N) \le \frac{1}{q}\frac{|\mathcal{B}(k,\ell,0)|}{q|\mathcal{B}(m+\ell-1,\ell-1,0)|}N.$$

Therefore,

$$\frac{\triangle(N)}{N} \le \frac{1}{q}\left(1 + \frac{(q-1)q^{\ell-1} - (q-1)|\mathcal{B}(\ell-1,\ell-m-1,0)|}{q^{\ell-1}(q+(\ell-1)(q-1)) - (q-1)\sum\limits_{i=1}^{\ell-m-1} q^{\ell-m-i}|\mathcal{B}(m+i-1,i-1,0)|}\right).$$

We have obtained this formula using the equality

$$\operatorname{shad}(\mathcal{B}(m+\ell+1,\ell,1)) = \mathcal{B}(m+\ell,\ell,0).$$

One can prove that for such N

$$\frac{(q-1)\left(q^{\ell-1}-|\mathcal{B}(\ell-1,\ell-m-1,0)|\right)}{q^{\ell-1}((\ell-1)(q-1)+q)-(q-1)\sum\limits_{i=1}^{\ell-m-1}q^{\ell-m-i}|\mathcal{B}(m+i-1,i-1,0)|}\le\frac{1}{\ell}.$$

Indeed,

$$\begin{aligned}(q^{\ell-1}&-|\mathcal{B}(\ell-1,\ell-m-1,0)|)(q-1)\ell\\&\le q^{\ell-1}((\ell-1)(q-1)+q)-(q-1)\sum_{i=1}^{\ell-m-1}q^{\ell-m-i}|\mathcal{B}(m+i-1,i-1,0)|\end{aligned}$$

if

$$\sum_{i=1}^{\ell-m-1}q^{\ell-m-i}|\mathcal{B}(m+i-1,i-1,0)|\le|\mathcal{B}(\ell-1,\ell-m-1,0)|\ell.$$

It is clear that for any natural u

$$q|\mathcal{B}(m+u-1,u-1,0)|<|\mathcal{B}(m+u,u,0)|.$$

Therefore,

$$\sum_{i=1}^{\ell-m-1}q^{\ell-m-i}|\mathcal{B}(m+i-1,i-1,0)|\le|\mathcal{B}(\ell-1,\ell-m-1,0)|(\ell-m-1),$$

which proves the theorem. $\triangle$

Remark 4. For the binary case one can prove that for $N=|\mathcal{B}(k,\ell,1)|=4|\mathcal{B}(k-2,\ell-1,0)|$ and $\ell>m$ one has

$$\frac{\triangle(N)}{N}\le\frac{1}{2}\left(1+\frac{1}{\ell+1}\right).$$

Extended basic sets. For basic sets $\mathcal{B}(k,\ell,r)$ we used building sets

$$\mathcal{X}^{\ell}0^{m}\mathcal{X}^{r} \tag{25}$$

and took unions of such sets. Now we define a dual building set as

$$0^{m}\mathcal{X}^{k-2m}0^{m}.$$

We add these dual building sets to the basic set and define an extended basic set $\widetilde{\mathcal{B}}(k,\ell,1)$ as

$$\widetilde{\mathcal{B}}(k,\ell,1)=\left(\bigcup_{s=0}^{\ell-1}\mathcal{X}^{\ell-s}0^{m}\mathcal{X}^{1+s}\right)\cup 0^{m}\mathcal{X}^{k-2m}0^{m}=\mathcal{B}(k,\ell,r)\cup 0^{m}\mathcal{X}^{k-2m}0^{m}. \tag{26}$$

The set $\widetilde{\mathcal{B}}(k,\ell,1)$ has a larger cardinality than $|\mathcal{B}(k,\ell,1)|$, but their shadows coincide!

Theorem 3. *For $\ell\ge m$ we have*

(i) $\widetilde{\mathcal{B}}(k,\ell,1)=|\mathcal{B}(k,\ell,1)|+1$ *for* $\ell=m$,
(ii) $\widetilde{\mathcal{B}}(k,\ell,1)=|\mathcal{B}(k,\ell,1)|+2^{\ell-m-1}$ *for* $m<\ell\le 2m$,
(iii) $\widetilde{\mathcal{B}}(k,\ell,1)=|\mathcal{B}(k,\ell,1)|+2^{\ell-m-1}-|\mathcal{B}(\ell-m-1,\ell-2m-1,0)|$ *for* $\ell>2m$,
(iv) $\operatorname{shad}(\widetilde{\mathcal{B}}(k,\ell,1))=|\mathcal{B}(k-1,\ell,0)|$.

Proof. For $\ell = m$ we add a new word $0^m 1 0^m$ to the basic set. In case (ii), a new block is $0^m 1 \mathcal{X}^{\ell-m-1} 1 0^m$. Since it has a 1 in the $(m+1)$st coordinate, it is not covered by the last m rows of the basic matrix $[\mathcal{B}(k,\ell,1)]$. Since it has a 1 in the $(\ell+1)$st coordinate, it is not covered by the first m rows of the basic matrix $[\mathcal{B}(k,\ell,1)]$. In total, we have ℓ rows in $[\mathcal{B}(k,\ell,1)]$; hence, this proves case (ii). In the case of $\ell > 2m$, there is also a 1 in both the $(m+1)$st and $(\ell+1)$st rows, but in this case we obtain $H(\ell-2m-1, m, \ell-2m)$ new sequences. Using (24), this proves case (iii). Dual building sets $0^m \mathcal{X}^{k-2m} 0^m$ yield a shadow which is a subset of the basic set $\mathcal{B}(k-1,\ell,0)$, whence follows (iv). $\triangle$

6. SHADOWS, UP-SHADOWS, AND THEIR INTERRELATION

Consider a word $b^{k-1} \in \mathcal{X}^{k-1}$. Then

$$\text{up-shad}(b^{k-1}) = \{a^k :\ a^k \in \mathcal{X}^k,\ b^{k-1} \in \text{shad}(a^k)\}.$$

Now for any subset $B \subset \mathcal{X}^{k-1}$ we define its up-shadow:

$$\text{up-shad}(B) = \bigcup_{b^{k-1} \in B} \text{up-shad}(b^{k-1}).$$

For a fixed k we are interested in the function

$$\nabla(M) = \min\{|\text{up-shad}(B)| :\ B \subset \mathcal{X}^{k-1},\ |B| = M\}.$$

The following function is important for finding a relation between the shadows.

Definition 2. Consider a set C of sequences of length n with cardinality M. Let $s_n(C,M)$ be the number of pairs (z, x^n), $z \in \mathcal{X}$, $x^n = (x_1, x_2, \ldots, x_n) \in C$, such that $(z, x_1, x_2, \ldots, x_{n-1}) \in C$. Denote

$$s_n(M) = \max_C s_n(C,M). \tag{27}$$

Lemma 5. *The following conditions are equivalent for* $C \subseteq \mathcal{X}^n$:

(i) $|C| \neq q^n$;
(ii) $\exists z \in \mathcal{X}$ *and* $c^n = (c_1, c_2, \ldots, c_n) \in C$ *such that* $(z, c_1, c_2, \ldots, c_{n-1}) \notin C$;
(iii) $\triangle(\nabla(C)) \neq C$;
(iv) $\nabla(\triangle(C)) \neq C$.

Proof. (ii) $\Rightarrow$ (iii). Consider $c \in C$ satisfying (ii). Then $(z, c_1, c_2, \ldots, c_n) \in \nabla(C)$, and therefore $y^n = (z, c_1, c_2, \ldots, c_{n-1}) \in \triangle(\nabla(C))$. However, (ii) implies $y^n \notin C$. Hence, $\triangle(\nabla(C)) \neq C$.

(iii) $\Rightarrow$ (i). The set $\nabla(c^n)$ consists of $\mathcal{X}c_1, c_2, \ldots, c_n \cup c_1, c_2, \ldots, c_n\mathcal{X}$. Hence,

$$\triangle(\nabla(c^n)) = c_1, c_2, \ldots, c_n \cup \mathcal{X}c_1, c_2, \ldots, c_{n-1} \cup c_2, \ldots, c_n\mathcal{X}.$$

Therefore, $C \subseteq \triangle(\nabla(C))$. Thus, (i) is proved.

(ii) $\Rightarrow$ (iv). Consider $c^n \in C$ satisfying (ii). Then we have $(c_1, c_2, \ldots, c_{n-1}) \in \triangle(C)$, and therefore $y^n = (z, c_1, c_2, \ldots, c_{n-1}) \in \nabla(\triangle(C))$. However, (ii) implies $y^n \notin C$.

(iv) $\Rightarrow$ (i). For any $c^n \in C$ we have

$$\triangle(c^n) = c_2, \ldots, c_n \cup c_1, c_2, \ldots, c_{n-1}$$

and

$$\nabla(\triangle(c^n)) = \mathcal{X}c_2, \ldots, c_n \cup c_2, \ldots, c_n\mathcal{X} \cup \mathcal{X}c_1, c_2, \ldots, c_{n-1} \cup c_1, c_2, \ldots, c_{n-1}\mathcal{X}.$$

Thus,

$$C \subseteq \triangle(\nabla(C)) \subseteq \nabla(\triangle(C)).$$

Hence, we get (i).

(i) $\Rightarrow$ (ii). Assume that for all $z \in \mathcal{X}$ and all $c^n \in C$ property (ii) is fulfilled. Then $\mathcal{X}c_1, c_2, \ldots, c_{n-1} \in C$. Hence, $\mathcal{X}\mathcal{X}c_1, c_2, \ldots, c_{n-2} \in C$, $\mathcal{X}\mathcal{X}\mathcal{X}c_1, c_2, \ldots, c_{n-3} \in C$, etc. Therefore, $\mathcal{X}\mathcal{X}\mathcal{X}\ldots\mathcal{X}\mathcal{X} \in C$, and we get a contradiction to (i). $\triangle$

Property (ii) and Definition 2 immediately imply the following result.

Corollary 3. *If* $M' < M$, *then*

$$s(M') < s(M).$$

Thus, $s(M)$ *is a strictly monotone increasing function.*

Theorem 4. *For any* q, k, *and* $M \le q^{k-1}$, *we have*

$$\triangle_k(s_{k-1}(M)) = M.$$

Proof. Let C ($|C| = M$) be a set maximizing (27). We add a sequence $(z, x_1, x_2, \ldots, x_n)$ to the set D, if the condition from Definition 2 holds for this z and $(x_1, x_2, \ldots, x_n) \in C$. Then $|D| = s_n(M)$ and $\mathrm{shad}(D) = C$. Hence,

$$\triangle_k(s_{k-1}(M)) \le M.$$

If there existed a set C' of a smaller cardinality M', $M' < M$, and such that $s(M') = s(M)$, this would contradict Corollary 3. Hence, $\triangle_k(s_{k-1}(M)) = M$. $\triangle$

From this and Lemma 5, we have the following fact.

Corollary 4. *If* $N < q^k$, *then*

$$\frac{1}{q}N < \triangle_k(q, N). \tag{28}$$

7. ISOPERIMETRIC NUMBERS OF GRAPHS

Problems on isoperimetric numbers of graphs have been studied for a long time (see, e.g., [7,8]).

Consider a graph $G(V, E)$ with the set of vertices V and set of edges E. If $X \subseteq V$ is some set of vertices, then ∂X denotes the set of edges that have one end in X and the other in $V \setminus X$. Thus,

$$\partial X = \{(x, y) \in E;\ x \in X,\ y \in V \setminus X\}.$$

The edge-isoperimetric number of this graph is defined to be

$$i(G) = \min \frac{|\partial X|}{|X|},$$

where the minimum is over all nonempty subsets $X \subset V$ satisfying $|X| \le |V|/2$.

Denote by $N(X)$ the set of vertices of $V \setminus X$ adjacent to some vertex in X. Thus,

$$N(X) = \{y \in V \setminus X;\ x \in X,\ (x, y) \in E\}.$$

The vertex-isoperimetric number of this graph is defined to be

$$i_v(G) = \min \frac{|N(X)|}{|X|},$$

where the minimum is over all nonempty subsets $X \subset V$ satisfying $|X| \le |V|/2$.

We want to consider graphs related to the word-subword relation: a U-D graph and a D-U graph. Vertices of these graphs are all sequences of $\mathcal{X}^n$.

For the U-D graph, vertices $a^n = a_1a_2\dots a_n$ and $b^n = b_1b_2\dots b_n$ ($a^n \neq b^n$) are adjacent if and only if there exists c^{n+1} such that

$$c^{n+1} \in \text{up-shad}(a^n), \qquad b^n \in \text{shad}(c^{n+1}).$$

We would like to have a bijection between edges of the U-D graph and all words from $\mathcal{X}^{n+1}$.

To this end, for $a^n = b^n$ we draw a *single* edge (loop) in the graph if and only if $a_1 = a_2 = a_3 = \ldots = a_n$. Then we get a bijection between edges of the U-D graph and all words from $\mathcal{X}^{n+1}$. Under this definition of the graph, its edges can be identified with sequences of length $n+1$, and vertices connected by an edge are the right and left shadows of this sequence. Such a definition of the graph seems to be extremely natural.

Note that a vertex degree in this graph is $2q-1$ for a^n with $a_1 = a_2 = a_3 = \ldots = a_n$, and $2q$ for all other vertices.

For the D-U graph, an edge connects vertices $a^n = a_1a_2\dots a_n$ and $b^n = b_1b_2\dots b_n$ with $a^n \neq b^n$ if and only if there exists a word d^{n-1} such that

$$d^{n-1} \in \text{shad}(a^n), \qquad b^n \in \text{up-shad}(d^{n-1}).$$

The total number of edges in the D-U graph is $q^{n-1}q(q-1) + q^{n+1} = q^n(2q-1)$.

In this paper we do not consider properties of the D-U graph.

8. RELATION TO DE BRUIJN GRAPHS

Recall that for a fixed n and $k = n+1$ we are interested in

$$\triangle(N) = \min\{|\text{shad}(A)| : \ A \subset \mathcal{X}^k, \ |A| = N\}.$$

In graph theory, an n-dimensional De Bruijn graph of q symbols is a directed graph with q^n vertices consisting of all possible n-sequences of the given symbols. If one of the vertices can be obtained from another by shifting all symbols by one position to the left and adding a new symbol at the end, then the latter vertex has a directed edge to the former. Thus, the set of (directed) edges is

$$E = \{((v_1, v_2, \ldots, v_n), (w_1, w_2, \ldots, w_n)) : \ v_2 = w_1, \ v_3 = w_2, \ \ldots, \ v_n = w_{n-1}\}.$$

Each vertex has exactly q incoming and q outgoing edges. Consider an undirected $(k-1)$-dimensional De Bruijn graph. The graph is very close to the U-D graph. (Sequences a^k from $\mathcal{X}^k$ are edges in the graph. The left shadow $\text{shad}^L(a^k)$ and right shadow $\text{shad}^R(a^k)$ are vertices incident to this edge.) For a^n with $a_1 = a_2 = a_3 = \ldots = a_n$ we have a loop in the U-D graph and two loops in the De Bruijn graph.

The minimal shadow problem is equivalent to the problem of finding N edges incident to a minimum possible number of vertices. Theorem 4 shows that the problem of finding M vertices in the U-D graph that give the maximum possible number of edges between them is the inverse problem. Thus, the function $s_{k-1}(M)$ is very important for us. It is also related to the up-shadow problem.

Theorem 5. *For any q, k, and $M \le q^{k-1}$, we have*

$$\nabla(M) = 2qM - s_{k-1}(M).$$

Proof. The de Bruijn graph is regular. The vertex degree is $2q$. Hence, there are $2qM$ edges incident to M vertices from a set C, but some of them were calculated twice. The number of edges calculated twice is $s_{k-1}(M)$, and therefore the number of edges incident to C is $2qM - s_{k-1}(M)$. △

Theorem 6. *For any* $M \le q^{k-1}$ *we have*

$$s_{k-1}(q^{k-1} - M) = q^k - 2qM + s_{k-1}(M).$$

Proof. Let C ($|C| = M$) be a set of vertices that maximizes (27). Let a set B of cardinality $|B| = \nabla_{k-1}(M)$ consist of edges incident to M vertices of C. Then any edge out of B gives a shadow out of C. Hence,

$$\triangle_k(q^k - \nabla(M)) \le q^{k-1} - M.$$

Theorem 5 implies

$$\nabla(M) = 2qM - s_{k-1}(M).$$

From this and Corollary 4, we obtain

$$s_{k-1}(q^{k-1} - M) \ge q^k - 2qM + s_{k-1}(M).$$

If we do the same with the set $\mathcal{X}^{k-1} \setminus C$, we get

$$\triangle_k(q^k - \nabla_{k-1}(q^{k-1} - M)) \le M.$$

Therefore,

$$s_{k-1}(M) \ge q^k - 2q(q^{k-1} - M) + s_{k-1}(q^{k-1} - M),$$

whence we find

$$s_{k-1}(q^{k-1} - M) \le q^k - 2qM + s_{k-1}(M). \quad \triangle$$

Using this theorem, we can compute the rate $R = \triangle(N)/N$ for large N.

Proposition 1. *For* $N = 2^k - 2^\ell(\ell + 3)$ *and* $\ell < k/2$ *in the binary case we have*

$$R \le 1/2\left(1 + \frac{1}{2^{k-\ell} - \ell - 3}\right).$$

Proof. For $\ell < k/2$ and $M = 2^{\ell-1}(\ell + 2)$ we obtain $s(M) = 2^\ell(\ell + 1)$. Theorem 6 implies

$$s_{k-1}(2^{k-1} - 2^{\ell-1}(\ell + 2)) = 2^k - 2^{\ell-1}4(\ell + 2) + 2^\ell(\ell + 1) = 2^k - 2^\ell(\ell + 3).$$

Hence,

$$R \le \frac{2^{k-1} - 2^{\ell-1}(\ell + 2)}{2^\ell(2^{k-\ell} - \ell - 3)} = 1/2\left(1 + \frac{1}{2^{k-\ell} - \ell - 3}\right). \quad \triangle$$

Proposition 2. *For* $N = q^k - q^\ell(q + (q - 1)(\ell + 1))$ *and* $\ell < k/2$ *in a* q*-ary case we have*

$$R \le \frac{1}{q}\left(1 + \frac{q - 1}{q^{k-\ell} - (q + (q - 1)(\ell + 1))}\right).$$

Proof. For $\ell < k/2$ and $M = q^{\ell-1}(q + \ell(q - 1))$ we obtain $s(M) = q^\ell(q + (q - 1)(\ell - 1))$. Theorem 6 implies

$$\begin{aligned} s_{k-1}(q^{k-1} - q^{\ell-1}(q + \ell(q - 1))) &= q^k - 2qq^{\ell-1}(q + \ell(q - 1)) + q^\ell(q + (q - 1)(\ell - 1)) \\ &= q^k - q^\ell(q + (q - 1)(\ell + 1)). \end{aligned}$$

Hence,

$$R \le \frac{q^{k-1} - q^{\ell-1}(q + \ell(q - 1))}{q^k - q^\ell(q + (q - 1)(\ell + 1))} = \frac{1}{q}\left(1 + \frac{q - 1}{q^{k-\ell} - (q + (q - 1)(\ell + 1))}\right). \quad \triangle$$

Denote by i(U-D) the edge-isoperimetric number of the U-D graph. For any q and k we have the following fact.

Theorem 7. *For* $|N| \le q^k/2 - i(\text{U-D})q^{k-1}/4$ *we have*

$$\frac{\triangle_k(N)}{N} \ge \frac{1}{q}\left(1 + \frac{i(\text{U-D})}{2q - i(\text{U-D})}\right). \tag{29}$$

Proof. Since

$$\min\{|\partial X| : |X| = M\} = \nabla(M) - s(M) = 2qM - 2s(M), \tag{30}$$

for $M \le q^{k-1}/2$ we have

$$2q - \frac{2s(M)}{M} \ge i(\text{U-D}).$$

Hence,

$$s(M) \le qM - i(\text{U-D})M/2.$$

Therefore,

$$\frac{\triangle_k(qM - i(\text{U-D})M/2)}{qM - i(\text{U-D})M/2} \ge \frac{M}{qM - i(\text{U-D})M/2} = \frac{1}{q - i(\text{U-D})/2} = \frac{1}{q}\left(1 + \frac{i(\text{U-D})}{2q - i(\text{U-D})}\right). \quad \triangle$$

In [9] it was proved that

$$i(\text{U-D}) \ge \frac{q}{2(n-1)}.$$

Hence we get the following result.

Corollary 5. *For* $|N| \le q^k/2 - \dfrac{q^k}{8(k-2)}$ *we have*

$$\frac{\triangle_k(N)}{N} \ge \frac{1}{q}\left(1 + \frac{1}{4k-9}\right). \tag{31}$$

9. EDGE-ISOPERIMETRIC NUMBER OF THE DE BRUIJN GRAPH

In [9] there was obtained the following upper bound for the edge-isoperimetric number of the De Bruijn graph:

$$i(B(n,q)) \le \frac{2q\pi}{n+1}.$$

Here is an improvement of this bound.

Theorem 8. *The isoperimetric number of the de Bruijn graph satisfies the inequality*

$$i(B(n,q)) \le \frac{2q}{n - 2\log_q n + 1},$$

and in the binary case,

$$i(B(n,q)) \le \frac{4}{n - \log n + 2}.$$

Proof. From (30) it follows that for $M = |\mathcal{B}(n,\ell,0)|$ we obtain

$$i(B(n,q)) \le 2q - \frac{2s(M)}{M}.$$

It follows from Theorem 2 that

$$\frac{M}{s(M)} \le \frac{|\mathcal{B}(n,\ell,0)|}{|\mathcal{B}(k,\ell,1)|} \le \frac{1}{q}\left(1+\frac{1}{\ell}\right),$$

and in the binary case,

$$\frac{M}{s(M)} \le \frac{|\mathcal{B}(n,\ell,0)|}{|\mathcal{B}(k,\ell,1)|} \le \frac{1}{2}\left(1+\frac{1}{\ell+1}\right).$$

Hence,

$$i(B(n,q)) \le 2q - \frac{2q\ell}{\ell+1} = \frac{2q}{\ell+1},$$

and in the binary case,

$$i(B(n,2)) \le 4 - \frac{4(\ell+1)}{\ell+2} = \frac{4}{\ell+2}.$$

From Lemma 3 we obtain

$$|\mathcal{B}(n,\ell,0)| \le 2^{\ell-1}(\ell+2).$$

Hence, for $m \ge \log n$ we have $\ell \le n - \log n$, and for $n \ge 4$,

$$|\mathcal{B}(n,\ell,0)| \le \frac{2^n(n-\log n+4)}{2n} \le \frac{2^n}{2}.$$

Hence, in the binary case we obtain

$$i(B(n,2)) \le \frac{4}{n-\log n+2}.$$

Lemma 4 implies

$$|\mathcal{B}(n,\ell,0)| \le q^{\ell-1}(\ell(q-1)+q).$$

Hence, for $m \ge 2\log n$ we have $\ell \le n - 2\log n$, and for $n \ge q$,

$$|\mathcal{B}(n,\ell,0)| \le \frac{q^n((n-2\log n+1)q)}{n^2q} \le \frac{q^n}{2}.$$

Therefore,

$$i(B(n,q)) \le \frac{2q}{n-2\log_q n+1}. \quad \triangle$$

10. VERTEX-ISOPERIMETRIC NUMBER

In [9] there was given the the following upper bound for the vertex-isoperimetric number:

$$i_v(B(n,q)) \le \frac{2\sqrt{q}\pi}{(n+1)\sqrt{1-((2q\pi)/(n+1))^2}}.$$

In [10] it was improved as follows:

$$i_v(B(n,q)) \le \frac{4}{n-2},$$

for $n \ge 9$.

Consider a basic set $\mathcal{B}(n,\ell,1)$, where $\ell+m+1=n$:

$$N(\mathcal{B}(n,\ell,1)) = \mathcal{X}^\ell 10^m \cup 0^m 1\mathcal{X}^\ell.$$

Then

$$|N(\mathcal{B}(n,\ell,1))| = 2^\ell + |0^m 1\mathcal{X}^{\ell-m-1}0\mathcal{X}^m| + |0^m 1\mathcal{X}^{\ell-m-1}1\mathcal{X}^m|.$$

Therefore,

$$|N(\mathcal{B}(n,\ell,1))| = 2^\ell + 2^{\ell-1} + 2^{\ell-m-1}(2^m - 1) = 2^{\ell+1} - 2^{\ell-m-1}.$$

From bounds on the cardinality of basic sets, we have

$$|\mathcal{B}(n,\ell,1)| \geq 2^\ell(\ell+1) - 2^{\ell-m-1}(\ell-m+1)(\ell-m-1).$$

Hence,

$$\frac{|N(\mathcal{B}(n,\ell,1))|}{|\mathcal{B}(n,\ell,1)|} \leq \frac{2^{\ell+1} - 2^{\ell-m-1}}{2^\ell(\ell+1) - 2^{\ell-m-1}(\ell-m+1)(\ell-m-1)}.$$

Put $m = 2\log n$; then $\ell = n - 2\log n - 1$, and we obtain as $n \to \infty$

$$\frac{|N(\mathcal{B}(n,\ell,1))|}{|\mathcal{B}(n,\ell,1)|} \leq \frac{2}{n}(1+o(1)).$$

Clearly,

$$|\mathcal{B}(n,\ell,1)| \leq 2^\ell(\ell+1).$$

Therefore, for $m \geq \log n$ we get $\ell \leq n - \log n - 1$, and

$$|\mathcal{B}(n,\ell,1)| \leq \frac{2^n(n-\log n)}{2n} \leq \frac{2^n}{2}.$$

Hence, as $n \to \infty$, we get

$$i_v(B(n,2)) \leq \frac{2}{n}(1+o(1)).$$

Theorem 9. *The vertex-isoperimetric number of the de Bruijn graph $B(n,q)$ satisfies the following inequality as $n \to \infty$:*

$$i_v(B(n,q)) \leq \frac{q+2}{qn}(1+o(1)).$$

Proof. For the binary case, this is already proved above. For a q-ary case, we again consider a basic set $\mathcal{B}(n,\ell,1)$ with $\ell + m + 1 = n$:

$$N(\mathcal{B}(n,\ell,1)) = \mathcal{X}^\ell\overline{\mathcal{X}}0^m \cup 0^m\overline{\mathcal{X}}\mathcal{X}^\ell,$$

where $\overline{\mathcal{X}}$ denotes any nonzero element.

Then

$$|N(\mathcal{B}(n,\ell,1))| = q^\ell(q-1) + |0^m\overline{\mathcal{X}}\mathcal{X}^{\ell-m-1}0\mathcal{X}^m| + |0^m\overline{\mathcal{X}}\mathcal{X}^{\ell-m-1}\overline{\mathcal{X}}(\mathcal{X}^m \setminus 0^m)|.$$

Hence,

$$|N(\mathcal{B}(n,\ell,1))| = q^\ell(q-1) + 2q^{\ell-1}(q-1) - q^{\ell-m-1}(q-1).$$

Put $m = 2\log n$; then, as in the binary case, one can check that for large n we have $|\mathcal{B}(n,\ell,1)| \leq \frac{q^n}{2}$ and

$$i_v(B(n,q)) \leq \frac{q+2}{qn}(1+o(1)). \quad \triangle$$

11. SHADOWS FROM $\mathcal{X}^k$ TO $\mathcal{X}^n$

Definition 3. A sequence $x^n = (x_1, x_2, \ldots, x_n)$ is an *n-subword* of $y^k = (y_1, y_2, \ldots, y_k)$ if there exists i, $i \in \{0, 1, \ldots, k-n\}$, such that

$$y_{i+1} = x_1, \quad y_{i+2} = x_2, \quad \ldots, \quad y_{i+n} = x_n.$$

Equivalently: x^n is an *n-subword* of y^k if there exist a^i and b^{k-n-i} such that $y^k = a^i x^n b^{k-n-i}$, where $i \in \{0, 1, \ldots, k-n\}$.

The shadow of y^k is the set of all its n-subwords:

$$\mathrm{shad}_{k,n}(y^k) = \{x^n :\ x^n \text{ is an } n\text{-subword of } y^k\}.$$

Now for any subset $A \subset \mathcal{X}^k$ we define its shadow

$$\mathrm{shad}_{k,n}(A) = \bigcup_{a^k \in A} \mathrm{shad}_{k,n}(a^k).$$

For fixed n and k we are interested in the function

$$\triangle_{k,n}(q, N) = \min\{|\mathrm{shad}_{k,n}(A)| :\ A \subset \mathcal{X}^k,\ |A| = N\}.$$

The up-shadow of a sequence x^n is the following set:

$$\text{up-shad}(x^n) = \{y^k :\ x^n \text{ is an } n\text{-subword of } y^k\}.$$

Now for any set $B \subset \mathcal{X}^n$ we define its up-shadow

$$\text{up-shad}(B) = \bigcup_{b^n \in B} \text{up-shad}(b^n).$$

For fixed n and k we are interested in the function

$$\nabla(M) = \min\{|\text{up-shad}(B)| :\ B \subset \mathcal{X}^n,\ |B| = M\}.$$

Let $v = k - n$. For any $\ell \ge r \ge v$ such that $m + r > \ell$ (or $k = \ell + m + r > 2\ell$), we have

$$\mathrm{shad}_{k,n} \mathcal{B}(k, \ell, r) = \mathcal{B}(k - v, \ell, r - v).$$

Hence, we have the following result.

Theorem 10. *For $N = q^{\ell+v} + q^{\ell+v-1}(\ell - v)(q-1)$ and $k = \ell + m + v > 2\ell \ge 2v$ we have*

$$\frac{1}{q^v} N \le \triangle_{k,n}(q, N) \le \frac{1}{q^v}\left(1 + \frac{v}{\ell - v + 1}\right) N.$$

REFERENCES

1. Katona, G., A Theorem of Finite Sets, *Theory of Graphs (Proc. Colloq. on Graph Theory, Tihany, Hungary, 1966)*, Erdős, P. and Katona, G., Eds., New York: Academic, 1968, pp. 187–207.

2. Kruskal, J.B., The Number of Simplices in a Complex, *Mathematical Optimization Techniques (Proc. Sympos. on Mathematical Optimization Techniques, Santa Monica, CA, 1960)*, Bellman, R., Ed., Berkeley: Univ. of California Press, 1963, pp. 251–278.

3. Lindström, B.A. and Zetterström, H.O., A Combinatorial Problem in the k-adic Number Systems, *Proc. Amer. Math. Soc.*, 1967, vol. 18, no. 1, pp. 166–170.

4. Schützenberger, M.P., *A Characteristic Property of Certain Polynomials of E.F. Moore and C.E. Shannon*, RLE Quarterly Progress Report no. 55, MIT Research Laboratory of Electronics, 1959, pp. 117–118.

5. Ahlswede, R. and Cai, N., Shadows and Isoperimetry under the Sequence-Subsequence Relation, *Combinatorica*, 1997, vol. 17, no. 1, pp. 11–29.

6. Chowdhury, A. and Patkós, B., Shadows and Intersections in Vector Spaces, *J. Combin. Theory, Ser. A*, 2010, vol. 117, no. 8, pp. 1095–1106.

7. Mohar, B., Isoperimetric Numbers of Graphs, *J. Combin. Theory, Ser. B*, 1989, vol. 47, no. 3, pp. 274–291.

8. Bezrukov, S.L., Edge Isoperimetric Problems on Graphs, *Graph Theory and Combinatorial Biology*, Lovasz, L., Gyarfas, A., Katona, G.O.H., Recski, A., and Szekely, L., Eds., Budapest: János Bolyai Math. Soc., 1999, pp. 157–197.

9. Delorme, C. and Tillich, J-P., The Spectrum of de Bruijn and Kautz Graphs, *European J. Combin.*, 1998, vol. 19, no. 3, pp. 307–319.

10. Bultermann, J., A New Upper Bound for the Isoperimetric Numbers of de Bruijn Networks, *Appl. Math. Lett.*, 1997, vol. 10, no. 6, pp. 97–100.