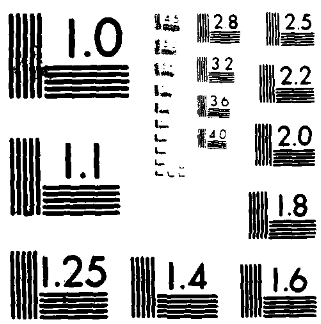
AD-A112 972 STANFORD UNIV CA DEPT OF COMPUTER SCIENCE F/G 12/1
THE LOWER BOUNDS ON THE ADDITIVE COMPLEXITY OF BILINEAR PROBLEM--ETC(U)
JUN 81 V Y PAN N00014-81-K-0269
UNCLASSIFIED STAN-CS-81-862 NL

1 OF 1
AD A
112972

END
DATE
FILMED
04-82
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

June 1981

Report. No. STAN-CS-81-862

The Lower Bounds on the Additive Complexity of Bilinear Problems in Terms of Some Algebraic Quantities

by

V. Ya. Pan

Research sponsored in part by

National Science Foundation
Office of Naval Research

Department of Computer Science

Stanford University
Stanford, CA 94305

The Lower Bounds on the Additive Complexity of Bilinear Problems in Terms of Some Algebraic Quantities

V. Ya. Pan*

The Institute of Advanced Study
Princeton, New Jersey 08540

and

Computer Science Department
Stanford University
Stanford, California 94305

Abstract. The lower bounds on the additive complexity of a bilinear problem are expressed in terms of the rank of the problem and also as a minimum number of elementary steps for the transformation of the identity matrix into a strongly regular one.

Key Words. Additive complexity, bilinear algorithms, tensor rank.

* Research supported in part by National Science Foundation grant MSC-8003347, Institute for Advanced Study; by National Science Foundation grant MCS-77-23738 and by Office of Naval Research connract N00014-81-K-0269. Reproduction in whole or in part is permitted for any purpose of the United States government.

As is well known, the basic part of the theory of algebraic computational complexity had been shaped by 1966; cf. [1,2,3]. In particular, until very recently the lower bounds on the additive complexity, $C(\pm)$, of intensively studied linear and bilinear arithmetic algorithms for arithmetic computational problems (such as DFT and matrix and polynomial multiplication, MM, PM) have relied on the active operation-basic substitution argument due to [1,2,3]; cf. also [4]. Consequently, those bounds have not exceeded D, the dimension of the problems that is the total number of input variables and outputs. In the present paper we consider another algebraic approach that generalizes the ingenious method of [5]. This enables us to reduce the problem to estimating the ranks of multidimensional tensors that we associate with the given computational problems. The successful solution of a similar problem in [6] gives some ground for optimism in the attempts to establish nonlinear lower bounds on $C(\pm)$ along this line. We also present another direction to attack the problem which reduces it to the study of a strong regularity of matrices; see Definition 2 and the Theorem below.

Notation. I, J, K are positive integers. $v_h = (\underline{V})_h$, $\mu_{js} = (\mu)_{js}$ are the entry h of a vector $\underline{V}$ and the entry (j, s) of a matrix μ, respectively. F is a field of constants. $\underline{X}$ is a vector of indeterminates, x_i, $i = 0, 1, \ldots, I-1$. $L(\underline{X}, F)$ is the set of all homogeneous linear forms of $x_0, \ldots, x_{I-1}$ with the coefficients from F.

Any $K \times J$ matrix, $\mu = \mu(\underline{X})$ with the entries from $L(\underline{X}, F)$ defines a bilinear arithmetic problem that is the set of bilinear forms $\{b_k(\underline{X}, \underline{Y})\}$ whose Y-coefficients form the matrix $\mu(\underline{X})$; cf. [7,8]. A bilinear arithmetic algorithm, A, that solves such a problem can be represented as a chain of matrices $\langle \mu(0), \mu(1), \ldots, \mu(C) \rangle$ (cf. [5,7,8]) such that $\mu(0)$ is the $J \times J$ identity matrix, μ is a submatrix of $\mu(C)$, each $\mu(q)$ is a $J \times (J+q)$ matrix such that

$$\mu(q+1) = (\mu(q) \mid \underline{V}(q+1)) \quad \text{for } q = 0, 1, \ldots, C-1, \tag{1}$$

where for all j either

$$(\underline{V}(q+1))_j = L(q)(\mu(q))_{js} \quad \text{for some } s = s(q) \le q + J \tag{2}$$

or

$$(\underline{V}(q+1))_j = (\mu(q))_{jp} + \delta(\mu(q))_{js} \\ \text{for some } p = p(q) \le q + J, \quad s = s(q) \le q + J. \tag{3}$$

In (3), $\delta = 1$ or $\delta = -1$. In (2) either $L(q) \in F$ or otherwise: $L(q) \in L(\underline{X}, F)$ and $(\mu(q))_{js} \in F$ for $s = s(q)$ and for all j. I[illegible] $C_A(\pm)$ designates the number of q such that (3) holds.

Definition 1 (cf. [9,10]). Given $P(\underline{X})$, a homogeneous polynomial in $x_0, \ldots, x_{I-1}$ of degree d, then $r(P(\underline{X}))$, the rank of $P(X)$, is the minimum integer $r \ge 0$ such that

$$P(\underline{X}) = \sum_{g=1}^{r} \prod_{h=1}^{d} r_{gh}(\underline{X}), \quad L_{gh}(\underline{X}) \in L(\underline{X}, F).$$

Let $D = D(M(\underline{X}))$ designate the set of all minors of a matrix $M(\underline{X})$ with the entries from $L(\underline{X}, F)$, $r(M) = \max_{m \in D} r(m)$. (We say that $r(M)$ is the rank of the bilinear problem associated with the matrix $M(\underline{X})$.) Then the next lemma is easily verified.

Lemma 1. *Equation (2) implies that $r(\mu(q+1)) = r(\mu(q))$, Equation (3) implies that $r(\mu(q+1)) \leq 2r(\mu(q))$.*

Corollary. *Given a bilinear algorithm, A (cf. (1)–(3)), for the bilinear problem defined by a matrix $\mu = \mu(\underline{X})$, then $C_A(\pm) \geq \log_2 r(\mu)$.*

Hence, $\log r(\mu) = \Omega(n \log n)$ for the general $n \times n$ Toëplitz matrix ($J = K = n$, $\mu_{jk} = x_{j-k+n-1}$, $j, k = 0, 1, \ldots, n-1$) would imply nonlinear lower bounds on the complexity of PM and DFT.

Remark. If $P_n(\underline{X}) = P_n(\underline{X}_1, \ldots, \underline{X}_n)$ is an n-linear form in n vectors of indeterminates, $\underline{X}_1, \ldots, \underline{X}_n$, then the polylinear rank, $R(P_n(\underline{X}))$, can be defined as the minimum integer R such that

$$P_n(\underline{X}) = \sum_{g=1}^{R} \prod_{h=1}^{n} L_{gh}(\underline{X}_h), \quad L_{gh}(\underline{X}_h) \in L(\underline{X}_h, F).$$

As is obvious, $R(P_n(\underline{X})) \geq r(P_n(\underline{X}))$. $R(P_2(\underline{X}_1, \underline{X}_2))$ equals the "usual" rank of the matrix of coefficients of the bilinear form $P_2(\underline{X}_1, \underline{X}_2)$. $R(P_3(\underline{X}_1, \underline{X}_2, \underline{X}_3))$ equals the multiplicative complexity of the three bilinear computational problems associated with $P_3(X_1, X_2, X_3)$, (cf. [11,12]). If $\mu(\underline{X})$ is an $n \times n$ matrix with row-vectors of indeterminates $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_n$, then $\log_2 R(\text{per } \mu(X)) \leq n$ (cf. [13]). Because of the latter estimate the inequality $\log_2 r(M) > n$ seems to be either false or very hard to prove even in the case of a general $n \times n$ matrix μ.

Despite the latter remark, we hope that the reader will be challenged to look for a better modification of the above approach and for new methods for establishing lower bounds on $C(\pm)$. Here is another example of natural approaches to this problem.

Definition 2. A matrix is strongly regular if it contains no singular submatrix. Given a $J \times s$ matrix μ and a field F then the elementary additive augmentation step consists of adding a new column-vector to μ which is a linear combination with the coefficients from F of two columns of μ. $C_{\pm}(J)$, the regularization number of order J is the minimum number of elementary additive augmentation steps required to transform the $J \times J$ identity matrix into a matrix that has a strongly regular $J \times J$ submatrix.

Theorem. *Let Y be the J-dimensional vector of indeterminates, μ be a $J \times J$ matrix over F that has a strongly regular $s \times s$ submatrix. Then the additive complexity of the evaluation of μY is at least $C_{\pm}(s)$.*

In particular, the general Töeplitz matrices are strongly regular. Hence any nonlinear lower bound on $C_{\pm}(s)$ would imply a nonlinear lower bound on $C(\pm)$ in the cases of PM and DFT.

I wish to thank Evelyn Laurent, IAS, and Phyllis Winkler, Stanford University, for typing this paper.

References

[1] E. G. Belaga, *Dokl. Akad. Nauk. SSSR* (in Russian), 123 (1958), 775–777.

[2] V. Ya. Pan, *Russian Math. Surv.* 21, 1 (1966), 105–136.

[3] V. Y. Pan, Ph.D. Thesis, Dept. of Mechanics and Math., Moscow State University (1964).

[4] Z. Kedem and D. Kirkpatrick, *SIAM J. on Comp.* 6, 1 (1977), 188–199.

[5] J. Morgenstern, *Journal ACM* 20, 2 (1973), 305–306.

[6] L. Hyafil, *Proc. 18th Ann. FOCS Symposium* (1977), 171–174.

[7] A. Borodin and I. Munro, *The Computational Complexity of Algebraic and Numeric Problems*, American Elsevier (1975).

[8] D. E. Knuth, *The Art of Computer Programming* 2 (1981).

[9] F. L. Hitchcock, *J. Math. and Phyics* 6 (1927), 164–189.

[10] N. Bourbaki, *Algèbre* 2, Hermann, Paris (1970), A2. 111–112.

[11] V. Ya. Pan, *Russ. Math. Surveys* 27, 5 (1972), 249–250.

[12] V. Strassen, *J. für die reine und angew. Math.* 264 (1973), 84–202.

[13] H. J. Ryser, Buffalo: Math. Assoc. of America, (1963), 26–28.

LMED

-8