stage (stage indexed by $u - 3$) of the hierarchy. The corresponding merging rule for the $(u - 1)$th stage is as follows: if there are an integral power of two of successive macro-shells with equal cardinality, these are merged into a single, larger macro-shell. One can also apply this rule successively several times. The number of successive times is denoted by S. The performance and complexity of this approach is shown in Tables I–III. These tables correspond to $S = 1, 2, 3$, and each table contains all the possible combinations of ℓ_i's, $i = 0,\cdots,7$. For example, the first row in each table means that: $(\ell_i,\ i = 0,\cdots,7) = (7,7,6,5,4,3,2,1)$ and the second row means that: $(\ell_i, i = 0,\cdots,7) = (6,6,5,4,3,2,2,2)$. The cases of special interest (good performance and low complexity) are underlined.

We have also examined: i) the case of $S = 0$, and ii) applying the nonuniform merging in the $(u - 1)$'th stage. In both cases the results were inferior to those presented here.

VI. Numerical Comparisons

A four state trellis diagram of [6] achieves $\gamma_s = 0.95$ dB, CER$_s$ = 1.5. In [14], an example for $N = 64$ is given which needs 1440 multiply-adds (assuming a 16 bit processor) and a memory of 1.5 kilo-bytes to achieve a tradeoff point with $\gamma_s = 1.15$ dB, CER$_2$ = 1.5

For a given CER$_s$ by appropriately choosing the merging parameters, we achieve nearly all of the shaping gain possible using a small amount of memory (refer to Table IV). Computation of the optimum γ_s is based on 3.

Table IV can be compared to Table V, which shows the method applied when an equal number of points is used in the macro-shells at each stage (this becomes the method discussed in [11]). The cases of special interest are underlined. The present schemes offer a reduction in complexity by a factor of 5 to 10.

VII. Summary and Conclusions

We have presented efficient addressing schemes based on partitioning the subconstellations into nonuniform shaping macro-shells of integer bit rate. The corresponding shaping performance is computed using the weight distribution of an optimally shaped constellation. As an example of performance in a 32-D space, we use about 0.8 k-bytes of memory to achieve trade-off points very close to the optimum performance. It seems that this is the simplest known method to achieve shaping gains in the order of 1.0 dB. Note that this method needs only a small number of table lookups and no arithmetic operation is needed.

References

[1] G. D. Forney, Jr. and L. F. Wei, "Multidimensional constellations—Part I: Introduction, figures of merit, and generalized cross constellations," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 877–892, Aug. 1989.

[2] J. H. Conway and N. J. A. Sloane, "A fast encoding method for lattice codes and quantizers," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 106–109, Jan. 1985.

[3] L. F. Wei, "Trellis coded modulation with multidimensional constellations," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 483–501, July 1987.

[4] G. D. Forney, Jr., "Multidimensional constellations—Part II: Voronoi constellations," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 941–958, Aug. 1989.

[5] A. R. Calderbank and L. H. Ozarow, "Nonequiprobable signaling on the Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 36, pp. 726–740, July 1990.

[6] G. D. Forney, Jr., "Trellis shaping," *IEEE Trans. Inform. Theory*, vol. 38, pp. 281–300, Mar. 1992.

[7] G. R. Lang and F. M. Longstaff, "A leech lattice modem," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 968–973, Aug. 1989.

[8] F. R. Kschischang and S. Pasupathy, "Optimal nonuniform signaling for Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 39, May 1993.

[9] J. R. Livingston, "Shaping using variable-size regions," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1347–1353, July 1992.

[10] A. R. Calderbank and M. Klimesh, "Balanced codes and nonequiprobable signaling," *IEEE Trans. Inform. Theory*, vol. IT-38, pp. 1119–1122, May 1992.

[11] A. K. Khandani and P. Kabal, "Shaping multi-dimensional signal spaces—Part II: Shell-addressed constellations," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1809–1819, Nov. 1993.

[12] F. R. Kschischang and S. Pasupathy, "Optimal shaping properties of the truncated polydisc," *IEEE Trans. Inform. Theory*, vol. IT-40, pp. 892–903, May 1994.

[13] F. R. Kschischang, "Shaping and coding gain criteria in signal constellation design," Ph.D. dissertation, Toronto Univ., Toronto, Ont., Canada, June 1991.

[14] R. Laroia, N. Farvardin, and S. A. Tretter "On optimal shaping of multi-dimensional constellations," *IEEE Trans. Inform. Theory*, vol. IT-40, pp. 1044–1056, July 1994.

[15] A. K. Khandani and P. Kabal, "Shaping multi-dimensional signal spaces—Part I: Optimum shaping, shell mapping," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1799–1808, Nov. 1993.

On Communication Complexity of Vector-Valued Functions

Rudolf Ahlswede and Ning Cai

***Abstract*—New upper and lower bounds on the two-way communication complexity of abstract functions $g: \mathscr{X} \times \mathscr{Y} \to \mathscr{Z}$ give tight bounds, when applied to vector-valued functions $f^n = (f_1,\cdots,f_n): \mathscr{X}^n \times \mathscr{Y}^n \to \mathscr{Z}^n$, if the alphabets are small. For the set-intersection function, an optimal protocol is presented. It is based on a simple new idea applicable also to abstract functions. The two-way communication complexities of all other Boolean functions are also determined. The results are extended to meets in abstract lattices and to a probabilistic model.**

***Index Terms*—Two-way communication complexity, vector-valued functions, Kronecker product, prefix codes, correlated sources, rank, alternating partitions.**

I. Introduction

Let $\mathscr{X}$, $\mathscr{Y}$, and $\mathscr{Z}$ be finite sets. For any function $f: \mathscr{X} \times \mathscr{Y} \to \mathscr{Z}$, we consider the (vector-valued) functions $f^n: \mathscr{X}^n \times \mathscr{Y}^n \to \mathscr{Z}^n$ defined by

$$f^n(x^n, y^n) = (f(x_1, y_1), \cdots, f(x_n, y_n)), \tag{1.1}$$

for $x^n = (x_1, \cdots, x_n) \in \mathscr{X}^n$ and $y^n = (y_1, \cdots, y_n) \in \mathscr{Y}^n$ and study their two-way communication complexity $C(f^n; 1 \leftrightarrow 2)$; that is, the minimal number of bits which need to be exchanged for any argument (x^n, y^n) between a person $P_{\mathscr{X}}$ knowing x^n and a person $P_{\mathscr{Y}}$ knowing y^n so that both can calculate $f^n(x^n, y^n)$.

We also consider cases of nonidentical component functions; that is, we are given sequences $(\mathscr{X}_t)_{t=1}^{\infty}$, $(\mathscr{Y}_t)_{t=1}^{\infty}$, and $(\mathscr{Z}_t)_{t=1}^{\infty}$ of finite sets, a sequence $(f_t)_{t=1}^{\infty}$ of functions $f_t: \mathscr{X}_t \times \mathscr{Y}_t \to \mathscr{Z}_t$, and

Manuscript received December 30, 1991; revised September 1, 1993. This paper was presented at the 1991 IEEE International Symposium on Information Theory, Budapest, Hungary, June 24–28.

The authors are with Universität Bielefeld, Fakultät für Mathematik, Postfach 100131, 33501 Bielefeld, Germany.

IEEE Log Number 9406009.

stage (stage indexed by $u - 3$) of the hierarchy. The corresponding merging rule for the $(u - 1)$th stage is as follows: if there are an integral power of two of successive macro-shells with equal cardinality, these are merged into a single, larger macro-shell. One can also apply this rule successively several times. The number of successive times is denoted by S. The performance and complexity of this approach is shown in Tables I–III. These tables correspond to $S = 1, 2, 3$, and each table contains all the possible combinations of ℓ_i's, $i = 0, \cdots, 7$. For example, the first row in each table means that: $(\ell_i,\ i = 0, \cdots, 7) = (7, 7, 6, 5, 4, 3, 2, 1)$ and the second row means that: $(\ell_i,\ i = 0, \cdots, 7) = (6, 6, 5, 4, 3, 2, 2, 2)$. The cases of special interest (good performance and low complexity) are underlined.

We have also examined: i) the case of $S = 0$, and ii) applying the nonuniform merging in the $(u - 1)$'th stage. In both cases the results were inferior to those presented here.

VI. Numerical Comparisons

A four state trellis diagram of [6] achieves $\gamma_s = 0.95$ dB, $\mathrm{CER}_s = 1.5$. In [14], an example for $N = 64$ is given which needs 1440 multiply-adds (assuming a 16 bit processor) and a memory of 1.5 kilo-bytes to achieve a tradeoff point with $\gamma_s = 1.15$ dB, $\mathrm{CER}_2 = 1.5$

For a given CER_s by appropriately choosing the merging parameters, we achieve nearly all of the shaping gain possible using a small amount of memory (refer to Table IV). Computation of the optimum γ_s is based on 3.

Table IV can be compared to Table V, which shows the method applied when an equal number of points is used in the macro-shells at each stage (this becomes the method discussed in [11]). The cases of special interest are underlined. The present schemes offer a reduction in complexity by a factor of 5 to 10.

VII. Summary and Conclusions

We have presented efficient addressing schemes based on partitioning the subconstellations into nonuniform shaping macro-shells of integer bit rate. The corresponding shaping performance is computed using the weight distribution of an optimally shaped constellation. As an example of performance in a 32-D space, we use about 0.8 k-bytes of memory to achieve trade-off points very close to the optimum performance. It seems that this is the simplest known method to achieve shaping gains in the order of 1.0 dB. Note that this method needs only a small number of table lookups and no arithmetic operation is needed.

References

[1] G. D. Forney, Jr. and L. F. Wei, "Multidimensional constellations—Part I: Introduction, figures of merit, and generalized cross constellations," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 877–892, Aug. 1989.

[2] J. H. Conway and N. J. A. Sloane, "A fast encoding method for lattice codes and quantizers," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 106–109, Jan. 1985.

[3] L. F. Wei, "Trellis coded modulation with multidimensional constellations," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 483–501, July 1987.

[4] G. D. Forney, Jr., "Multidimensional constellations—Part II: Voronoi constellations," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 941–958, Aug. 1989.

[5] A. R. Calderbank and L. H. Ozarow, "Nonequiprobable signaling on the Gaussian channel," *IEEE Trans. Inform. Theory*, vol. 36, pp. 726–740, July 1990.

[6] G. D. Forney, Jr., "Trellis shaping," *IEEE Trans. Inform. Theory*, vol. 38, pp. 281–300, Mar. 1992.

[7] G. R. Lang and F. M. Longstaff, "A leech lattice modem, *Select. Areas Commun.*, vol. 7, pp. 968–973, Aug. 1989.

[8] F. R. Kschischang and S. Pasupathy, "Optimal nonunifo ing for Gaussian channels," *IEEE Trans. Inform. Theor* May 1993.

[9] J. R. Livingston, "Shaping using variable-size regions," *IE Inform. Theory*, vol. 38, pp. 1347–1353, July 1992.

[10] A. R. Calderbank and M. Klimesh, "Balanced c nonequiprobable signaling," *IEEE Trans. Inform. Theor* 38, pp. 1119–1122, May 1992.

[11] A. K. Khandani and P. Kabal, "Shaping multi-dimensio spaces—Part II: Shell-addressed constellations," *IEEE form. Theory*, vol. 39, pp. 1809–1819, Nov. 1993.

[12] F. R. Kschischang and S. Pasupathy, "Optimal shaping of the truncated polydisc," *IEEE Trans. Inform. Theory*, pp. 892–903, May 1994.

[13] F. R. Kschischang, "Shaping and coding gain criteria constellation design," Ph.D. dissertation, Toronto Univ., Ont., Canada, June 1991.

[14] R. Laroia, N. Farvardin, and S. A. Tretter "On optimal multi-dimensional constellations," *IEEE Trans. Inform* vol. IT-40, pp. 1044–1056, July 1994.

[15] A. K. Khandani and P. Kabal, "Shaping multi-dimensio spaces—Part I: Optimum shaping, shell mapping," *IE Inform. Theory*, vol. 39, pp. 1799–1808, Nov. 1993.

On Communication Complexity of Vector-Valued Functions

Rudolf Ahlswede and Ning Cai

***Abstract*—New upper and lower bounds on the two-way co tion complexity of abstract functions $g: \mathscr{X} \times \mathscr{Y} \to \mathscr{Z}$ give tigh when applied to vector-valued functions $\mathbf{f}^n = (\mathbf{f}_1, \cdots, \mathbf{f}_n): \mathscr{X}^n$ $\mathscr{Z}^n$, if the alphabets are small. For the set-intersection fu optimal protocol is presented. It is based on a simple new ide ble also to abstract functions. The two-way communication co of all other Boolean functions are also determined. The re extended to meets in abstract lattices and to a probabilistic m**

***Index Terms*—Two-way communication complexity, vect functions, Kronecker product, prefix codes, correlated sou alternating partitions.**

I. Introduction

Let $\mathscr{X}$, $\mathscr{Y}$, and $\mathscr{Z}$ be finite sets. For any function $f: \mathscr{X}$ $\mathscr{Z}$, we consider the (vector-valued) functions $f^n: \mathscr{X}^n \times \mathscr{Y}$ defined by

$$f^n(x^n, y^n) = (f(x_1, y_1), \cdots, f(x_n, y_n)),$$

for $x^n = (x_1, \cdots, x_n) \in \mathscr{X}^n$ and $y^n = (y_1, \cdots, y_n) \in \mathscr{Y}^n$ a their two-way communication complexity $C(f^n; 1 \leftrightarrow 2)$: the minimal number of bits which need to be exchange argument (x^n, y^n) between a person $P_{\mathscr{X}}$ knowing x person $P_{\mathscr{Y}}$ knowing y^n so that both can calculate $f^n(x$

We also consider cases of nonidentical component f that is, we are given sequences $(\mathscr{X}_t)_{t=1}^{\infty}$, $(\mathscr{Y}_t)_{t=1}^{\infty}$, and (finite sets, a sequence $(f_t)_{t=1}^{\infty}$ of functions $f_t: \mathscr{X}_t \times \mathscr{Y}_t \to$

Manuscript received December 30, 1991; revised Septembe This paper was presented at the 1991 IEEE International Symp Information Theory, Budapest, Hungary, June 24–28.

The authors are with Universität Bielefeld, Fakultät für M Postfach 100131, 33501 Bielefeld, Germany.

IEEE Log Number 9406009.

the vector-valued functions $f^n: \mathscr{X}^n \times \mathscr{Y}^n \to \mathscr{Z}^n$, where $\mathscr{X}^n = \prod_{t=1}^n \mathscr{X}_t$, $\mathscr{Y}^n = \prod_{t=1}^n \mathscr{Y}_t$, $\mathscr{Z}^n = \prod_{t=1}^n \mathscr{Z}_t$, and $f^n(x^n, y^n) = (f_1(x_1, y_1), \cdots, f_n(x_n, y_n))$.

In particular, we sometimes assume that

$$\mathscr{X}_t = \mathscr{Y}_t = \mathscr{Z}_t = \{0, 1, \cdots, \alpha_t - 1\}. \tag{1.2}$$

We denote the set of all protocols for calculating f^n by $\mathscr{Q}_{f^n}$ and for any protocol $Q \in Q_{f^n}$, $l_Q(x^n, y^n)$ is the number of bits exchanged for inputs x^n and y^n. In the worst case, we get

$$L(Q) \triangleq \max_{x^n \in \mathscr{X}^n, y^n \in \mathscr{Y}^n} l_Q(x^n, y^n), \tag{1.3}$$

the length of the protocol Q. In this terminology,

$$C(f^n; 1 \leftrightarrow 2) = \min_{Q \in \mathscr{Q}_{f^n}} L(Q). \tag{1.4}$$

For abstract functions $g: \mathscr{X} \times \mathscr{Y} \to \mathscr{Z}$ with $\mathscr{X}, \mathscr{Y}, \mathscr{Z}$ finite, Yao [1] introduced a k-decomposition as a partition $\mathscr{P} = \{U_1 \times V_1, \cdots, U_k \times V_k\}$ of $\mathscr{X} \times \mathscr{Y}$ into g-monochromatic rectangles. Those are rectangles $U \times V$ ($U \subset \mathscr{X}$, $V \subset \mathscr{Y}$) on which g is constant. Let $\mathscr{D}_g$ be the set of all such decompositions. For the decomposition number

$$D(g) \triangleq \min_{\mathscr{P} \in \mathscr{D}_g} |\mathscr{P}|, \tag{1.5}$$

Yao's inequality (in the improved form of [9]) states that

$$C(g; 1 \leftrightarrow 2) \geq \log_2 D(g). \tag{1.6}$$

The function table (or matrix) associated with g shall be denoted by M_g. We refer to the k-decompositions above also as k-decompositions of M_g. Mehlhorn and Schmidt [2] gave an important lower bound for $D(g)$. They associated with M_g the class of matrices $\{\Delta_z: z \in \mathscr{Z}\}$, where

$$\Delta_z(x, y) = \begin{cases} 1, & \text{iff } M_g(x, y) = z, \\ 0, & \text{otherwise,} \end{cases} \tag{1.7}$$

and established the inequality

$$D(g) \geq \sum_{z \in \mathscr{Z}} \operatorname{rank}_{\mathbb{F}}(\Delta_z), \tag{1.8}$$

where the rank is taken over any field $\mathbb{F}$. Here we always choose the field of rationals and therefore omit the index $\mathbb{F}$.

There is a simple but useful lower bound on $D(g)$. We call the set $\{(x_1, y_1), \cdots, (x_i, y_i)\}$ an independent set of Δ_z if $\Delta_z(x_t, y_t) = 1$ for $t = 1, 2, \cdots, i$, but no two members of the set are in the same monochromatic rectangle of Δ_z. The maximum cardinality of such a set is the independence number $\operatorname{ind}(\Delta_z)$. Clearly,

$$D(g) \geq \sum_z \operatorname{ind}(\Delta_z) \triangleq I(g). \tag{1.9}$$

The matrix $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$ with independence number 4 and rank 3 is an example for "$\operatorname{rank}(\Delta_z) \ngeq \operatorname{ind}(\Delta_z)$".

For the information of the reader, we mention that other authors (for example [11]) speak of fooling sets and sizes instead of independent sets and independent numbers.

There are three seemingly basic observations in the present paper.

1) *The protocol* $\bar{Q}$: For the set-intersection function $\wedge^n$ (Boolean "and"), the decomposition number is 3^n (Lemma 1 in Section II) and we found a protocol $\bar{Q}$ whose length achieves this bound (Theorem 1 in Section II). Other Boolean functions are easy to analyze. Moreover, all functions $f^n = (f, \cdots, f)$ with $|\mathscr{X}| = |\mathscr{Y}| = 2$, $|\mathscr{Z}| \leq 4$ fall into four classes, which can be reduced to Boolean cases (Corollary 1 in Section III). Here and also for the analysis of meets in abstract lattices, variations of the basic protocol $\bar{Q}$ are used (Theorems 7, 8, 9 in Section IX).

2) *General upper bound:* Partitions arising in protocols are more special than k-decompositions. For a function g, we denote the class of those "protocol"-generated partitions by $\mathscr{A}_g$ and introduce the protocol partition number

$$D^*(g) = \min_{\mathscr{P} \in \mathscr{A}_g} |\mathscr{P}|. \tag{1.10}$$

Whereas obviously

$$C(g; 1 \leftrightarrow 2) \geq \log D^*(g), \tag{1.11}$$

we found an equally simple upper bound involving the depth $d(\mathscr{P})$ (defined in Section IV):

$$C(g; 1 \leftrightarrow 2) \leq \min_{\mathscr{P} \in \mathscr{A}_g} (\log |\mathscr{P}| + d(\mathscr{P})). \tag{1.12}$$

Explicit definitions, the result (Theorem 2), and proof can be found in Section IV. Consequences for vector-valued functions, in particular with constant alphabets, gives Theorem 3 in Section V. Theorem 4 in Section VI states a probabilistic version of (1.12). Also in a probabilistic vector-values setting (for so-called correlated sources), there is an asymptotic result (Theorem 5, Section VII).

3) *Lower bound for vector-valued functions:* The independence number I and the rank function show multiplicative behavior under the Kronecker product (Lemmas 2, 3 in Section VIII). Our new observation is that the bounds can be combined on a letter-by-letter basis, yielding a bound better than the maximum of the two original bounds (Theorem 6 in Section VIII). Finally, this bound along with our upper bound (1.12) enables us to determine $C(f^n; 1 \leftrightarrow 2)$ within $o(n)$ for small alphabets. The precise results are stated in Theorem 10 in Section X.

II. Exact Solutions for all Boolean Functions

In the Boolean case, that is, $\mathscr{X} = \mathscr{Y} = \mathscr{Z} = \{0, 1\}$, one readily verifies that every Boolean function can be transformed by exchanges of 0 and 1 into one of the following functions (defined by tables):

$$B_1: \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 0 \end{array}, \qquad B_2: \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix},$$

$$B_3: \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad B_4: \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Here B_1 is a constant function, B_2 is the projection on the first argument, B_4 is the symmetric difference Δ, and finally B_3 stands for the logical "and" $\wedge$, which is also called the intersection function.

Now, obviously $C(B_1^n; 1 \leftrightarrow 2) = 0$ and $C(B_2^n; 1 \leftrightarrow 2) = n$, because $P_{\mathscr{X}}$ has to inform $P_{\mathscr{Y}}$ about x^n and this also suffices. Furthermore, $C(B_4^n; 1 \leftrightarrow 2) = 2n$ because $P_{\mathscr{X}}$ (resp., $P_{\mathscr{Y}}$) can recover y^n (resp., x^n) from $\Delta(x^n, y^n)$ and x^n (resp., y^n). The analysis of B_3^n is less obvious, but the answer is again "smooth." We summarize our findings.

Lemma 1: For the decomposition number, we have $D(B_3^n) = 3^n$.

Theorem 1: For the four types of Boolean functions, we have, for the two-way communication complexities,

$$C(B_i^n; 1 \leftrightarrow 2) = \lceil n \log i \rceil, \quad \text{for } i = 1, 2, 3, 4. \tag{2.1}$$

Proof of Lemma 1: For fixed z^n and $x^n \geq z^n$, that is, $x_t \geq z_t$ for $t = 1, 2, \cdots, n$, consider the sets

$$S(x^n, z^n) = \{(x^n, y^n): x^n \wedge y^n = z^n\}. \tag{2.2}$$

the vector-valued functions f^n: $\mathcal{X}^n \times \mathcal{Y}^n \to \mathcal{Z}^n$, where $\mathcal{X}^n = \prod_{t=1}^n \mathcal{X}_t$, $\mathcal{Y}^n = \prod_{t=1}^n \mathcal{Y}_t$, $\mathcal{Z}^n = \prod_{t=1}^n \mathcal{Z}_t$, and $f^n(x^n, y^n) = (f_1(x_1, y_1), \cdots, f_n(x_n, y_n))$.

In particular, we sometimes assume that

$$\mathcal{X}_t = \mathcal{Y}_t = \mathcal{Z}_t = \{0, 1, \cdots, \alpha_t - 1\}. \tag{1.2}$$

We denote the set of all protocols for calculating f^n by $\mathcal{Q}_{f^n}$ and for any protocol $Q \in Q_{f^n}$, $l_Q(x^n, y^n)$ is the number of bits exchanged for inputs x^n and y^n. In the worst case, we get

$$L(Q) \triangleq \max_{x^n \in \mathcal{X}^n,\, y^n \in \mathcal{Y}^n} l_Q(x^n, y^n), \tag{1.3}$$

the length of the protocol Q. In this terminology,

$$C(f^n; 1 \leftrightarrow 2) = \min_{Q \in \mathcal{Q}_{f^n}} L(Q). \tag{1.4}$$

For abstract functions $g: \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ with $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ finite, Yao [1] introduced a k-decomposition as a partition $\mathcal{P} = \{U_1 \times V_1, \cdots, U_k \times V_k\}$ of $\mathcal{X} \times \mathcal{Y}$ into g-monochromatic rectangles. Those are rectangles $U \times V$ ($U \subset \mathcal{X}$, $V \subset \mathcal{Y}$) on which g is constant. Let $\mathcal{D}_g$ be the set of all such decompositions. For the decomposition number

$$D(g) \triangleq \min_{\mathcal{P} \in \mathcal{D}_g} |\mathcal{P}|, \tag{1.5}$$

Yao's inequality (in the improved form of [9]) states that

$$C(g; 1 \leftrightarrow 2) \geq \log_2 D(g). \tag{1.6}$$

The function table (or matrix) associated with g shall be denoted by M_g. We refer to the k-decompositions above also as k-decompositions of M_g. Mehlhorn and Schmidt [2] gave an important lower bound for $D(g)$. They associated with M_g the class of matrices $\{\Delta_z: z \in \mathcal{Z}\}$, where

$$\Delta_z(x, y) = \begin{cases} 1, & \text{iff } M_g(x, y) = z, \\ 0, & \text{otherwise,} \end{cases} \tag{1.7}$$

and established the inequality

$$D(g) \geq \sum_{z \in \mathcal{Z}} \operatorname{rank}_{\mathbb{F}} (\Delta_z), \tag{1.8}$$

where the rank is taken over any field $\mathbb{F}$. Here we always choose the field of rationals and therefore omit the index $\mathbb{F}$.

There is a simple but useful lower bound on $D(g)$. We call the set $\{(x_1, y_1), \cdots, (x_i, y_i)\}$ an independent set of Δ_z if $\Delta_z(x_t, y_t) = 1$ for $t = 1, 2, \cdots, i$, but no two members of the set are in the same monochromatic rectangle of Δ_z. The maximum cardinality of such a set is the independence number $\operatorname{ind}(\Delta_z)$. Clearly,

$$D(g) \geq \sum_z \operatorname{ind}(\Delta_z) \triangleq I(g). \tag{1.9}$$

The matrix $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$ with independence number 4 and rank 3 is an example for "$\operatorname{rank}(\Delta_z) \ngeq \operatorname{ind}(\Delta_z)$".

For the information of the reader, we mention that other authors (for example [11]) speak of fooling sets and sizes instead of independent sets and independent numbers.

There are three seemingly basic observations in the present paper.

1) *The protocol* $\overline{Q}$: For the set-intersection function $\wedge^n$ (Boolean "and"), the decomposition number is 3^n (Lemma 1 in Section II) and we found a protocol $\overline{Q}$ whose length achieves this bound (Theorem 1 in Section II). Other Boolean functions are easy to analyze. Moreover, all functions $f^n = (f, \cdots, f)$ with $|\mathcal{X}| = |\mathcal{Y}| = 2$, $|\mathcal{Z}| \leq 4$ fall into four classes, which can be reduced to Boolean cases (Corollary 1 in Section III). Here and also for the analysis of meets in abstract lattices, variations of the basic protocol $\overline{Q}$ are used (Theorems 7, 8, 9 in Section IX).

2) *General upper bound:* Partitions arising in protocols are more special than k-decompositions. For a function g, we denote the class of those "protocol"-generated partitions by $\mathcal{A}_g$ and introduce the protocol partition number

$$D^*(g) = \min_{\mathcal{P} \in \mathcal{A}_g} |\mathcal{P}|. \tag{1.10}$$

Whereas obviously

$$C(g; 1 \leftrightarrow 2) \geq \log D^*(g), \tag{1.11}$$

we found an equally simple upper bound involving the depth $d(\mathcal{P})$ (defined in Section IV):

$$C(g; 1 \leftrightarrow 2) \leq \min_{\mathcal{P} \in \mathcal{A}_g} (\log |\mathcal{P}| + d(\mathcal{P})). \tag{1.12}$$

Explicit definitions, the result (Theorem 2), and proof can be found in Section IV. Consequences for vector-valued functions, in particular with constant alphabets, gives Theorem 3 in Section V. Theorem 4 in Section VI states a probabilistic version of (1.12). Also in a probabilistic vector-values setting (for so-called correlated sources), there is an asymptotic result (Theorem 5, Section VII).

3) *Lower bound for vector-valued functions:* The independence number I and the rank function show multiplicative behavior under the Kronecker product (Lemmas 2, 3 in Section VIII). Our new observation is that the bounds can be combined on a letter-by-letter basis, yielding a bound better than the maximum of the two original bounds (Theorem 6 in Section VIII). Finally, this bound along with our upper bound (1.12) enables us to determine $C(f^n; 1 \leftrightarrow 2)$ within $o(n)$ for small alphabets. The precise results are stated in Theorem 10 in Section X.

II. Exact Solutions for all Boolean Functions

In the Boolean case, that is, $\mathcal{X} = \mathcal{Y} = \mathcal{Z} = \{0, 1\}$, one readily verifies that every Boolean function can be transformed by exchanges of 0 and 1 into one of the following functions (defined by tables):

$$B_1: \begin{array}{c|cc} & 0 & 1 \\ \hline 0 & 0 & 0 \\ 1 & 0 & 0 \end{array}, \qquad B_2: \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix},$$

$$B_3: \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad B_4: \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Here B_1 is a constant function, B_2 is the projection on the first argument, B_4 is the symmetric difference Δ, and finally B_3 stands for the logical "and" $\wedge$, which is also called the intersection function.

Now, obviously $C(B_1^n; 1 \leftrightarrow 2) = 0$ and $C(B_2^n; 1 \leftrightarrow 2) = n$, because $P_{\mathcal{X}}$ has to inform $P_{\mathcal{Y}}$ about x^n and this also suffices. Furthermore, $C(B_4^n; 1 \leftrightarrow 2) = 2n$ because $P_{\mathcal{X}}$ (resp., $P_{\mathcal{Y}}$) can recover y^n (resp., x^n) from $\Delta(x^n, y^n)$ and x^n (resp., y^n). The analysis of B_3^n is less obvious, but the answer is again "smooth." We summarize our findings.

Lemma 1: For the decomposition number, we have $D(B_3^n) = 3^n$.

Theorem 1: For the four types of Boolean functions, we have, for the two-way communication complexities,

$$C(B_i^n; 1 \leftrightarrow 2) = \lceil n \log i \rceil, \quad \text{for } i = 1, 2, 3, 4. \tag{2.1}$$

Proof of Lemma 1: For fixed z^n and $x^n \geq z^n$, that is, $x_t \geq z_t$ for $t = 1, 2, \cdots, n$, consider the sets

$$S(x^n, z^n) = \{(x^n, y^n): x^n \wedge y^n = z^n\}. \tag{2.2}$$

Then partition the set

$$S(z^n) = \bigcup_{x^n \geq z^n} S(x^n, z^n) \tag{2.3}$$

into 2^{n-k} monochromatic rectangles, when $k = w(z^n)$ counts the 1's in z^n. Since there are $\binom{n}{k}$ sequences z^n with weight $w(z^n) = k$, we can therefore partition the whole table into $\Sigma_{k=0}^{n}\binom{n}{k}2^{n-k} = 3^n$ monochromatic rectangles.

Conversely, it suffices to show that the set $S(z^n)$ requires $2^{n-w(z^n)}$ rectangles in any partition. But this is the case, because

$$\{(x^n, \overline{x^n}): x_t = 1, \text{if } z_t = 1,\ 1 \leq t \leq n\},$$

is an independent (fooling) set of size $2^{n-w(2^n)}$ in $S(z^n)$, if $\overline{x^n} = (1 - x_1, \cdots, 1 - x_n)$.

Proof of Theorem 1: After the discussion preceding Lemma 1, we need to consider only the function B_3^n. Using inequality (1.6), Lemma 1 implies

$$C(B_3^n; 1 \leftrightarrow 2) \geq \lceil n \log 3 \rceil. \tag{2.4}$$

Since the lower bound in (1.6) is based on prefix codes, it is natural to look at how close to the truth protocols based on such codes are in the present case. Surprisingly, these protocols give an exact bound.

To fix ideas, let us recall a trivial protocol: $\mathscr{P}_{\mathscr{X}}$ sends x and $\mathscr{P}_{\mathscr{Y}}$ sends $g(x, y)$. This shows that $C(g; 1 \leftrightarrow 2) \leq \lceil \log|\mathscr{X}| \rceil + \lceil \log|\mathscr{Z}| \rceil$. There is a smarter two-rounds protocol!

Knowing x, $\mathscr{P}_{\mathscr{X}}$ also knows $\{g(x, y): y \in \mathscr{Y}\}$. In case this set is large, the length $l(x)$ of the codeword for x should be small, and in case this set is small, x can get a longer codeword in order to minimize the worst-case total number of bits. $\mathscr{P}_{\mathscr{Y}}$ knows when to start transmission if $\mathscr{P}_{\mathscr{X}}$ uses a prefix code. In the present case, if $\mathscr{P}_{\mathscr{X}}$ describes x^n, then $\mathscr{P}_{\mathscr{Y}}$ must describe all y_t's of y^n where x_t is 1. The goal is to have a prefix code encoding x^n with $l(x^n)$ bits such that $L = \max_{x^n \in \mathscr{X}^n}(l(x^n) + w(x^n))$ is minimal. By Kraft's inequality, L must satisfy $\Sigma_{x^n} 2^{-(L-w(x^n))} \leq 1$. Since $\Sigma_{x^n} 2^{w(x^n)} = 3^n$, the minimal L equals $\lceil n \log 3 \rceil$.

III. Exact Solutions for the Cases $|\mathscr{X}| = |\mathscr{Y}| = 2$

In case $|\mathscr{Z}| \leq 2$, one readily verifies that the functions are equivalent to one of the Boolean functions $B_1, \cdots, B_4$. Also, in case $|\mathscr{Z}| = 4$ for an f taking on four values, $C(f_n; 1 \leftrightarrow 2)$ equals $2n$, because here any one of the processors can recover the other input from the value of f_n. Thus we are left with the case $\mathscr{Z} = \{0, 1, 2\}$ and without loss of generality (w.l.o.g.), with the tables

$$M_g: \begin{pmatrix} 0 & 0 \\ 2 & 1 \end{pmatrix}, \quad M_h: \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}.$$

With the map φ_j, which sends 2 to j and i to i otherwise, we can write $\Lambda = \varphi_0(g)$ and $\Delta = \varphi_1(h)$ and thus, by Theorem 1,

$$C(g_n; 1 \leftrightarrow 2) \geq C(B_3^n; 1 \leftrightarrow 2) \geq \lceil n \log 3 \rceil,$$

$$C(h_n; 1 \leftrightarrow 2) \geq C(B_2^n; 1 \leftrightarrow 2) \geq 2n.$$

On the other hand, $C(h^n; 1 \leftrightarrow 2) \leq 2n$ follows with the simple protocol, where $P_{\mathscr{X}}$ sends x^n and $P_{\mathscr{Y}}$ sends y^n.

The bound $C(g^n; 1 \leftrightarrow 2) \leq \lceil n \log 3 \rceil$ can be derived with the protocol $\overline{Q}$ for B_s^n in the proof of Theorem 1, because g_n is a function of x^n and of B_3^n because $P_{\mathscr{Y}}$ knows B_3^n after learning x^n from $P_{\mathscr{X}}$.

We summarize our findings.

Corollary 1: The functions $f^n = (f, \cdots, f)$ with $|\mathscr{X}| = |\mathscr{Y}| = 2$, $|\mathscr{Z}| \leq 4$, fall into classes $\mathscr{Z}_1, \mathscr{Z}_2, \mathscr{Z}_3, \mathscr{Z}_4$. The members in $\mathscr{Z}_i$ have the same two-way complexity as B_i.

IV. A General Upper Bound via Protocol-Generated Partitions

The present approach lifts the upper-bound proof of Theorem 1 to a general function $f: \mathscr{X} \times \mathscr{Y} \to \mathscr{Z}$ (all sets finite with table or matrix M_f). Whereas Yao considered the set $\mathscr{D}_f$ of all (abstract) partitions of M_f into monochromatic rectangles, here we use the set $\mathscr{A}_f$ of protocol-generated partitions of M_f into monochromatic rectangles. They are defined as follows. Suppose the partitions $\mathscr{P}_1 = \{A_1 \times B_1, \cdots, A_k \times B_k\}$ and $\mathscr{P}_2 = \{A'_1 \times B'_1, \cdots, A'_n \times B'_n\}$ divide a rectangle into subrectangles and, for all $A'_i \times B'_i$, there exist $i_1, \cdots, i_j$ such that $A'_i = A_{i_1} \cup A_{i_2} \cup \cdots \cup A_{i_j}$ and $B_{i_1} = \cdots = B_{i_j} = B'_i$; then we call $\mathscr{P}_1$ a row-refinement of $\mathscr{P}_2$ and write $\mathscr{P}_2 \triangleright_r \mathscr{P}_1$. The column-refinement is defined analogously. We say that a partition $\mathscr{P}$ of M_f into monochromatic rectangles is protocol generated, if there are partitions $\mathscr{P}_i$ ($i = 1, 2, \ldots, k$) with

$$\{M_f\} = \mathscr{P}_0 \triangleright_1 \mathscr{P}_1 \cdots \triangleright_k \mathscr{P}_k = \mathscr{P},$$

where $\triangleright_i$ means a row-refinement iff $\triangleright_{i-1}$ is a column-refinement. The minimal k of such a representation of $\mathscr{P}$ is called the depth $d(\mathscr{P})$.

Obviously, any $\mathscr{P} \in \mathscr{A}_f$ can be produced by a two-way communication protocol. A basic quantity is $D^*(f) = \min\{|\mathscr{P}|: \mathscr{P} \in \mathscr{A}_f\}$. The lower bound implicit in [9] is

$$C(f; 1 \leftrightarrow 2) \geq \log D^*(f). \tag{4.1}$$

Here is our simple upper bound.

Theorem 2: For all $\mathscr{P} \in \mathscr{A}_f$, there is a protocol Q with $L(Q) \leq \log|\mathscr{P}| + d(\mathscr{P})$. In particular,

$$C(f; 1 \leftrightarrow 2) \leq \min_{\mathscr{P} \in \mathscr{A}_f} (\log|\mathscr{P}| + d(\mathscr{P})).$$

Proof: We proceed by induction on d. If $d(\mathscr{P}) = k$, then there are partitions $\mathscr{P}_1, \cdots, \mathscr{P}_{k-1}$ with $\{M_f\} \triangleright_1 \mathscr{P}_1 \triangleright_2 \mathscr{P}_2 \triangleright \cdots \triangleright \mathscr{P}_{k-1} \triangleright_k \mathscr{P}$. Assuming w.l.o.g. that $\triangleright_1$ is a row-refinement, we have $\mathscr{P}_1 = \{A_i \times \mathscr{Y}\}_{i=1}^{m}$. Suppose now that $A_i \times \mathscr{Y}$ contains a_i subrectangles in $\mathscr{P}$. Then $|\mathscr{P}| = \Sigma_i a_i$ and the partition $\mathscr{P}_{(i)}$ of $A_i \times \mathscr{Y}$ (into the a_i subrectangles) induced by $\mathscr{P}$ has depth not greater than $d(\mathscr{P}) - 1$.

By the induction hypothesis for all $A_i \times \mathscr{Y}$, there are protocols $Q_{(i)}$ with

$$L(Q_{(i)}) \leq \log|\mathscr{P}_{(i)}| + d(\mathscr{P}_{(i)}) \leq \log a_i + d(\mathscr{P}) - 1. \tag{4.2}$$

Since $(a_1/\Sigma_j a_j, a_2/\Sigma_j a_j, \cdots, a_m/\Sigma_j a_j)$ is a probability distribution, by the Noiseless Coding Theorem there is a binary prefix-code $\{c_1, c_2, \cdots, c_m\}$ with length $(c_i) \leq \log \Sigma_j a_j/a_i + 1 = \log \Sigma_j a_j - \log a_i + 1$. With $b_i \triangleq \lfloor \log \Sigma_j a_j - \log a_i \rfloor + 1$ bits, $P_{\mathscr{X}}$ can send i, the label of $A_i \times \mathscr{Y}$, to $P_{\mathscr{Y}}$. By (4.2), $b_i + L(Q_{(i)}) \leq \log \Sigma_j a_j + d(\mathscr{P}) = \log|\mathscr{P}| + d(\mathscr{P})$.

V. Application of the Upper Bound to Vector-Valued Functions

The functions $f^n: \mathscr{X}^n \times \mathscr{Y}^n \to \mathscr{Z}^n$, $f^n = (f_1, f_2, \cdots, f_n)$ were defined in the Introduction. We shall apply the bound of Theorem 2. A simple observation is that, for partitions $\mathscr{P}_t = \{D_{t,j}\}_j \in \mathscr{A}_{f_t}$ ($t = 1, 2, \cdots, n$), we have

$$\mathscr{P}^n = \left\{\prod_{t=1}^{n} D_{t,j_t}\right\}_{(j_1, \cdots, j_n)} \in \mathscr{A}_{f^n}. \tag{5.1}$$

Furthermore, we show next that $d(\mathscr{P}^n)$ is controlled by

$$d(\mathscr{P}^n) \leq \max_{1 \leq t \leq n} \{2 \cdot \min(|\mathscr{X}_t|, |\mathscr{Y}_t|) - 2\} \triangleq \theta. \tag{5.2}$$

This is the case, because we can first cut M_{f^n} row-wise by cutting all component spaces for which the first round in their protocol partition are row-refinements. In the next step, there must be column-refinements in all components, etc. We thus reach $\mathscr{P}^n$ after at most $\theta + 1$ rounds.

The one extra round is not needed, because if $f_t = f$ $(t = 1, 2, \cdots, n)$, then Theorem 2 in conjunction with (5.1) and (5.2) imply the following "single-letter" upper bound.

Theorem 3: For $n \in \mathbb{N}$,

$$C(f^n; 1 \leftrightarrow 2) \leq \sum_{t=1}^{n} \log D^*(f_t) + \theta + 1.$$

If $f_t = f$ for $t = 1, 2, \cdots, n$, then

$$C(f^n; 1 \leftrightarrow 2) \leq n \log D^*(f) + \theta.$$

Even though $n \log(D^*(f))$ may be much larger than $\log D^*(f^n)$, we show in Section X that, quite surprisingly, this is not the case if one of the alphabets $\mathcal{X}$ or $\mathcal{Y}$ is smaller than 5.

VI. Application of the Upper Bound to a Probabilistic Model

Let (X, Y) be a pair of random variables with X taking values in $\mathcal{X}$, Y taking values in $\mathcal{Y}$, and joint distribution P_{XY}. For $f: \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ and a protocol Q, $l_Q(x, y)$ is the length of the protocol in the input (x, y). We are interested in the expected two-way communication complexity

$$\tilde{C}(f; 1 \leftrightarrow 2) \triangleq \min_{Q \in \mathcal{Q}_f} \mathbb{E} l_Q(X, Y).$$

For $\mathscr{P} = \{A_1, \cdots, A_m\} \in \mathscr{A}_f$, consider the entropy

$$H_{\mathscr{P}} \triangleq H(P_{XY}(A_1), \cdots, P_{XY}(A_m)).$$

Theorem 4:

$$\tilde{C}(f; 1 \leftrightarrow 2) \leq H_{\mathscr{P}} + d(\mathscr{P})$$

for all $\mathscr{P} \in \mathscr{A}_f$.

Proof: The derivation is almost the same as that of Theorem 2. The difference is as follows. When we employ the induction hypothesis for $\{A_i \times \mathcal{Y}: i = 1, 2, \cdots, m\}$, we give a protocol Q_i for the restriction of f on $A_i \times \mathcal{Y}$. This happens with probability $P_{XY}(A_i \times \mathcal{Y})$. Conditional on the event $A_i \times \mathcal{Y}$ for $\mathscr{P}_i$ (defined as in the proof of Theorem 2), by the induction hypothesis,

$$\mathbb{E} l_{Q_i} \leq H_{\mathscr{P}_i | A_i \times \mathcal{Y}} + d(\mathscr{P}) - 1.$$

We have to add the prefix-free coding of $\{A_i \times \mathcal{Y}\}_i$ with an average length not greater than $H(P_{XY}(A_1 \times \mathcal{Y}), \cdots, P_{XY}(A_m \times \mathcal{Y}))$. Now, using the grouping axiom of entropy,

$$H_{\mathscr{P}} = \sum_i P_{XY}(A_i \times \mathcal{Y}) H_{\mathscr{P}_i | A_i \times \mathcal{Y}} + H(P_{XY}(A_1 \times \mathcal{Y}), \cdots, P_{XY}(A_m \times \mathcal{Y})).$$

VII. An Asymptotic Result for Correlated Sources

We consider the vector-valued case with identical components, that is,

$$f^n = (f, \cdots, f).$$

Furthermore, we assume that $(X_t, Y_t)_{t=1}^{\infty}$ is a sequence of independent, identically distributed pairs of random variables.

Theorem 5: Under the preceding probabilistic assumptions,

$$\lim_{n \to \infty} \frac{1}{n} \tilde{C}(f^n; 1 \leftrightarrow 2) = \lim_{n \to \infty} \frac{1}{n} \min_{\mathscr{P} \in \mathscr{A}_{f^n}} H_{\mathscr{P}}.$$

Proof: For any $\epsilon > 0$, first choose $n(\epsilon)$ such that

$$\frac{1}{n(\epsilon)} \min_{\mathscr{P} \in \mathscr{A}_{f^{n(\epsilon)}}} H_{\mathscr{P}} \leq \frac{\epsilon}{2} + \underline{\lim}_{n \to \infty} \frac{1}{n} \min_{\mathscr{P} \in \mathscr{A}_{f^n}} H_{\mathscr{P}},$$

and then choose $m^*(\epsilon)$ and $\mathscr{P}^* = \{A_1, \cdots, A_{m^*(\epsilon)}\} \in \mathscr{A}_{f^{n(\epsilon)}}$ such that

$$\frac{1}{n(\epsilon)} H_{\mathscr{P}^*} \leq \epsilon + \underline{\lim}_{n \to \infty} \frac{1}{n} \min_{\mathscr{P} \in \mathscr{A}_{f^n}} H_{\mathscr{P}}. \tag{7.1}$$

Use now for any r, the partition $\mathscr{P}^{*(r)} = \{A_{j_1} \times \cdots \times A_{j_r}: 1 \leq j_1, \cdots, j_r \leq m^{*(\epsilon)}\}$ obtained as product from $\mathscr{P}^*$. Then $\mathscr{P}^{*(r)} \in \mathscr{A}_{f^{n(\epsilon)r}}$ and

$$H_{\mathscr{P}^{*(r)}} = r H_{\mathscr{P}^*}. \tag{7.2}$$

By Theorem 4, $\tilde{C}(f^{n(\epsilon)r}; 1 \leftrightarrow 2) \leq H_{\mathscr{P}^{*(r)}} + d(\mathscr{P}^{*(r)})$, and thus by (5.2) and (7.2),

$$\frac{1}{n(\epsilon)r} \tilde{C}(f^{n(\epsilon)r}; 1 \leftrightarrow 2) \leq \frac{1}{n(\epsilon)} H_{\mathscr{P}^*} + \frac{\theta}{n(\epsilon)r}. \tag{7.3}$$

Using the inequality (7.1), we continue with

$$\frac{1}{n(\epsilon)r} \tilde{C}(f^{n(\epsilon)r}; 1 \leftrightarrow 2) \leq \epsilon + \underline{\lim}_{n \to \infty} \frac{1}{n} \min_{\mathscr{P} \in \mathscr{A}_{f^n}} H_{\mathscr{P}} + \frac{\theta}{n(\epsilon)r}, \tag{7.4}$$

and

$$\overline{\lim}_{r \to \infty} \frac{1}{n(\epsilon)r} \tilde{C}(f^{n(\epsilon)r}; 1 \leftrightarrow 2) \leq \epsilon + \underline{\lim}_{n \to \infty} \frac{1}{n} \min_{\mathscr{P} \in \mathscr{A}_{f^n}} H_{\mathscr{P}}. \tag{7.5}$$

Since for $n_1 \leq n_2$,

$$\tilde{C}(f^{n_1}; 1 \leftrightarrow 2) \leq \tilde{C}(f^{n_2}; 1 \leftrightarrow 2), \tag{7.6}$$

then we also have, for $n \in [n(\epsilon)(r-1), n(\epsilon)r]$,

$$\frac{1}{n} \tilde{C}(f^n; 1 \leftrightarrow 2) \leq \frac{r}{r-1} \frac{1}{n(\epsilon)r} \tilde{C}(f^{n(\epsilon)r}; 1 \leftrightarrow 2). \tag{7.7}$$

Thus, by (7.5)–(7.7) and since ϵ was arbitrary,

$$\overline{\lim}_{n \to \infty} \frac{1}{n} \tilde{C}(f^n; 1 \leftrightarrow 2) \leq \underline{\lim}_{n \to \infty} \frac{1}{n} \min_{\mathscr{P} \in \mathscr{A}_{f^n}} H_{\mathscr{P}}. \tag{7.8}$$

The lower bound is immediate from the Noiseless Coding Theorem.

VIII. Lower Bounds for Vector-Valued Functions

Using Yao's bound $C(f^n; 1 \leftrightarrow 2) \geq \log D(f^n)$ and

$$D(f^n) \geq I(f^n) \geq \prod_{t=1}^{n} I(f_t), \tag{8.1}$$

we get the following.

Lemma 2: $C(f^n; 1 \leftrightarrow 2) \geq \sum_{t=1}^{n} \log I(f_t)$.

Recall now the definition of Δ_z in (1.7) and set

$$r(f_t) = \sum_{z \in \mathcal{Z}_t} \operatorname{rank}(\Delta_{z_t}). \tag{8.2}$$

Next observe that, in analogous notation,

$$\Delta_{z^n} = \bigotimes_{t=1}^{n} \Delta_{z_t}, \tag{8.3}$$

where $\otimes$ denotes the Kronecker product of matrices, i.e.,

$$\Delta_{z^n}(x^n, y^n) = \Delta_{z_1}(x_1, y_1) \Delta_{z_2}(x_2, y_2) \cdots \Delta_{z_n}(x_n, y_n).$$

Therefore, $\Sigma_{x^n \in \mathscr{Z}^n} z^n \Delta_{z^n} = \Sigma_{z^n \in \mathscr{Z}^n} z^n \otimes_{t=1}^n \Delta_{z_t}$, and

$$\begin{aligned} r(f^n) &= \sum_{z^n \in \mathscr{Z}^n} \operatorname{rank}(\Delta_{z^n}) = \sum_{z^n \in \mathscr{Z}^n} \operatorname{rank}\left(\bigotimes_{t=1}^n \Delta_{z_t}\right) \\ &= \sum_{z^n \in \mathscr{Z}^n} \prod_{t=1}^n \operatorname{rank}(\Delta_{z_t}) = \prod_{t=1}^n \sum_{z_t \in \mathscr{Z}_t} \operatorname{rank}(\Delta_{z_t}) \\ &= \prod_{t=1}^n r(f_t). \end{aligned}$$

We have derived an identity.

Lemma 3: $r(f^n) = \prod_{t=1}^n r(f_t)$.

From (1.6), (1.8), and this identity follows a useful lower bound.

Lemma 4: $C(f^n; 1 \leftrightarrow 2) \geq \Sigma_{t=1}^n \log r(f_t)$.

Especially for the stationary case $f^n = (f, \cdots, f)$, the bound depends only on one component f. In information theory, this is called a "single-letter characterization."

The bounds in Lemma 2 and Lemma 4 can be combined into one bound,

$$C(f^n; 1 \leftrightarrow 2) \geq \max\left(\sum_{t=1}^n \log I(f_t), \sum_{t=1}^n \log r(f_t)\right). \quad (8.4)$$

Whereas this bound is canonical and perhaps known, we now derive an improvement to this bound.

Using the independence numbers $\operatorname{ind}(\Delta_z)$, we can write

$$I(f_t) = \sum_{z_t \in \mathscr{Z}_t} \operatorname{ind}(\Delta_{z_t}). \quad (8.5)$$

We need the local maxima

$$j(\Delta_{z_t}) = \max(\operatorname{rank}(\Delta_{z_t}), \operatorname{ind}(\Delta_{z_t})), \quad (8.6)$$

and their sums

$$J(f_t) = \sum_{z_t \in \mathscr{Z}_t} j(\Delta_{z_t}).$$

Theorem 6:

$$C(f^n; 1 \leftrightarrow 2) \geq \log D(f^n) \geq \sum_{t=1}^n \log J(f_t). \quad (8.7)$$

Proof: Denote Yao's decomposition number of Δ_{z^n} by $\delta(\Delta_{z^n})$. Since

$$D(f^n) = \sum_{z^n} \delta(\Delta_{z^n}), \quad (8.8)$$

and since $\prod_{t=1}^n J(f_t) = \Sigma_{t=1}^n \Sigma_{z_t \in \mathscr{Z}_t} j(\Delta_{z_t}) = \Sigma_{z^n} \prod_{t=1}^n j(\Delta_{z_t})$, (8.7) follows by summation over z^n if we know that

$$\delta(\Delta_{z^n}) \geq \prod_{t=1}^n j(\Delta_{z_t}). \quad (8.9)$$

In order to verify this inequality, we first introduce an auxiliary matrix Δ_z^* as follows. Let Ω_z be an independent set of Δ_z with $|\Omega_z| = \operatorname{ind}(\Delta_z)$. Then set $\Delta_z^* = \Delta_z$ if $\operatorname{rank}(\Delta_z) \geq \operatorname{ind}(\Delta_z)$, and otherwise define

$$\Delta_z^*(x, y) = \begin{cases} 1, & \text{if } (x, y) \in \Omega_z, \\ 0, & \text{otherwise}. \end{cases}$$

If we can prove that

$$\delta(\Delta_{z^n}) \geq \delta\left(\bigotimes_{t=1}^n \Delta_{z_t}^*\right), \quad (8.10)$$

then we are done, because

$$\begin{aligned} \delta\left(\bigotimes_{t=1}^n \Delta_{z_t}^*\right) &\geq \prod_{t=1}^n \operatorname{rank}(\Delta_{z_t}^*) \\ &= \prod_{t=1}^n \max(\operatorname{rank}(\Delta_{z_t}), \operatorname{ind}(\Delta_{z_t})) \\ &= \prod_{t=1}^n j(\Delta_{z_t}). \end{aligned}$$

Now just observe that a monochromatic partition of $\Delta_{z_1} \times \Delta_{z_2}$ restricted to $\Delta_{z_1}^* \times \Delta_{z_2}$ (in case $\Delta_{z_1} \neq \Delta_{z_1}^*$) gives again a monochromatic partition, and thus $\delta(\Delta_{z_1} \otimes \Delta_{z_2}) \geq \delta(\Delta_{z_1}^* \otimes \Delta_{z_2})$. Finally, (8.10) follows by repetition of this argument.

IX. Generalizations of the Result on the Set-Intersection Problem to Meets in a Lattice

First consider the case of chains, that is, $\mathscr{X}_t = \mathscr{Y}_t = \mathscr{Z}_t = \{0, 1, 2, \cdots, \alpha_t - 1\}$ and

$$\Lambda^n(x^n, y^n) = (x_1 \wedge y_1, \cdots, x_n \wedge y_n). \quad (9.1)$$

Theorem 7:

$$\begin{aligned} \sum_t \log(2\alpha_t - 1) &\leq C(\Lambda^n; 1 \leftrightarrow 2) \\ &\leq \sum_t \log(2\alpha_t - 1) + 2 \max_t \alpha_t - 3. \end{aligned}$$

Remark: When $\alpha_t = 2$ for $t = 1, 2, \cdots, n$, this result is only slightly weaker than that of Theorem 1, the difference being one bit in the upper bound.

Proof: The lower bound follows from either Lemma 2 or Theorem 6. The upper bound is one bit better than what follows from Theorem 3. This one bit can be saved, because parts of the second to last partition have sizes which are powers of 2 (cf. proof of Theorem 1).

Suppose now that $\mathscr{L}$ is a finite lattice. Let $(l] = \{x: x \leq l\}$, $[l) = \{x: l \leq x\}$ for all $l \in \mathscr{L}$, and $S = \{(l, l'): l \leq l'\} \subset \mathscr{L}^2$. Clearly, $\Sigma_{l \in \mathscr{L}} |[l)| = \log |S|$.

Considering the family of monochromatic rectangles $\{(l, x): x \in [l)\}$, $l \in \mathscr{L}$, Theorem 2 yields

$$C(\Lambda; 1 \leftrightarrow 2) \leq \log \sum_{l \in \mathscr{L}} |[l)| + 1 = \log |S| + 1. \quad (9.2)$$

We use an additional structure to also get a lower bound.

The element $l^* \in \mathscr{L}$ is called a pseudocomplement of $l \in \mathscr{L}$ iff $l^* \wedge l = 0$ and $l \wedge x = 0$ implies $x \leq l^*$.

$\mathscr{L}$ is called pseudocomplemented if every element has a pseudocomplement. Notice that for pseudocomplemented $\mathscr{L}$, for any $l \in \mathscr{L}$, $[l)$ is also pseudocomplemented ($a \in [l)$ has a relative pseudocomplement $a^* \vee l$, where a^* is a pseudocomplement of a in $\mathscr{L}$).

Now the proof of Theorem 1 can be generalized (see also [12]) to yield the following.

Theorem 8: For a finite pseudocomplemented lattice,

$$\lceil \log |S| \rceil \leq C(\Lambda; 1 \leftrightarrow 2) \leq \lfloor \log |S| \rfloor + 1.$$

For general (not necessarily pseudocomplemented) lattices, we now derive a lower bound on $C(\Lambda; 1 \leftrightarrow 2)$ along the lines initiated by Hajnal, Maass, and Turán [3], who introduced the Möbius function to the study of two-way communication complexity for estimating the rank of the function table. In particular, we rely upon an idea of Lovász and Saks [4], who used the following result.

Theorem (Lindström [5] and Dowling and Wilson [6]): For the disjointness indicator $T_{\mathscr{L}}$ of a lattice $\mathscr{L}$, that is, a matrix whose rows and columns are labeled by the elements of $\mathscr{L}$ and $T_{\mathscr{L}}(l, k) = 1$ if l and k are disjoint, and $T_{\mathscr{L}}(l, k) = 0$ otherwise, the $\operatorname{rank}(T_{\mathscr{L}}) = |\{l \in \mathscr{L}: \mu(\phi, l) \neq 0\}|$, where μ is the Möbius inverse.

Consider now $\Lambda: \mathscr{L}^2 \to \mathscr{L}$ and for all $l \in \mathscr{L}$, $\Delta_l = T_{[l)}$. From $C(\Lambda; 1 \leftrightarrow 2) \geq \log r(\Lambda)$, it follows from the previous theorem that, for $S^* \triangleq \{(l, l'): \mu(l, l') \neq 0\} \subset S$,

$$C(\Lambda; 1 \leftrightarrow 2) \geq \log r(\Lambda) \geq \lceil \log |S^*| \rceil. \tag{9.3}$$

In summary, we thus have the following.

Theorem 9:

$$\lceil \log |S^*| \rceil \leq C(\Lambda; 1 \leftrightarrow 2) \leq \lfloor \log |S| \rfloor + 1.$$

For Boolean lattices, subspaces of finite fields, and the partition lattice, $|S^*| = |S|$. For multisets, $|S^*| \neq |S|$, but Theorem 7 shows that the lower bound in (9.3) is "almost tight."

This investigation has been continued in [12].

X. Sharp Results for Vector-Valued Functions over Small Alphabets on One Side

For small alphabets, Theorem 6 in Section VIII can now be used to derive a "single-letter" lower bound for the complexity $C(f^n; 1 \leftrightarrow 2)$, which asymptotically approaches the upper bound of Theorem 3. In this case, we have the surprising fact that equality holds in

$$D(f^n) \leq \prod_{t=1}^{n} D(f_t). \tag{10.1}$$

Theorem 10: Suppose that $|\mathscr{X}_t| \leq \min(4, |\mathscr{Y}_t|)$ for $t \in \mathbb{N}$, then

$$\text{(i)} \quad D^*(f_t) = D(f_t), \qquad \text{(ii)} \quad J(f_t) = D(f_t), \tag{10.2}$$

and

$$\sum_{t} \log D(f_t) + \theta + 1 \geq C(f^n; 1 \leftrightarrow 2) \geq \sum_{t=1}^{n} \log D(f_t). \tag{10.3}$$

Particularly, for $f^n = (f, \cdots, f)$,

$$n \log D(f) + \theta \geq C(f^n; 1 \leftrightarrow 2) \geq n \log D(f). \tag{10.4}$$

Proof: The upper bounds follow from Theorem 3 and identity (i) in (10.2). The lower bounds in (10.3) and (10.4) follow from Theorem 6 and the identity (ii) in (10.2), which we now prove.

Clearly, it suffices to show that

$$\delta(\Delta_z) = j(\Delta_z), \quad \text{for } z \in \mathscr{Z}. \tag{10.5}$$

If $\operatorname{rank}(\Delta_z) = |\mathscr{X}|$, then $\delta(\Delta_z) = |\mathscr{X}| = j(\Delta_z)$. We need to consider only those z for which $\operatorname{rank}(\Delta_z) < |\mathscr{X}|$. For then we can actually show that

$$\delta(\Delta_z) = \operatorname{ind}(\Delta_z), \tag{10.6}$$

and thus *a fortiori*, (10.5) holds.

We can assume that there is no all-0 row or column and that there are no two identical rows or columns in Δ_z, because otherwise rows or columns can be removed without changing any of the parameters of interest.

Case $|\mathscr{X}| = 3$: Since $\operatorname{rank}(\Delta_z) \leq 2$, w.l.og. the first row is a linear combination of the other two. Therefore, the matrix is of the form

$$\Delta_z = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and (10.6) holds.

Case $|\mathscr{X}| = 4$: The case $\operatorname{rank}(\Delta_z) \leq 2$ being obvious again, we assume that $\operatorname{rank}(\Delta_z) = 3$. We also assume that the last three rows in Δ_z are linearly independent. The following subcases arise:

a) The first row vector V_1 is a linear combination of two others, say V_2 and V_3. Thus, we must have

$$V_1 = V_2 + V_3. \tag{10.7}$$

b) $V_1 = \lambda_2 V_2 + \lambda_3 V_3 + \lambda_4 V_4$, $\lambda_i > 0$.

c) $\lambda_1 V_1 + \lambda_2 V_2 = \lambda_3 V_3 + \lambda_4 V_4$, $\lambda_i > 0$.

In case a), Δ_z has one of the forms

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ ? & ? & 0 & 1 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ ? & ? & ? & ? & 1 \end{pmatrix},$$

and in both cases, (10.6) holds.

In case b), there are the forms

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix},$$

and again (10.6) holds.

Finally, in case c), we have the forms

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

and (10.6) holds in all cases!

References

[1] A. Yao, "Some complexity questions related to distributive computing," in *Proc. 11th Annu. ACM Symp. Theory Comput.*, 1979, pp. 209–213.

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

[3] A. Hajnal, W. Maass, and G. Turán, "On the communication complexity of graph properties," in *Proc. 20th ACM STOC*, 1988, pp. 186–191.

[4] L. Lovász and M. Saks, "Lattices, Möbius functions and communication complexity," in *Proc. 29th IEEE FOCS*, 1988, pp. 81–90.

[5] B. Lindström, "Determinants on semilattices," *Proc. AMS*, vol. 20, pp. 207–208, 1969.

[6] T. Dowling and R. Wilson, "Whitney number inequalities for geometric lattices," *Proc. AMS*, vol. 47, pp. 504–512, 1975.

[7] R. Brualdi, N. Cai, and V. Pless, "Orphan structure of the first order Reed-Muller codes," *Discr. Math.*, vol. 102, pp. 239–247, 1992.

[8] U. Tamm, "On the communication complexity of sum-type functions invariant under translation," preprint 91-016 SFB 343, to appear in *Inf. Comput.*

[9] C. H. Papadimitriou and M. Sipser, "Communication complexity," in *Proc. 14th Annu. ACM Symp. Theory Comput.*, 1982, pp. 201–214.

[10] R. Ahlswede, N. Cai, and Z. Zhang, "A general 4-words inequality with consequences for 2-way communication complexity," *Adv. Appl. Math.*, vol. 10, pp. 75–94, 1989.

[11] J. D. Ullman, *Computational Aspects of VLSI.* Rockville, Maryland, Computer Science Press, 1984.

[12] R. Ahlswede, N. Cai, and U. Tamm, "Communication complexity in lattices," *Appl. Math. Lett*, vol. 6, no. 6, pp. 53–58, 1993.