SIAM J. COMPUT.
Vol. 43, No. 1, pp. 144–149

NEW LOWER BOUNDS FOR THE RANK OF MATRIX MULTIPLICATION*

J. M. LANDSBERG†

Abstract. The rank of the matrix multiplication operator for $\mathbf{n}\times\mathbf{n}$ matrices is one of the most studied quantities in algebraic complexity theory. I prove that the rank is at least $3\mathbf{n}^2 - o(\mathbf{n}^2)$. More precisely, for any integer $p \le \mathbf{n}-1$ the rank is at least $(3-\frac{1}{p+1})\mathbf{n}^2 - (1+2p\binom{2p}{p-1})\mathbf{n}$. The previous lower bound, due to Bläser, was $\frac{5}{2}\mathbf{n}^2 - 3\mathbf{n}$ (the case $p=1$). The new bounds improve Bläser's bound for all $\mathbf{n} > 84$. I also prove lower bounds for rectangular matrices that are significantly better than the previous bound.

Key words. rank, matrix multiplication

AMS subject classification. 68Q17

DOI. 10.1137/120880276

1. Introduction.

Let $X = (x^i_j)$, $Y = (y^i_j)$ be $\mathbf{n}\times\mathbf{n}$ matrices with indeterminant entries. The *rank* of matrix multiplication, denoted $\mathbf{R}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{n}\rangle})$, is the smallest number r of products $p_\rho = u_\rho(X)v_\rho(Y)$, $1 \le \rho \le r$, where u_ρ, v_ρ are linear forms, such that the entries of the matrix product XY are contained in the linear span of the p_ρ. This quantity is also called the bilinear complexity of $\mathbf{n}\times\mathbf{n}$ matrix multiplication. More generally, one may define the rank $\mathbf{R}(b)$ of any bilinear map b; see section 2.

From the point of view of geometry, rank is badly behaved as it is not semicontinuous. Geometers usually prefer to work with the *border rank* of matrix multiplication, which fixes the semicontinuity problem by fiat: the border rank of a bilinear map b, denoted $\underline{\mathbf{R}}(b)$, is the smallest r such that b can be approximated to arbitrary precision by bilinear maps of rank r. By definition, one has $\mathbf{R}(b) \ge \underline{\mathbf{R}}(b)$. A more formal definition is given in section 2.

Let $M_{\langle \mathbf{n},\mathbf{m},\mathbf{l}\rangle}$ denote the multiplication of an $\mathbf{n}\times\mathbf{m}$ matrix by an $\mathbf{m}\times\mathbf{l}$ matrix. In [5] G. Ottaviani and I gave new lower bounds for the border rank of matrix multiplication; namely, for all $p \le \mathbf{n}-1$, $\underline{\mathbf{R}}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{m}\rangle}) \ge \frac{2p+1}{p+1}\mathbf{nm}$. Taking $p = \mathbf{n}-1$ gives the bound $\underline{\mathbf{R}}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{m}\rangle}) \ge 2\mathbf{nm} - \mathbf{m}$. In this paper it will be advantageous to work with a smaller value of p. The results of [5] are used here to prove the following theorem.

THEOREM 1.1. *Let $p \le \mathbf{n}-1$ be a natural number. Then*

$$\mathbf{R}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{m}\rangle}) \ge \frac{2p+1}{p+1}\mathbf{nm} + \mathbf{n}^2 - \left(1 + 2p\binom{2p}{p-1}\right)\mathbf{n}.$$

The previous bound, due to M. Bläser [2], was $\mathbf{R}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{m}\rangle}) \ge 2\mathbf{nm} - \mathbf{m} + 2\mathbf{n} - 2$. For square matrices Theorem 1.1 specializes to the following theorem.

THEOREM 1.2. *Let $p \le \mathbf{n}-1$ be a natural number. Then*

$$\mathbf{R}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{n}\rangle}) \ge \left(3 - \frac{1}{p+1}\right)\mathbf{n}^2 - \left(1 + 2p\binom{2p}{p-1}\right)\mathbf{n}.$$

In particular, $\mathbf{R}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{n}\rangle}) \ge 3\mathbf{n}^2 - o(\mathbf{n}^2)$.

*Received by the editors June 11, 2012; accepted for publication (in revised form) September 23, 2013; published electronically February 4, 2014. This work was supported by NSF grant DMS-1006353.
http://www.siam.org/journals/sicomp/43-1/88027.html
†Department of Mathematics, Texas A&M University, College Station, TX 77843 (jml@math.tamu.edu).

The "in particular" sentence follows by setting, e.g., $p = \lfloor \sqrt{\log(\mathbf{n})} \rfloor$.

This improves Bläser's bound [1] of $\frac{5}{2}\mathbf{n}^2 - 3\mathbf{n}$ (the case $p = 1$) for all $\mathbf{n} > 84$. Working under my direction, Massarenti and Raviolo [8, 7] improved the error term in Theorem 1.2. In a preprint of this article I made a mistake in computing the error term. Unfortunately this mistake was not noticed before Massarenti and Raviolo's paper [8], which repeated the error, was published, although their contribution is completely correct, and their correct bound will appear in [7].

Remark 1.3. If T is a tensor of border rank r, where the approximating curve of rank r tensors limits in such a way that q derivatives of the curve are used, then the rank of T is at most $(2q-1)r$; see [3, Prop. 15.26]. In [6] the authors give explicit, but very large, upper bounds on the order of approximation h needed to write a tensor of border rank r as lying in the h-jet of a curve of tensors of rank r.

The language of tensors will be used throughout. In section 2 the language of tensors is introduced, and previous work of Bläser and others is rephrased in a language suitable for generalizations. In section 3 I describe the equations of [5] and give a very easy proof of a slightly weaker result than Theorem 1.1. In section 4 I express the equations in coordinates and prove Theorem 1.1. I work over the complex numbers throughout.

2. Ranks and border ranks of tensors. Let A, B, C be vector spaces of dimensions $\mathbf{a}, \mathbf{b}, \mathbf{c}$ and with dual spaces A^*, B^*, C^*. That is, A^* is the space of linear maps $A \to \mathbb{C}$. Write $A^* \otimes B$ for the space of linear maps $A \to B$ and $A^* \otimes B^* \otimes C$ for the space of bilinear maps $A \times B \to C$. To avoid extra $*$'s, I work with bilinear maps $A^* \times B^* \to C$, i.e., elements of $A \otimes B \otimes C$. Let $T : A^* \times B^* \to C$ be a bilinear map. One may also consider T as a linear map $T : A^* \to B \otimes C$ (and similarly with the roles of A, B, C exchanged) or as a trilinear map $A^* \times B^* \times C^* \to \mathbb{C}$.

The *rank* of a bilinear map $T : A^* \times B^* \to C$, denoted $\mathbf{R}(T)$, is the smallest r such that there exist $a_1, \dots, a_r \in A$, $b_1, \dots, b_r \in B$, $c_1, \dots, c_r \in C$ such that $T(\alpha, \beta) = \sum_{i=1}^{r} a_i(\alpha) b_i(\beta) c_i$ for all $\alpha \in A^*$ and $\beta \in B^*$. The *border rank* of T, denoted $\underline{\mathbf{R}}(T)$, is the smallest r such that T may be written as a limit of a sequence of rank r tensors. Since the set of tensors of border rank at most r is closed, one can use polynomials to obtain lower bounds on border rank. That is, let P be a polynomial on $A \otimes B \otimes C$ such that P vanishes on all tensors of border rank at most r: if $T \in A \otimes B \otimes C$ is such that $P(T) \neq 0$, then $\underline{\mathbf{R}}(T) > r$.

The following proposition is a rephrasing of part of the proof in [1].

PROPOSITION 2.1. *Let P be a polynomial of degree d on $A \otimes B \otimes C$ such that $P(T) \neq 0$ implies $\underline{\mathbf{R}}(T) > r$. Let $T \in A \otimes B \otimes C$ be a tensor such that $P(T) \neq 0$ and $T : A^* \to B \otimes C$ is injective. Then $\mathbf{R}(T) \geq r + \mathbf{a} - d$.*

As stated, the proposition is useless, as the degrees of polynomials vanishing on all tensors of border rank at most r are greater than r. (A general tensor of border rank r also has rank r.) However, the conclusion still holds if one can find, for a given tensor T, a polynomial, or collection of polynomials on smaller spaces, such that the nonvanishing of P on T is equivalent to the nonvanishing of the new polynomials. Then one substitutes the smaller degree into the statement to obtain the nontrivial lower bound.

In our situation, first I will show $P(M_{\langle \mathbf{n},\mathbf{n},\mathbf{m} \rangle}) \neq 0$ if and only if $\tilde{P}(\tilde{M}) \neq 0$, where $\tilde{M}$ is a tensor in a smaller space of tensors and $\tilde{P}$ is a polynomial of lower degree than P; see (2). More precisely, note that in the course of the proof, $B \otimes C$ does not play a role, and we will see that the relevant polynomial, when applied to matrix multiplication $M \in A \otimes B \otimes C = A \otimes \mathbb{C}^{\mathbf{n}^2} \otimes \mathbb{C}^{\mathbf{n}^2}$, will not vanish if and only if a polynomial $\tilde{P}$ applied to $\tilde{M} \in A\otimes\mathbb{C}^{\mathbf{n}}\otimes\mathbb{C}^{\mathbf{n}}$ with $\deg(\tilde{P}) = \deg(P)/\mathbf{n}$ does not vanish, so the proof below works in this case. Then, in section 4, I show that $\tilde{P}(\tilde{M}) \neq 0$ is implied by the nonvanishing of two polynomials of even smaller degrees.

This is why both Bläser's result and the result of this paper improve the bound of border rank by $\mathbf{a} = \mathbf{n}^2$ minus an error term, where Bläser improves Strassen's bound and I improve the bound of [5]. (Bläser shows Strassen's equations for border rank, when applied to the matrix multiplication tensor, are equivalent to the nonvanishing of three polynomials of degree $\mathbf{n}$, hence the error term of $3\mathbf{n}$. See [4, section 11.5] for an exposition.)

To prove the proposition, we need a standard lemma, also used in [2], which appears in this form in [4, Lemma 11.5.0.2].

LEMMA 2.2. *Let $\mathbb{C}^{\mathbf{a}}$ be given a basis. Given a polynomial P of degree d on $\mathbb{C}^{\mathbf{a}}$, there exists a set of at most d basis vectors such that P restricted to their span is not identically zero.*

The lemma follows by simply choosing a monomial that appears in P, as it can involve at most d basis vectors.

Proof of Proposition 2.1. Let $\mathbf{R}(T) = r$, and assume we have written T as a sum of r rank one tensors. Since $T : A^* \to B\otimes C$ is injective, we may write $T = T' + T''$ with $\mathbf{R}(T') = \mathbf{a}$, $\mathbf{R}(T'') = r - \mathbf{a}$, and $T' : A^* \to B\otimes C$ injective. Now consider the $\mathbf{a}$ elements of $A\otimes B\otimes C$ appearing in T'. Since they are linearly independent, by Lemma 2.2 we may choose a subset of d of them such that P, evaluated on the sum of terms in T whose A terms are in the span of these d elements, is not identically zero. Let T_1 denote the sum of the terms in T' not involving the (at most) d basis vectors needed for nonvanishing, so $\mathbf{R}(T_1) \geq \mathbf{a} - d$. Let $T_2 = T - T_1 + T''$. Now $\underline{\mathbf{R}}(T_2) \geq r$ because $P(T_2) \neq 0$. Finally, $\mathbf{R}(T) = \mathbf{R}(T_1) + \mathbf{R}(T_2)$. ◻

Let $G(k, V) \subset \mathbb{P}\Lambda^k V$ denote the Grassmannian of k-planes through the origin in V in its Plücker embedding. That is, if a k-plane is spanned by $v_1, \ldots, v_k$, we write it as $[v_1 \wedge \cdots \wedge v_k]$. One says a function on $G(k, V)$ is a polynomial of degree d if, as a function in the Plücker coordinates, it is a degree d polynomial. The Plücker coordinates (x^μ_α), $k + 1 \leq \mu \leq \dim V = \mathbf{v}$, $1 \leq \alpha \leq k$, are obtained by choosing a basis $e_1, \ldots, e_{\mathbf{v}}$ of V, centering the coordinates at $[e_1 \wedge \cdots \wedge e_k]$, and writing a nearby k-plane as $[(e_1 + \sum x^\mu_1 e_\mu) \wedge \cdots \wedge (e_k + \sum x^\mu_k e_\mu)]$. If the polynomial is also homogeneous in the x^μ_α, this is equivalent to it being the restriction of a homogeneous degree d polynomial on $\Lambda^k V$. (The ambiguity of the scale does not matter as we are concerned only with its vanishing.)

LEMMA 2.3. *Let A be given a basis. Given a homogeneous polynomial of degree d on the Grassmannian $G(k, A)$, there exists at least dk basis vectors such that, denoting their (at most) dk-dimensional span by A', P restricted to $G(k, A')$ is not identically zero.*

Proof. Consider the map $f : A^{\times k} \to G(k, A)$ given by $(a_1, \ldots, a_k) \mapsto [a_1 \wedge \cdots \wedge a_k]$. Then f is surjective. Take the polynomial P and pull it back by f. (The pullback $f^*(P)$ is defined by $f^*(P)(a_1, \ldots, a_k) := P(f(a_1, \ldots, a_k))$.) The pullback is of degree d in each copy of A. (That is, fixing $k - 1$ parameters, it becomes a degree d polynomial in the kth.) Now simply apply Lemma 2.2 k times to see that the pulled back polynomial is not identically zero restricted to A', and thus P restricted to $G(k, A')$ is not identically zero. ◻

Remark 2.4. The bound in Lemma 2.3 is sharp, as we give A a basis $a_1, \ldots, a_{\mathbf{a}}$ and consider the polynomial on $\Lambda^k A$ with coordinates $x^I = x^{i_1}, \ldots, x^{i_k}$ corresponding to the vector $\sum_I x^I a_{i_1} \wedge \cdots \wedge a_{i_k}$: $P = x^{1,\ldots,k} x^{k+1,\ldots,2k} \cdots x^{(d-1)k+1,\ldots,dk}$. Then

P restricted to $G(k,\langle a_1,\dots,a_{dk}\rangle)$ is nonvanishing, but there is no smaller subspace spanned by basis vectors on which it is nonvanishing.

3. Matrix multiplication and its rank. Let $M_{\langle \mathbf{m},\mathbf{n},\mathbf{l}\rangle} : Mat_{\mathbf{m}\times\mathbf{n}} \times Mat_{\mathbf{n}\times\mathbf{l}} \to Mat_{\mathbf{m}\times\mathbf{l}}$ denote the matrix multiplication operator. Write $M=\mathbb{C}^{\mathbf{m}}$, $N=\mathbb{C}^{\mathbf{n}}$, and $L=\mathbb{C}^{\mathbf{l}}$. Then

$$M_{\langle \mathbf{m},\mathbf{n},\mathbf{l}\rangle} : (N{\otimes}L^*)\times(L{\otimes}M^*)\to N{\otimes}M^*$$

has the interpretation as $M_{\langle \mathbf{m},\mathbf{n},\mathbf{l}\rangle}=Id_N{\otimes}Id_M{\otimes}Id_L\in(N^*{\otimes}L){\otimes}(L^*{\otimes}M){\otimes}(N{\otimes}M^*)$, where $Id_N\in N^*{\otimes}N$ is the identity map. If one thinks of $M_{\langle \mathbf{m},\mathbf{n},\mathbf{l}\rangle}$ as a trilinear map $(N{\otimes}L^*)\times(L{\otimes}M^*)\times(N{\otimes}M^*)\to\mathbb{C}$, in bases it is $(X,Y,Z)\mapsto \mathrm{trace}(XYZ)$. If one thinks of $M_{\langle \mathbf{m},\mathbf{n},\mathbf{l}\rangle}$ as a linear map $N{\otimes}L^*\to(L^*{\otimes}M){\otimes}(N{\otimes}M^*)$, it is just the identity map tensored with Id_M. In particular, if $\alpha\in N{\otimes}L^*$ is of rank q, its image, considered as a linear map $L{\otimes}M^*\to N{\otimes}M^*$, is of rank $q\mathbf{m}$.

Returning to general tensors $T\in A{\otimes}B{\otimes}C$, from now on assume that $\mathbf{b}=\mathbf{c}$. When $T=M_{\langle \mathbf{m},\mathbf{n},\mathbf{l}\rangle}$, one has $A=N^*{\otimes}L$, $B=L^*{\otimes}M$, $C=N{\otimes}M^*$, so $\mathbf{b}=\mathbf{c}$ is equivalent to $\mathbf{l}=\mathbf{n}$.

The equations of [5] are as follows: given $T\in A{\otimes}B{\otimes}C$, with $\mathbf{b}=\mathbf{c}$, take $A'\subset A$ of dimension $2p+1\le\mathbf{a}$. Define a linear map

$$T^{\wedge p}_{A'} : \Lambda^p A'{\otimes}B^*\to\Lambda^{p+1}A'{\otimes}C \tag{1}$$

by first considering $T|_{A'\otimes B\otimes C} : B^*\to A'{\otimes}C$ tensored with the identity map on $\Lambda^p A'$, which is a map $\Lambda^p A{\otimes}B^*\to\Lambda^p A{\otimes}A{\otimes}C$, and then projecting the image to $\Lambda^{p+1}A'{\otimes}C$. Then if the determinant of this linear map is nonzero, the border rank of T is at least $\frac{2p+1}{p+1}\mathbf{b}$. If there exists an A' such that the determinant is nonzero, we may think of the determinant as a nontrivial homogeneous polynomial of degree $\binom{2p+1}{p}\mathbf{b}$ on $G(2p+1,A)$.

Now consider the case $T=M_{\langle \mathbf{n},\mathbf{n},\mathbf{m}\rangle}$, and recall that $B=L^*{\otimes}M$, $C=N{\otimes}M^*$. The map $(M_{\langle \mathbf{n},\mathbf{n},\mathbf{m}\rangle})^{\wedge p}_{A'} : \Lambda^p A'{\otimes}L{\otimes}M^*\to\Lambda^{p+1}A'{\otimes}N{\otimes}M^*$ is actually a reduced map

$$\tilde{M}^{\wedge p}_{A'} : \Lambda^p A'{\otimes}L\to\Lambda^{p+1}A'{\otimes}N \tag{2}$$

tensored with the identity map $M^*\to M^*$, and thus its determinant is nonvanishing if and only if the determinant of $\tilde{M}^{\wedge p}_{A'}$ is nonvanishing. *But this is a polynomial of degree* $\binom{2p+1}{p}\mathbf{n} << \binom{2p+1}{p}\mathbf{n}^2$ *on* $G(2p+1,\mathbf{n}^2)$. Proposition 2.1, with $d=\binom{2p+1}{p}\mathbf{n}$, $\mathbf{a}=\mathbf{n}^2$, and $r=\frac{2p+1}{p+1}\mathbf{mn}$, combined with Lemma 2.3, gives the bound

$$\mathbf{R}(M_{\langle \mathbf{n},\mathbf{n},\mathbf{m}\rangle})\ge\frac{2p+1}{p+1}\mathbf{nm}+\mathbf{n}^2-(2p+1)\binom{2p+1}{p}\mathbf{n}.$$

Note that this already gives the $3\mathbf{n}^2-o(\mathbf{n}^2)$ asymptotic lower bound. The remainder of the paper is dedicated to improving the error term.

4. The equations of [5] in coordinates. Let $\mathbf{a}=3$ (so $p=1$) and $\mathbf{b}=\mathbf{c}$; the map (1) expressed in bases is a $3\mathbf{b}\times3\mathbf{b}$ matrix. If a_0,a_1,a_2 is a basis of A and one chooses bases of B,C, then elements of $B{\otimes}C$ may be written as matrices, and $T=a_0{\otimes}X_0+a_1{\otimes}X_1+a_2{\otimes}X_2$, where the X_j are size $\mathbf{b}$ square matrices. Order the basis of A by a_0, a_1, a_2 and of $\Lambda^2 A$ by $a_1 \wedge a_2, a_0 \wedge a_1, a_0 \wedge a_2$. We compute

$$\begin{aligned} T_A^{\wedge 1}(a_0 \otimes \beta) &= \beta(X_0)\otimes a_0 \wedge a_0 + \beta(X_1)\otimes a_1 \wedge a_0 + \beta(X_2)\otimes a_2 \wedge a_0 \\ &= -\beta(X_1)\otimes a_0 \wedge a_1 - \beta(X_2)\otimes a_0 \wedge a_2, \\ T_A^{\wedge 1}(a_1 \otimes \beta) &= \beta(X_0)\otimes a_0 \wedge a_1 - \beta(X_2)\otimes a_1 \wedge a_2, \\ T_A^{\wedge 1}(a_2 \otimes \beta) &= \beta(X_0)\otimes a_0 \wedge a_2 + \beta(X_1)\otimes a_1 \wedge a_2, \end{aligned}$$

so the corresponding matrix for $T_A^{\wedge 1}$ is the block matrix

$$Mat(T_A^{\wedge 1}) = \begin{pmatrix} 0 & -X_2 & X_1 \\ -X_1 & X_0 & 0 \\ -X_2 & 0 & X_0 \end{pmatrix}.$$

Now assume X_0 is invertible, and change bases such that it is the identity matrix. Recall the formula for block matrices,

$$\det \begin{pmatrix} X & Y \\ Z & W \end{pmatrix} = \det(W)\det(X - YW^{-1}Z), \tag{3}$$

assuming W is invertible. Then, using the $(\mathbf{b}, 2\mathbf{b}) \times (\mathbf{b}, 2\mathbf{b})$ blocking (so $X = 0$ in (3)),

$$\det Mat(T_A^{\wedge 1}) = \det(X_1X_2 - X_2X_1) = \det([X_1, X_2]).$$

When $\dim A > 3$, if there exists a three dimensional subspace A' of A, such that $\det Mat(T_{A'}^{\wedge 1}) \neq 0$, then $\underline{\mathbf{R}}(T) \geq \frac{3}{2}\mathbf{b}$ as this is (1) in the case $p = 1$. These are Strassen's equations [9].

I now phrase the equations of [5] in coordinates. Let $\dim A = 2p+1$. Write $T = a_0 \otimes X_0 + \cdots + a_{2p} \otimes X_{2p}$. The expression of (1) in bases is as follows: write $a_I := a_{i_1} \wedge \cdots \wedge a_{i_p}$ for $\Lambda^p A$, require that the first $\binom{2p}{p-1}$ basis vectors have $i_1 = 0$ and that the second $\binom{2p}{p}$ do not, and call these multi-indices $0J$ and K. Order the bases of $\Lambda^{p+1}A$ such that the first $\binom{2p}{p+1}$ multi-indices do not have 0 and the second $\binom{2p}{p}$ do, and furthermore that the second set of indices is ordered the same way as K, except that we write $0K$ since a zero index is included. Then the resulting matrix is of the form

$$\begin{pmatrix} 0 & Q \\ \tilde{Q} & R \end{pmatrix}, \tag{4}$$

where this matrix is blocked $(\binom{2p}{p+1}\mathbf{b}, \binom{2p}{p}\mathbf{b}) \times (\binom{2p}{p+1}\mathbf{b}, \binom{2p}{p}\mathbf{b})$,

$$R = \begin{pmatrix} X_0 & & \\ & \ddots & \\ & & X_0 \end{pmatrix},$$

and $Q, \tilde{Q}$ have entries in blocks consisting of $X_1, \ldots, X_{2p}$ and zero. Thus if X_0 is the identity matrix, so is R, and the determinant equals the determinant of $Q\tilde{Q}$. If X_0 is the identity matrix, when $p = 1$ we have $Q\tilde{Q} = [X_1, X_2]$, and when $p = 2$

$$Q\tilde{Q} = \begin{pmatrix} 0 & [X_1, X_2] & [X_1, X_3] & [X_1, X_4] \\ [X_2, X_1] & 0 & [X_2, X_3] & [X_2, X_4] \\ [X_3, X_1] & [X_3, X_2] & 0 & [X_3, X_4] \\ [X_4, X_1] & [X_4, X_2] & [X_4, X_3] & 0 \end{pmatrix}. \tag{5}$$

In general, when X_0 is the identity matrix, $Q\tilde{Q}$ is a block $\binom{2p}{p-1}\mathbf{b} \times \binom{2p}{p-1}\mathbf{b}$ matrix whose block entries are either zero or commutators $[X_i, X_j]$.

To prove Theorem 1.1 we work with $\tilde{M}_{A'}^{\wedge p}$ of (2), so $\mathbf{b} = \mathbf{n}$. First apply Lemma 2.2 to choose $\mathbf{n}$ basis vectors such that restricted to them $\det(X_0)$ is nonvanishing, and then we consider our polynomial $\det(Q\tilde{Q})$ as defined on $G(2p, (2p+1)\mathbf{n}^2 - 1)$ and apply Lemma 2.3, using $2p\binom{2p}{p-1}\mathbf{n}$ basis vectors to ensure it is nonvanishing. Our error term is thus $\mathbf{n} + 2p\binom{2p}{p-1}\mathbf{n}$, and the theorem follows.

Remark 4.1. In [8, 7], the authors show that the matrix $Q\tilde{Q}$ can be made to have a nonzero determinant by a subtle combination of factoring and splitting it into a sum of two matrices that carries a lower cost than just taking its determinant.

Acknowledgments. I thank the anonymous referee for useful suggestions and C. Ikenmeyer for help with the exposition.

REFERENCES

[1] M. Bläser, *A $\frac{5}{2}n^2$-lower bound for the rank of $n \times n$-matrix multiplication over arbitrary fields*, in Proceedings of the 40th Annual Symposium on Foundations of Computer Science (New York, 1999), IEEE Computer Society, Los Alamitos, CA, 1999, pp. 45–50.

[2] M. Bläser, *On the complexity of the multiplication of matrices of small formats*, J. Complexity, 19 (2003), pp. 43–60.

[3] P. Bürgisser, M. Clausen, and M. Amin Shokrollahi, *Algebraic Complexity Theory*, Grundlehren Math. Wiss. 315, Springer-Verlag, Berlin, 1997.

[4] J. M. Landsberg, *Tensors: Geometry and Applications*, Grad. Stud. Math. 128, AMS, Providence, RI, 2012.

[5] J. M. Landsberg and G. Ottaviani, *New Lower Bounds for the Border Rank of Matrix Multiplication*, preprint, arXiv:1112.6007.

[6] T. Lehmkuhl and T. Lickteig, *On the order of approximation in approximative triadic decompositions of tensors*, Theoret. Comput. Sci., 66 (1989), pp. 1–14.

[7] A. Massarenti and E. Raviolo, *Erratum to: The rank of $n \times n$ matrix multiplication is at least $3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - 3n$*, Linear Algebra Appl., to appear.

[8] A. Massarenti and E. Raviolo, *The rank of $n \times n$ matrix multiplication is at least $3n^2 - 2\sqrt{2}n^{\frac{3}{2}} - 3n$*, Linear Algebra Appl., 438 (2013), pp. 4500–4509.

[9] V. Strassen, *Rank and optimal computation of generic tensors*, Linear Algebra Appl., 52/53 (1983), pp. 645–685.