2–WAY COMMUNICATION COMPLEXITY OF SUM–TYPE FUNCTIONS FOR ONE PROCESSOR TO BE INFORMED

Rudolf Ahlswede and Ning Cai

Universität Bielefeld
Fakultät für Mathematik
Postfach 100131
33501 Bielefeld
Germany

Abstract.

The asymmetric 2–way communication complexity of a function is a measure of the minimal amount of information required to be communicated between two parties in order for one of them to compute the value of the function at the inputs supplied by the parties. We provide rather sharp lower bounds for this quantity in terms of the rank of a certain matrix transform of the function. For several sum–type functions such as the Hamming, Lee or Taxi metrics they are even tight. We emphasize that for this class of functions the familiar $\log$ rank of the function tables gives in general a poor lower bound.

keywords and phrases

Communication Complexity, sum type functions, exponential transform.

1. Introduction.

In this paper we consider the problem of bounding the amount of information that must be communicated between two parties that cooperate to compute a function. We consider a class of functions, which may be rather special from a practical point of view. On the other hand our modest goal made a theoretical analysis possible, which led to a new bounding technique and thus gives also a better understanding of the subject on a large scale. Previous work in [4] and [6] examined the computation of sum–type functions of the form $S_n(x^n, y^n) = \sum_{t=1}^{n} f(x_t, y_t)$, where a party $P_{\mathcal{X}}$ contributes an input $x^n = (x_1, \dots, x_n)$, and $P_{\mathcal{Y}}$ contributes an input $y^n = (y_1, \dots, y_n)$, and both parties seek to compute the value of S_n . In this paper, we extend these results in two directions. First, we consider the case where the function f is replaced by a sequence of functions f_t , and second, we consider protocols in which only $P_{\mathcal{Y}}$ is required to compute the function.

Specifically, suppose that we are given sequences $(\mathcal{X}_t)_{t=1}^{\infty}$, $(\mathcal{Y}_t)_{t=1}^{\infty}$ of finite sets and a sequence $(f_t)_{t=1}^{\infty}$ of functions $f_t : \mathcal{X}_t \times \mathcal{Y}_t \to G$, where G is an abelian group (in this paper the additive groups on $\mathbb{R}$ or the eydic groups of integers $\mod p$ for any positive integer p). The associated sum–type function $S_n : \mathcal{X}^n \times \mathcal{Y}^n \to G$ is defined by

$$S_n(x^n, y^n) = \sum_{t=1}^{n} f_t(x_t, y_t) \tag{1.1}$$

for all $x^n = (x_1, \dots, x_n) \in \mathcal{X}^n$ and $y^n = (y_1, \dots, y_n) \in \mathcal{Y}^n$. Typical examples are distance functions such as the Hamming distance.

We introduce now our complexity measure, which we denote by $C(S_n, 1 \leftrightarrow 2^+)$. As usual a person (or processor) $P_{\mathcal{X}}$ observes output x^n and another person $P_{\mathcal{Y}}$ observes output y^n . They agree in advance on a protocol Q for transmitting alternatively strings of bits to each other. At the end of this exchange $P_{\mathcal{Y}}$ must be able to calculate $S_n(x^n, y^n)$. If $\ell_Q(x^n, y^n)$ is the number of bits exchanged for inputs x^n and y^n , then

$$L(Q) = \max_{x^n \in \mathcal{X}^n, y^n \in \mathcal{Y}^n} \ell_Q(x^n, y^n) \tag{1.2}$$

is the (worst case) length of the protocol Q . Let $\mathcal{Q}_{S_n}$ denote the set of all protocols for S_n . Then we define the 2–way communication complexity with respect to an informed $P_{\mathcal{Y}}$ by

$$C(S_n; 1 \leftrightarrow 2^+) = \min_{Q \in \mathcal{Q}_{S_n}} L(Q) \ .$$

Earlier we considered the larger quantity $C(S_n; 1 \leftrightarrow 2)$ based on protocols which enabled both processors to calculate $S_n(x^n, y^n)$.

In order to get a feeling for the behaviour of $C(S_n; 1 \leftrightarrow 2^+)$ we state now for comparison known bounds for $C(S_n; 1 \leftrightarrow 2)$ and we also discuss an example. We thus learn that the lower bound for $C(f; 1 \leftrightarrow 2^+)$ of [2] (see (1.9) below) is not very good. An analysis of this phenomenon led to the new lower bound of this paper.

For a function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ and associated function table $M_f = \big(f(x,y)\big)_{x\in\mathcal{X}, y\in\mathcal{Y}}$ Yao introduced the decomposition number D_f and Mehlhorn and Schmidt considered the class of matrices $(\Delta_z)_{z\in Z}$, where

$$\Delta_z(x,y) = \begin{cases} 1 & \text{if } M_f(x,y) = z \\ 0 & \text{otherwise,} \end{cases}$$

with a rank $r_{\mathbb{F}}(M_f)$ over any field $\mathbb{F}$, which is defined by

$$r_{\mathbb{F}}(M_f) = \sum_{z\in Z} \text{rank}_{\mathbb{F}}(\Delta_z) \ . \tag{1.3}$$

Yao's inequality, in the improved form of Papadimitriou and Sipser [3], states

$$C(f; 1 \leftrightarrow 2) \geq \log_2 D_f \tag{1.4}$$

and Mehlhorn and Schmidt showed that

$$D_f \geq r_{\mathbb{F}}(M_f) \ . \tag{1.5}$$

Quite general and sharp bounds for $C(S_n; 1 \leftrightarrow 2)$ were derived in [4] via a 4–words inequality. For other classes of sum–type functions in many cases even exact results were obtained by Tamm by evaluating $r_{\mathbb{F}}(S_n)$ for the lower bound and by using the trivial upper bound $C(f; 1 \leftrightarrow 2) \leq \lceil \log |\mathcal{X}| \rceil + \lceil \log |\mathcal{Z}| \rceil$.

Now notice that in the inequality (1.4) and also the weaker

$$C(f; 1 \leftrightarrow 2) \geq \log r_{\mathbb{F}}(M_f) \tag{1.6}$$

we cannot replace $C(f; 1 \leftrightarrow 2)$ by $C(f; 1 \leftrightarrow 2^+)$. Inspection of the following example shows that

	0	1	2
0	0	1	2
1	0	1	2
2	4	5	0
3	4	5	0

$D_f = 6$, $r_{\mathbb{R}}(M_f) = 2 + 1 + 1 + 1 + 1 = 6$ and thus $C(f; 1 \leftrightarrow 2) \geq 3$. Actually, with 1 bit $P_{\mathcal{X}}$ can tell $P_{\mathcal{Y}}$ whether $x \in \{0,1\}$ or $x \in \{2,3\}$ and then $P_{\mathcal{Y}}$ can calculate $f(x,y)$. With 2 additional bits $P_{\mathcal{Y}}$ can inform $P_{\mathcal{X}}$ about y . Thus we have $C(f; 1 \leftrightarrow 2) = 3$. However, using only the first step of this protocol we see that $C(f; 1 \leftrightarrow 2^+) = 1$.

On the other hand, there are the relations

$$C(S_n; 1 \leftrightarrow 2^+) + \log \|S_n\| \geq C(S_n; 1 \leftrightarrow 2) \geq C(S_n; 1 \leftrightarrow 2^+), \tag{1.7}$$

where $\|S_n\|$ denotes the cardinality of the range of S_n .

Therefore, in some cases, such as the case $f_t : \mathcal{X}_t \times \mathcal{Y}_t \to \mathbb{N} \cup \{0\}$ with

$$\left| \bigcup_{t=1}^{\infty} \{ f_t(x_t, y_t) : (x_t, y_t) \in \mathcal{X}_t \times \mathcal{Y}_t \} \right| = b < \infty,$$

we can derive from $\|S_n\| \leq 0(n)$ and (1.7)

$$\lim_{n \to \infty} \frac{1}{n} \big(C(S_n; 1 \leftrightarrow 2) - C(S_n; 1 \leftrightarrow 2^+) \big) = 0 \ . \tag{1.8}$$

This holds in the particular case that all of the f_t are identical.

In these cases the bounds for one complexity measure apply asymptotically to the other. In [2] there is also for $\mathcal{Z}$ with field structure $\mathbb{F}$ the bound

$$C(f; 1 \leftrightarrow 2^+) \geq \log \operatorname{rank}_{\mathbb{F}}(M_f) \ . \tag{1.9}$$

In the example above $\log \operatorname{rank}_{\mathbb{R}}(M_f) = 1$ is a sharp bound. However, (1.9) is a very poor bound for most sum–type functions, because their rank is only linearly increasing in n (see [5]) and the row and column numbers of the matrices increase exponentially in n . The results in Section 3 show how bad the linear bound really is. From [7] we know that the Kronecker product has exponentially increasing rank. Our idea now was to connect sum–type functions S_n to functions of "product–type" via an exponential transform of M_{S_n} . This gives excellent, often even exact, results. Moreover, this quantity often can be evaluated, because it factorizes (see Lemma 2 below).

After the idea is there, the mathematics almost takes care of itself. The Main Theorem in Section 2 has a short proof. Several consequences, including exact results for distance functions, are derived in Section 3. Finally, in Section 4 we show how the approach of [4] fits into the new frame.

2. Exponential transform, Kronecker product and sum–type functions.

We analyse sum–type functions by looking at their exponential transform.

Formally, for any matrix $M = (m_{ij})_{i\in I, j\in J}$ with entries in $\mathbb{R}$ the exponential transform of M is defined as

$$\operatorname{Exp}(M, z) = \left(z^{m_{ij}}\right)_{i\in I, j\in J} \tag{2.1}$$

where $z \in \mathbb{C}$, the field of complex numbers, and we choose the principal branch at every opportunity.

We introduce two rank functions, which are associated with $\mathbb{R}$ resp. the p elements cyclic group of integers modulo p . In space saving notation the first quantity is

$$\mathcal{R}\mathrm{ank}_{\mathbb{R}}(M) = \max_{z\in\mathbb{R}} \operatorname{rank}_{\mathbb{R}}\big(\operatorname{Exp}(M, z)\big) \tag{2.2}$$

and the second quantity is

$$\mathcal{R}\mathrm{ank}_{p}(M) = \operatorname{rank}_{\mathbb{C}}\big(\operatorname{Exp}(M, e^{2\pi i/p}\big) \ . \tag{2.3}$$

Sum–type functions $S_n = \sum_{t=1}^{n} f_t$ suggest the outer product of matrices as a basic structure for them.

For two vectors $\vec{u} = (u_1, \ldots, u_\ell)$ and $\vec{v} = (v_1, \ldots, v_m)$, $\vec{u} \circ \vec{v}$ is defined as (an $\ell \times m$ vector) $\vec{w} = (u_1 + v_1, u_1 + v_2, \ldots, u_1 + v_m, u_2 + v_1, \ldots, u_2 + v_m, \ldots, u_e + v_m)$, that is, the (i,j)–th component is $u_i + v_j$ (see [5]). Now the sum–type outer product of 2 matrices

$$U = \begin{pmatrix} \vec{u}_1 \\ \vdots \\ \vec{u}_r \end{pmatrix}, V = \begin{pmatrix} \vec{v}_1 \\ \vdots \\ \vec{v}_s \end{pmatrix} \quad \text{is defined by} \quad U \circ V = \begin{pmatrix} \vec{u}_1 \circ \vec{v}_1 \\ \vec{u}_1 \circ \vec{v}_2 \\ \vdots \\ \vec{u}_r \circ \vec{v}_s \end{pmatrix} .$$

These definitions imply directly a first basic fact.

Lemma 1. *For* $S_n = \sum_{t=1}^{n} f_t$

$$M_{S_n} = M_{f_1} \circ M_{f_2} \circ \cdots \circ M_{f_n} \ .$$

A second key observation is, that for two matrices U and V

$$\operatorname{Exp}(U \circ V, z) = \operatorname{Exp}(U, z) \bigotimes \operatorname{Exp}(V, z), \tag{2.4}$$

where $\bigotimes$ denotes the operation “Kronecker product”. Inductively our next result follows.

Lemma 2.

$$Exp\,(M_{S_n}, z) = \bigotimes_{t=1}^{n} Exp\,(M_{f_t}, z) \ \ .$$

Now, for any positive number $z \neq 1$ the map τ_z defined by

$$\tau_z : f \to z^f \tag{2.5}$$

for all real–valued functions on a fixed domain is a bijection and therefore for real–valued $f_t (t = 1, 2, \dots)$

$$C(S_n; 1 \leftrightarrow 2^+) = C(z^{S_n}; 1 \leftrightarrow 2^+) \ . \tag{2.6}$$

Since obviously by the definitions

$$\mathrm{rank}_{\mathbb{R}} M_{z^{S_n}} = \mathrm{rank}_{\mathbb{R}} \mathrm{Exp}\,(M_{S_n}, z)\,,$$

and since $\mathrm{rank}_{\mathbb{R}}$ is multiplicative in the Kronecker product of matrices, using also Lemma 2 we get

Lemma 3.

$$rank_{\mathbb{R}} M_{z^{S_n}} = \prod_{t=1}^{n} rank_{\mathbb{R}} Exp\,(M_{f_t}, z) \ \ .$$

We are now prepared to derive our key result.

We have by (1.9) the inequality

$$C(f; 1 \leftrightarrow 2^+) \geq \lceil \log \mathrm{rank}_{\mathbb{R}} M_f \rceil \ . \tag{2.7}$$

Identity (2.6) and Lemmas 1,2,3 imply

$$C(S_n; 1 \leftrightarrow 2^+) \geq \sum_{t=1}^{n} \log \mathrm{rank}_{\mathbb{R}} \mathrm{Exp}(M_{f_t}, z) \ \text{ for } \ z \in \mathbb{R}^+ - \{1\} \ . \tag{2.8}$$

Also, with the choice $z = e^{2\pi i/p}$ we have

$$e^{S_n \cdot 2\pi i/p} = e^{(S_n \mod p) \cdot 2\pi i/p}$$

and therefore

$$C(S_n \mod p; 1 \leftrightarrow 2^+) \geq \sum_{t=1}^{n} \log \mathcal{R}\mathrm{ank}_p M_{f_t} \ . \tag{2.9}$$

We summarize our findings.

Main Theorem.

(a) For the functions $f_t : \mathcal{X}_t \times \mathcal{Y}_t \to \mathbb{R} \quad (t \in \mathbb{N})$ *with* $\mathcal{X}_t, \mathcal{Y}_t$ *finite we have for the sum–type function* $S_n = \sum_{t=1}^n f_t$

$$C(S_n; 1 \leftrightarrow 2^+) \geq \sum_{t=1}^n \log \mathcal{R}ank_{\mathbb{R}} M_{f_t}$$

(b) For the functions $f_t ; \mathcal{X}_t \times \mathcal{Y}_t \to \Gamma$ *, the ring of integers, we have for any positive integer* p

$$C(S_n \mod p; 1 \leftrightarrow 2^+) \geq \sum_{t=1}^n \log \mathcal{R}ank_p M_{f_t} \ .$$

3. Consequences of Main Theorem.

Our first Theorem covers all distance functions on sequence spaces such as the Hamming metric, the Lee metric, the Taxi metric etc.

Theorem 1. *If for all* $t \in \mathbb{N} \quad \mathcal{X}_t = \mathcal{Y}_t = \{0, 1, \dots, \alpha_t - 1)$ *and*

$$f_t(x,y) = \begin{cases} 0 & \text{if } x = y \\ > 0 & \text{if } x \neq y, \end{cases}$$

then

$$C(S_n; 1 \leftrightarrow 2^+) = \lceil \sum_{t=1}^n \log \alpha_t \rceil \ .$$

Proof. Recalling the definition (2.1) we see that

$$\mathcal{R}\text{ank}_{\mathbb{R}} M_{f_t} \geq \text{rank}_{\mathbb{R}} \text{Exp}(M_{f_t}, 0) = \alpha_t \ .$$

The Main Theorem implies therefore

$$C(S_n; 1 \leftrightarrow 2^+) \geq \lceil \sum_{t=1}^n \log \alpha_t \rceil \ .$$

The reverse inequality is also true, because $P_{\mathcal{X}}$ can encode any argument $x^n \in \mathcal{X}^n$ with $\lceil \log |\mathcal{X}| \rceil$ bits and send them to $P_{\mathcal{Y}}$, who then calculates $S_n(x^n, y^n)$.

Theorem 1 settles a case in which one–way communication is as good as two–way communication.

We choose now $G = \{0, 1, \dots, p-1\}$, that is, we consider $S_n \mod p$ and give precise results in some basic cases.

Theorem 2. *If for all* $t \in \mathbb{N}$ $\mathcal{X}_t = \mathcal{Y}_t$ *and all* f_t *are Hamming distances, that is,* $f_t(x,y) = \begin{cases} 0 \text{ if } x = y \\ 1 \text{ if } x \neq y \end{cases}$ *, then*

(a) $| \, C(S_n \mod 2; 1 \leftrightarrow 2^+) - \sum_{t \in T_n} \log |\mathcal{X}_t| \, | \leq \varepsilon$
where $T_n = \{t : 1 \leq t \leq n, |\mathcal{X}_t| > 2\}$,

$$\varepsilon = \begin{cases} 0 & \text{if } T_n = \{1, 2, \dots, n\} \\ 1 & \text{otherwise.} \end{cases}$$

In particular, if $T_n = \varnothing$ *, necessarily*

$$C(S_n \mod 2; 1 \leftrightarrow 2^+) = 1 \ .$$

(b) $C(S_n \mod p; 1 \leftrightarrow 2^+) = \lceil \sum_{t=1}^{n} \log |\mathcal{X}_t| \rceil$ *for* $p > 2$.

Proof.

a.) We determine ranks by computing $\det \operatorname{Exp}(M_{f_t}, z)$.

Since $M_{f_t} = \begin{pmatrix} 0 & 1 & & & \dots & & 1 \\ 1 & 0 & 1 & & \dots & & 1 \\ 1 & 1 & 0 & 1 & \dots & & 1 \\ \dots & \dots & \dots & \dots & \dots & & \\ 1 & 1 & & & \dots & 1 & 0 \end{pmatrix}$, we have

$$\det \operatorname{Exp}(M_{f_t}, z) = \begin{vmatrix} 1 & z & & \dots & & z \\ z & 1 & z & \dots & & z \\ \dots & \dots & \dots & \dots & \dots & \\ z & & & \dots & z & 1 \end{vmatrix}$$

$= \big(1 + (\alpha_t - 1)z\big)(1 - z)^{\alpha_t - 1}$ and therefore

$$\det \operatorname{Exp}(M_{f_t}, -1) = (1 - \alpha_t + 1)2^{\alpha_t - 1} = \begin{cases} 0 & \text{if } \alpha_t = 2 \\ \neq 0 & \text{if } \alpha_t > 2. \end{cases}$$

Furthermore, in case $\alpha_t = 2$ we have

$$\operatorname{rank}_{\mathbb{C}} \operatorname{Exp}(M_{f_t}, e^{\pi i}) = \operatorname{rank}_{\mathbb{C}} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = 1 \ .$$

By the Main Theorem we have

$$C(S_n \bmod 2; 1 \leftrightarrow 2^+) \geq \sum_{t \in T_n} \log |\mathcal{X}_t| + \varepsilon \ .$$

For the upper bound we use the following protocol. $P_{\mathcal{X}}$ sends for input $(x_1, \dots, x_n)$ the subword with letters in positions T_n and with one more bit he sends the value of $\displaystyle\sum_{t \in \{1,2,\dots,n\} \smallsetminus T_n} x_t \bmod 2$ to $P_{\mathcal{Y}}$.

$P_{\mathcal{Y}}$ can compute the parities of the Hamming distance for the letters in positions T_n and also outside T_n , because there $\alpha_t = 2$. He then just adds these parities mod 2 .

b.) By the previous calculation

$$\begin{aligned} \det \operatorname{Exp}\big(M_{f_t}, e^{2\pi i/p}\big) &= \big(1 - (\alpha_t - 1)e^{2\pi i/p}\big)(1 - e^{2\pi i/p})^{\alpha_t - 1} \\ &\neq 0 \ \text{ for } \ p > 2 \end{aligned}$$

and by the Main Theorem

$$C(S_n \bmod p; 1 \leftrightarrow 2^+) \geq \lceil \sum_{t=1}^{n} \log |\mathcal{X}_t| \rceil \ .$$

For the reverse inequality the protocol, in which $P_{\mathcal{X}}$ transmits his input, suffices.

Theorem 3. *If for all* $t \in \mathbb{N}$ $\mathcal{X}_t = \mathcal{Y}_t = \{0, 1, \ldots, \alpha_t - 1\}$ *and all* f_t *are Taxi metrics, that is,* $f_t(x, y) = |x - y|$ *, then*

(a) $C(S_n \mod 2; 1 \leftrightarrow 2^+) = 1$

(b) $C(S_n \mod p; 1 \leftrightarrow 2^+) = \sum_{t=1}^n \log |\mathcal{X}_t|$ *for* $p > 2$ *.*

Proof. (a) The protocol, in which $P_{\mathcal{X}}$ sends $\sum_{t=1}^n x_t \mod 2$ to $P_{\mathcal{Y}}$, does it (as in the special case of a Hamming distance over binary alphabets).

(b) Here we have $M_{f_t} = \begin{pmatrix} 0 & 1 & 2 & \ldots & \alpha_t - 1 \\ 1 & 0 & 1 & \ldots & \alpha_t - 2 \\ \vdots & & & & \vdots \\ \alpha_t - 1 & \alpha_t - 2 & & \ldots & 0 \end{pmatrix}$

$$\mathcal{T}_{\alpha_t} \triangleq \det\big(\mathrm{Exp}(M_{f_t}, z)\big) = \begin{vmatrix} 1 & z & z^2 & z^3 & \ldots & z^{\alpha_t - 1} \\ z & 1 & z & z^2 & \ldots & z^{\alpha_t - 2} \\ \vdots & & & & & \\ z^{\alpha_t - 1} & z^{\alpha_t - 2} & & & \ldots & 1 \end{vmatrix}$$

$$= \mathcal{T}_{\alpha_t - 1} - z^2 \mathcal{T}_{\alpha_t - 1} = (1 - z^2)\mathcal{T}_{\alpha_t - 1} \text{ and } \mathcal{T}_2 = 1 - z^2 \ .$$

For $z = e^{2\pi i/p}$ and $p > 2$ we have therefore $1 - z^2 \neq 0$ and thus $\mathrm{rank}_{\mathbb{C}} \mathrm{Exp}(M_{f_t} e^{2\pi i/p}) = |\mathcal{X}_t|$. The Main Theorem and the standard protocol give the result.

Theorem 4. *If for* $t \in \mathbb{N}$ $\mathcal{X}_t = \mathcal{Y}_t = \{0, 1, \ldots, \alpha_{t-1}\}$ *and* $f_t = \Lambda_t$ *, where* $\Lambda_t(x_t, y_t) = \min(x_t, y_t)$ *, then*

$$C(S_n \mod p; 1 \leftrightarrow 2^+) = \sum_{t=1}^n \log |\mathcal{X}_t|$$

for all $p \geq 2$ *.*

Proof.

$$M_{f_t} = \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 1 & \ldots & 1 & 1 \\ 0 & 1 & 2 & \ldots & 2 & 2 \\ \vdots & \vdots & \vdots & & & \\ 0 & 1 & 2 & \ldots & \alpha_t - 2 & \alpha_t - 2 \\ 0 & 1 & 2 & \ldots & \alpha_t - 2 & \alpha_t - 1 \end{pmatrix},$$

$$\det\big(\mathrm{Exp}(M_{f_t},z)\big) = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & z & z & \dots & z & z \\ 1 & z & z^2 & \dots & z^2 & z^2 \\ \vdots & \vdots & \vdots & & \vdots & \\ 1 & z & z^2 & \dots & z^{\alpha_t-2} & z^{\alpha_t-2} \\ 1 & z & z^2 & \dots & z^{\alpha_t-2} & z^{\alpha_t-1} \end{vmatrix} .$$

By subtracting the first row from the second one, the second one from the third, fourth and so on, then in the resulting matrix the third one from the fourth etc, we finally get

$$\begin{aligned} \det\big(\mathrm{Exp}(M_{f_t},z)\big) &= \begin{vmatrix} 1 & 1 & 1 & \dots & 1 & 1 \\ 0 & z-1 & z-1 & \dots & z-1 & z-1 \\ 0 & 0 & z^2-z & \dots & z^2-z & z^2-z \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & z^{\alpha_t-2}-z^{\alpha_t-3} & z^{\alpha_t-2}-z^{\alpha_t-2} \\ 0 & 0 & 0 & \dots & 0 & z^{\alpha_t-1}-z^{\alpha_t-2} \end{vmatrix} \\ &= \prod_{r=1}^{\alpha_t-1} \left(z^r - z^{r-1}\right) = \prod_{r=1}^{\alpha_t-1} (z-1) z^{r-1} . \end{aligned}$$

For $z = e^{2\pi i/p}$ the determinant is unequal to zero for all $p \geq 2$. Again, the Main Theorem and the standard protocol give the result.

Our last sum–type function is the inner product. We choose $\mathcal{X}_t = \mathcal{Y}_t = \{0, 1, \dots, \alpha - 1\}$, $f_t = f \ (t \in \mathbb{N})$, where $f(x,y) = x \cdot y$.

Then

$$M_f = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & \alpha-1 \\ 0 & 2 & \dots & 2(\alpha-1) \\ \vdots & & & \\ 0 & \alpha-1 & \dots & (\alpha-1)^2 \end{pmatrix} ,$$

$$\det\big(\mathrm{Exp}(M_f,z)\big) = \begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & z & z^2 & \dots & z^{\alpha-1} \\ 1 & z^2 & (z^2)^2 & \dots & (z^2)^{\alpha-1} \\ \vdots & & & & \vdots \\ 1 & z^{\alpha-1} & (z^{\alpha-1})^2 & \dots & (z^{\alpha-1})^{\alpha-1} \end{vmatrix} ,$$

which is a Vandermonde determinant

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \dots & a_1^{\alpha-1} \\ 1 & a_2 & a_2^2 & \dots & a_2^{\alpha-1} \\ \vdots & & & & \\ 1 & a_\alpha & a_\alpha^2 & \dots & a_\alpha^{\alpha-1} \end{vmatrix} = \prod_{m=2}^{\alpha} \prod_{\ell=1}^{m-1} (a_m - a_{m-\ell})$$

with $a_1 = 1, a_2 = z, a_3 = z^2, \dots, a_\alpha = z^{\alpha-1}$.

Its value is

$$\prod_{m=2}^{\alpha} \prod_{\ell=1}^{m-1} (z^{m-1} - z^{m-1-\ell}), \quad \text{which is}$$

unequal to zero for $z \in \mathbb{R} - \{0, +1, -1\}$.

$\mathcal{R}\mathrm{ank}_{\mathbb{R}}(M_f) = \alpha$ and with the Main Theorem we complete the proof of the following result.

Theorem 5. *If* $\mathcal{X}_t = \mathcal{Y}_t = \{0, 1, \dots, \alpha - 1\}$ *for* $t \in \mathbb{N}$ *and if* $S_n(x^n, y^n) = \sum_{t=1}^n x_t \cdot y_t$ *, then*

$$C(S_n; 1 \leftrightarrow 2^+) = \lceil n \log \alpha \rceil \ .$$

Remark. Notice that for the inner product $\mathrm{rank}_{\mathbb{R}} M_f = 1$. This shows how much better it is to work with $\mathrm{Exp}(M_f, z)$.

4. A new look at 4–words property and inequality.

In [4] we found a general 4–words inequality and used it for the analysis of $C(S_n; 1 \leftrightarrow 2)$, if $S_n(x^n, y^n) = \sum_{t=1}^n f(x_t, y_t)$ and f satisfies the 4–words property

$$f(x, y) + f(x', y') - f(x, y') - f(x', y) = 0 \tag{4.1}$$

for all $x, y \in \mathcal{X} = \mathcal{Y}$.

It is clear that (4.1) is equivalent to

$$z^{f(x,y)} \cdot z^{f(x',y')} = z^{f(x,y')} \cdot z^{f(x',y)} \tag{4.2}$$

for $x, y \in \mathcal{X}$.

Therefore a submatrix of M_f satisfying the 4–words property corresponds to a submatrix of $\mathrm{Exp}(M_f, z)$ with rank not exceeding 1. Therefore the 4–word inequality (Theorem 1 of [4]) can be expressed also in the following form:

If $N_1(M)$ is the maximal size (= number of rows times number of columns) of submatrices of matrix M , whose rank does not exceed 1, then

$$N_1\Big(\bigotimes_1^n M\Big) = N_1(M)^n \ . \tag{4.3}$$

The Decomposition Lemma of [4] also has a simple interpretation:

A $k \times \ell$ submatrix has rank smaller than 2 iff it can be expressed as a product of a $k \times 1$ and a $1 \times \ell$ matrix.

Acknowledgement.

We thank Boris Spieker for his critical comments on [10], the first version of this paper.

References.

[1] A. Yao, Some complexity questions related to distributive computing, in "Proceedings 11th Ann. ACM Sympos. Theory of Computing, 1979", 209–213.

[2] K. Mehlhorn and E.M. Schmidt, Las Vegas is better than determinism in VLSI and distributed computing, in "Proceedings 14th ACM STOC, 1982", 330–337.

[3] C.H. Papadimitriou and M. Sipser, Communication complexity, in "Proceedings 14th Ann. ACM Sympos. on Theory of Computing, 1982", 201–214.

[4] R. Ahlswede, N. Cai, and Z. Zhang, A general 4–words inequality with consequences for 2–way communication complexity, Advances in Applied Mathematics 10 (1989), 75–94.

[5] R. Brualdi, N. Cai, and V. Pless, Orphan structure of the first order Reed–Muller codes, Discrete Math. 102 (1992), 239–247.

[6] U. Tamm, On the communication complexity of sum–type functions invariant under translation, Preprint 91–016 SFB 343 "Diskrete Strukturen in der Mathematik", to appear in Information and Computation.

[7] R. Ahlswede and N. Cai, On communication complexity of vector–valued functions, to appear in IEEE Trans. on Information Theory.

[8] R. Ahlswede, On code pairs with specified Hamming distances, Colloquia Mathematica Societatis János Bolyai 52. Combinatorics, Eger (1987), 9–47.

[9] R. Ahlswede and N. Cai, Rank formulas for certain products of matrices, to appear in Applicable Algebra in Engineering, Communication and Computing 1993.

[10] R. Ahlswede and N. Cai, 2–way communication complexity of sum–type functions for one processor to be informed, Preprint 91–053 of SFB 343 "Diskrete Strukturen in der Mathematik", Bielefeld, July 1991.