Barriers for fast matrix multiplication from irreversibility

Christandl, Matthias; Vrana, Péter; Zuiddam, Jeroen

Published in:
34th Computational Complexity Conference, CCC 2019

DOI:
10.4230/LIPIcs.CCC.2019.26

Publication date:
2019

Document version
Publisher's PDF, also known as Version of record

Barriers for Fast Matrix Multiplication from Irreversibility

Matthias Christandl

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark
christandl@math.ku.dk

Péter Vrana

Department of Geometry, Budapest University of Technology and Economics, Egry József u. 1., 1111 Budapest, Hungary
MTA-BME Lendület (Momentum) Research group
vranap@math.bme.hu

Jeroen Zuiddam

Institute for Advanced Study, 1 Einstein Drive, Princeton 08540, NJ, USA
jzuiddam@ias.edu

Abstract

Determining the asymptotic algebraic complexity of matrix multiplication, succinctly represented by the matrix multiplication exponent ω, is a central problem in algebraic complexity theory. The best upper bounds on ω, leading to the state-of-the-art $\omega \leq 2.37..$, have been obtained via the laser method of Strassen and its generalization by Coppersmith and Winograd. Recent barrier results show limitations for these and related approaches to improve the upper bound on ω.

We introduce a new and more general barrier, providing stronger limitations than in previous work. Concretely, we introduce the notion of "irreversibility" of a tensor and we prove (in some precise sense) that any approach that uses an irreversible tensor in an intermediate step (e.g., as a starting tensor in the laser method) cannot give $\omega = 2$. In quantitative terms, we prove that the best upper bound achievable is lower bounded by two times the irreversibility of the intermediate tensor. The quantum functionals and Strassen support functionals give (so far, the best) lower bounds on irreversibility. We provide lower bounds on the irreversibility of key intermediate tensors, including the small and big Coppersmith–Winograd tensors, that improve limitations shown in previous work. Finally, we discuss barriers on the group-theoretic approach in terms of "monomial" irreversibility.

2012 ACM Subject Classification Theory of computation → Algebraic complexity theory

Keywords and phrases Matrix multiplication exponent, barriers, laser method

Digital Object Identifier 10.4230/LIPIcs.CCC.2019.26

Related Version `https://arxiv.org/abs/1812.06952`

Funding *Matthias Christandl*: European Research Council (ERC Grant Agreement No. 337603) and VILLUM FONDEN via the QMATH Centre of Excellence (Grant No. 10059).
Péter Vrana: National Research, Development and Innovation Fund of Hungary within the Quantum Technology National Excellence Program (Project Nr. 2017-1.2.1-NKP-2017-00001) and research grants K124152, KH129601.
Jeroen Zuiddam: National Science Foundation under Grant No. DMS-1638352.

34th Computational Complexity Conference (CCC 2019).
Editor: Amir Shpilka; Article No. 26; pp. 26:1–26:17
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

1.1 Matrix multiplication

Determining the asymptotic algebraic complexity of matrix multiplication is a central open problem in algebraic complexity theory. Several different methods for constructing fast matrix multiplication algorithms have been developed, but on a high level they typically consist of two parts: an efficient reduction of matrix multiplication to an intermediate problem (some bilinear map, i.e. 3-tensor) and an efficient algorithm for the intermediate problem. Recent results have shown "barriers" for such constructions to yield fast matrix multiplication algorithms [4, 7, 8, 2, 3]. We give a barrier, based on a new notion called irreversibility, that is more general and in some cases stronger than the barriers from previous work.

1.2 Matrix multiplication barriers

The matrix multiplication exponent ω is defined as the infimum over all real numbers β such that any two $n \times n$ matrices can be multiplied with $\mathcal{O}(n^\beta)$ algebraic operations, and thus ω represents the asymptotic algebraic complexity of matrix multiplication. Trivially holds $2 \leq \omega \leq 3$. Strassen published the first non-trivial upper bound $\omega \leq \log_2 7$ in 1969 [20]. In the decades that followed, through the development of several ingenious methods by several people, the upper bound was improved to the state-of-the-art bound $\omega \leq 2.37..$, and the pursuit to prove whether $\omega = 2$ or $\omega > 2$ has been ongoing [12, 19, 27, 17, 10, 11]. As mentioned before, these upper bound methods typically consist of a reduction of matrix multiplication to an intermediate problem and an efficient algorithm for the intermediate problem.

Ambainis et al. [4], for the first time, proved a "barrier" result for some collection of such methods. Namely, they showed that a variety of methods that go via the big Coppersmith–Winograd tensor as an intermediate problem cannot give $\omega = 2$, and in fact not even $\omega \leq 2.30...$ We call any lower bound for all upper bounds on ω that can be obtained by some method, a *barrier* for that method. In general, barriers in the sense of limitations to proof methods have a long history in computational complexity theory and recognizing barriers is a natural step towards finding proof methods that do solve the problem at hand.

Next, Alman and Williams [2, 3] extended the realm of barriers beyond the scope of the Ambainis et al. barrier, to a larger collection of methods. Also Blasiak et al. [7, 8] did a study of barriers, namely of barriers for a subset of the group-theoretic method. Both the Blasiak et al. and the Alman and Williams barriers rely on studying versions of "asymptotic subrank" of an intermediate problem.

We give a barrier that applies more generally than all previous barriers and that is in some cases stronger. Our barrier also relies on studying versions of asymptotic subrank, which together with the notion of asymptotic rank we combine into a single parameter called irreversibility. Our barrier simplifies and generalises previous barriers and connects the barrier literature to some central notions from the framework of Strassen [21, 22, 23, 9]. Alman reported similar independent results [1] shortly after our manuscript appeared on the arxiv.

1.3 Our barrier: intuitive explanation

An intuitive explanation of our barrier is as follows. In the language of tensors, the matrix multiplication exponent ω is the optimal "rate of transformation" from the "unit tensor" to the "matrix multiplication tensor",

$$\text{unit tensor} \xrightarrow{\omega} \text{matrix multiplication tensor.} \tag{1}$$

The rate of transformation naturally satisfies a triangle inequality and thus upper bounds on ω can be obtained by combining the rate of transformation α_1 from the unit tensor to some intermediate tensor and the rate of transformation α_2 from the intermediate tensor to the matrix multiplication tensor; this is the two-component approach alluded to earlier,

$$\text{unit tensor} \xrightarrow{\alpha_1} \text{intermediate tensor} \xrightarrow{\alpha_2} \text{matrix multiplication tensor.} \tag{2}$$

We define the irreversibility of the intermediate tensor as the necessary "loss" that will occur when transforming the unit tensor to the intermediate tensor followed by transforming the intermediate tensor back to the unit tensor. It is well-know that the transformation rate from the matrix multiplication tensor to the unit tensor is $\frac{1}{2}$, so we can extend (2) to

$$\text{unit tensor} \xrightarrow{\alpha_1} \text{intermediate tensor} \xrightarrow{\alpha_2} \text{matrix multiplication tensor} \xrightarrow{1/2} \text{unit tensor.} \tag{3}$$

We thus see that $\alpha_1\alpha_2$ is directly related to the irreversibility of the intermediate tensor, and hence the irreversibility of the intermediate tensor provides limitations on the upper bounds on ω that can be obtained from (2). In particular, any fixed irreversible intermediate tensor cannot show $\omega = 2$ via (2), since the matrix multiplication tensor is reversible when $\omega = 2$.

1.4 Explicit numerical barriers

To exemplify our barrier we show that the support functionals [23] and quantum functionals [9] give (so far, the best) lower bounds on the irreversibility of the following families of tensors:

- the small Coppersmith–Winograd tensors
$$\mathrm{cw}_q = \sum_{i=1}^{q} e_{0,i,i} + e_{i,0,i} + e_{i,i,0}$$
- the big Coppersmith–Winograd tensors
$$\mathrm{CW}_q = e_{0,0,q+1} + e_{0,q+1,0} + e_{q+1,0,0} + \sum_{i=1}^{q} e_{0,i,i} + e_{i,0,i} + e_{i,i,0}$$
- the reduced polynomial multiplication tensors
$$t_n = \sum_{\substack{i,j,k=0:\\ i+j=k}}^{n-1} e_{i,j,k}$$

which for small parameters lead to the following explicit barriers (Theorem 9 and Section 4.2):

q	cw_q-barrier
2	2
3	2.02..
4	2.06..
5	2.09..
6	2.12..
7	2.15..

q	CW_q-barrier
1	2.16..
2	2.17..
3	2.19..
4	2.20..
5	2.21..
6	2.23..

n	t_n-barrier
1	2.17..
2	2.16..
3	2.15..
4	2.15..
5	2.14..
6	2.14..

Indeed, as suggested by the values in the above tables, the cw_q-barrier and CW_q-barrier increase with q (converging to 3), whereas the t_n-barrier decreases with n (converging to 2).

1.5 Comparison and other applications

Compared to Ambainis, Filmus and Le Gall [4] our barriers are valid for a larger class of approaches (and naturally we obtain lower barriers). Compared to Alman and Williams [3] our barriers are valid for a larger class of approaches but our barriers are also higher. As a variation on our barrier we introduce a "monomial" version. Compared to Blasiak, Church, Cohn, Grochow, Naslund, Sawin and Umans [7], and Blasiak, Church, Cohn, Grochow and Umans [8] our monomial barriers are valid for a class of approaches that includes their STPP approach, and thus we provide a uniform view on the barriers that have appeared in the literature. We have not tried to optimise the barriers that we obtain, but focus instead on introducing the barrier itself. The barrier of Alman stated in [1] is very similar to ours, but makes use of asymptotic slice rank instead of asymptotic subrank. Since asymptotic subrank is at most asymptotic slice rank, our barriers are technically stronger. (It is not known whether asymptotic slice rank and asymptotic subrank are equal in general.)

It will become clear to the reader during the development of our ideas that they not only apply to the problem of fast matrix multiplication, but extend to give barriers for the more general problem of constructing fast rectangular matrix multiplication algorithms or even transformations between arbitrary powers of tensors. Such transformations may represent, for example, asymptotic slocc (stochastic local operations and classical communication) reductions among multipartite quantum states [5, 13, 25, 15].

We define irreversibility in Section 2. In Section 3 we introduce the irreversibility barrier. Finally, in Section 4 we present explicit irreversibility barriers.

2 Irreversibility

We begin by introducing some standard notation and terminology. Then we discuss a useful notion called the relative exponent and we define the irreversibility of a tensor. After that we introduce the monomial versions of these ideas and discuss so-called balanced tensors.

2.1 Standard definitions

We assume familiarity with tensors and with the tensor Kronecker product and direct sum. All our tensors will be 3-tensors over some fixed but arbitrary field $\mathbb{F}$. For two tensors $t \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3}$ and $s \in \mathbb{F}^{m_1} \otimes \mathbb{F}^{m_2} \otimes \mathbb{F}^{m_3}$ we write $t \geq s$ and say t restricts to s if there are linear maps $A_i : \mathbb{F}^{n_i} \to \mathbb{F}^{m_i}$ such that $(A_1, A_2, A_3) \cdot t = s$. For $n \in \mathbb{N}$ we define the diagonal tensor (also called the rank-n unit tensor) $\langle n \rangle \coloneqq \sum_{i=1}^{n} e_{i,i,i} \in \mathbb{F}^n \otimes \mathbb{F}^n \otimes \mathbb{F}^n$. The tensor rank of t is defined as $\mathrm{R}(t) \coloneqq \min\{n \in \mathbb{N} : t \leq \langle n \rangle\}$ (this coincides with the definition that $\mathrm{R}(t)$ is the smallest size of any decomposition of t into a sum of simple tensors) and the subrank of t is defined as $\mathrm{Q}(t) \coloneqq \max\{n \in \mathbb{N} : \langle n \rangle \leq t\}$. The asymptotic rank of t is defined as

$$\underset{\sim}{\mathrm{R}}(t) \coloneqq \lim_{n\to\infty} \mathrm{R}(t^{\otimes n})^{1/n} = \inf_n \mathrm{R}(t^{\otimes n})^{1/n} \tag{4}$$

and the asymptotic subrank of t is defined as

$$\underset{\sim}{\mathrm{Q}}(t) \coloneqq \lim_{n\to\infty} \mathrm{Q}(t^{\otimes n})^{1/n} = \sup_n \mathrm{Q}(t^{\otimes n})^{1/n}. \tag{5}$$

The above limits exist and equal the respective infimum and supremum by Fekete's lemma. For $a, b, c \in \mathbb{N}_{\geq 1}$ the matrix multiplication tensor $\langle a, b, c\rangle$ is defined as

$$\langle a, b, c\rangle \coloneqq \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} e_{(i,j),\,(j,k),\,(k,i)} \in (\mathbb{F}^a \otimes \mathbb{F}^b) \otimes (\mathbb{F}^b \otimes \mathbb{F}^c) \otimes (\mathbb{F}^c \otimes \mathbb{F}^a). \tag{6}$$

The matrix multiplication exponent is defined as $\omega \coloneqq \log_2 \underset{\sim}{\mathrm{R}}(\langle 2, 2, 2\rangle)$. The meaning of ω in terms of algorithms is: for any $\varepsilon > 0$ there is an algorithm that for any $n \in \mathbb{N}$ multiplies two $n \times n$ matrices using $\mathcal{O}(n^{\omega+\varepsilon})$ scalar additions and multiplications. The difficulty of determining the asymptotic rank of $\langle 2, 2, 2\rangle$ is to be contrasted with the situation for the asymptotic subrank; to put it in Strassen's words: *Unlike the cynic, who according to Oscar Wilde knows the price of everything and the value of nothing, we can determine the asymptotic value of $\langle h, h, h\rangle$ precisely* [22],

$$\underset{\sim}{\mathrm{Q}}(\langle h, h, h\rangle) = h^2. \tag{7}$$

2.2 Relative exponent

For a clean exposition of our barrier we will use the notion of relative exponent, which we will define in this section. This notion is inspired by the notion of rate from information theory and alternatively can be seen as a versatile version of the notion of the asymptotic preorder for tensors of Strassen. In the context of tensors, the relative exponent previously appeared in [28] and [26].

▶ **Assumption 1.** *To avoid irrelevant technicalities, we will from now on, without further mentioning, only consider tensors that are not of the form $u \otimes v \otimes w$.*

▶ **Definition 2.** *For two tensors $t \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3}$ and $s \in \mathbb{F}^{m_1} \otimes \mathbb{F}^{m_2} \otimes \mathbb{F}^{m_3}$ we define the* relative exponent *from t to s as*

$$\omega(t, s) \coloneqq \lim_{n\to\infty} \tfrac{1}{n} \min\{m \in \mathbb{N} : t^{\otimes m} \geq s^{\otimes n}\} \tag{8}$$

$$= \sup_{n} \tfrac{1}{n} \min\{m \in \mathbb{N} : t^{\otimes m} \geq s^{\otimes n}\}. \tag{9}$$

The limit is a supremum by Fekete's lemma. Let us briefly relate the relative exponent to the basic notions and results stated earlier. The reader verifies directly that the identities

$$\omega(\langle 2\rangle, t) = \log_2 \underset{\sim}{\mathrm{R}}(t) \tag{10}$$

$$\omega(t, \langle 2\rangle) = 1/(\log_2 \underset{\sim}{\mathrm{Q}}(t)) \tag{11}$$

hold. By definition of the matrix multiplication exponent ω holds

$$\omega(\langle 2\rangle, \langle 2, 2, 2\rangle) = \omega. \tag{12}$$

We know from (7) that

$$\omega(\langle 2, 2, 2\rangle, \langle 2\rangle) = \tfrac{1}{2}. \tag{13}$$

The relative exponent has the following two basic properties that the reader verifies directly.

▶ **Proposition 3.** *Let s, t and u be tensors.*

(i) $\omega(t, t) = 1$.

(ii) $\omega(s, t)\, \omega(t, u) \geq \omega(s, u)$ (triangle inequality).

2.3 Irreversibility

Our barrier framework relies crucially on the irreversibility of a tensor, a new notion that we define now.

▶ **Definition 4.** *We define the* irreversibility *of a tensor t as the product of the relative exponent from $\langle 2\rangle$ to t and the relative exponent from t to $\langle 2\rangle$, i.e.*

$$\mathbf{i}(t) := \omega(\langle 2\rangle, t)\,\omega(t, \langle 2\rangle). \tag{14}$$

Thus $\mathbf{i}(t)$ measures the extent to which the asymptotic conversion from $\langle 2\rangle$ to t is irreversible, explaining the name. Equivalently, the irreversibility is the ratio of the logarithms of the asymptotic rank and the asymptotic subrank, i.e.

$$\mathbf{i}(t) = \frac{\log_2 \underaccent{\sim}{\mathrm{R}}(t)}{\log_2 \underaccent{\sim}{\mathrm{Q}}(t)}. \tag{15}$$

From the basic properties of the relative exponent (Proposition 3) follows directly the inequality $\mathbf{i}(t) = \omega(\langle 2\rangle, t)\,\omega(t, \langle 2\rangle) \geq \omega(\langle 2\rangle, \langle 2\rangle) = 1$.

▶ **Proposition 5.** *For any tensor t holds that*

$$\mathbf{i}(t) \geq 1. \tag{16}$$

▶ **Definition 6.** *Let t be a tensor.*

- *If $\mathbf{i}(t) = 1$, then we say that t is* reversible*.*
- *If $\mathbf{i}(t) > 1$, then we say that t is* irreversible*.*

For example, for any $n \in \mathbb{N}$ the diagonal tensor $\langle n\rangle = \sum_{i=1}^{n} e_{i,i,i}$ is reversible. In fact, any reversible tensor t that we know of is equivalent to $\langle n\rangle$ for some n, in the sense that $\langle n\rangle \leq t \leq \langle n\rangle$.

For the matrix multiplication tensor $\langle 2,2,2\rangle$ we have $2\,\mathbf{i}(\langle 2,2,2\rangle) = \omega$ (using (13)). Thus if $\omega = 2$, then $\langle 2,2,2\rangle$ is reversible (and also any other $\langle n,n,n\rangle$). As we will see in Section 3, this is ultimately the source of our barrier.

Irreversible tensors exist. For example, $W = e_{0,0,1}+e_{0,1,0}+e_{1,0,0}$ is irreversible. Namely, it is well-known that $\log_2 \underaccent{\sim}{\mathrm{R}}(W) = 1$ and that $\log_2 \underaccent{\sim}{\mathrm{Q}}(W) = h(1/3) = 0.918..$ [23, Theorem 6.7], so $\mathbf{i}(W) = 1.088.. > 1$. In Section 4 we will compute lower bounds on the irreversibility of the small and big Coppersmith–Winograd tensors (that play a crucial role in the best upper bounds on ω).

2.4 Monomial relative exponent and monomial irreversibility

The following restrained version of relative exponent and irreversibility will be relevant. For two tensors $t \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3}$ and $s \in \mathbb{F}^{m_1} \otimes \mathbb{F}^{m_2} \otimes \mathbb{F}^{m_3}$ we write $t \geq_{\mathrm{M}} s$ and say t *monomially restricts* to s if there are linear maps $A_i : \mathbb{F}^{n_i} \to \mathbb{F}^{m_i}$, the corresponding matrices of which are generalised sub-permutation matrices in the standard basis, such that $(A_1, A_2, A_3) \cdot t = s$ [21, Section 6]. Replacing the preorder $\geq$ by $\geq_{\mathrm{M}}$ in Section 2 gives the notions of monomial subrank Q_{M}, monomial asymptotic subrank $\underaccent{\sim}{\mathrm{Q}}_{\mathrm{M}}$ and monomial relative exponent ω_{M}. (For simplicity we will use monomial restriction here, but our results will also hold with $\geq_{\mathrm{M}}$ replaced by monomial degeneration $\trianglerighteq_{\mathrm{M}}$ defined in [21, Section 6].) Note that the notions Q_{M} and $\underaccent{\sim}{\mathrm{Q}}_{\mathrm{M}}$ only depend on the support of the tensor, and not on the particular values of the nonzero coefficients. We define the monomial irreversibility $\mathbf{i}_{\mathrm{M}}(t)$ of t as the product of the (normal) relative exponent from $\langle 2\rangle$ to t and the monomial relative exponent from t to $\langle 2\rangle$,

$$\mathbf{i}_{\mathrm{M}}(t) \coloneqq \omega(\langle 2\rangle, t)\, \omega_{\mathrm{M}}(t, \langle 2\rangle). \tag{17}$$

Equivalently, we have

$$\mathbf{i}_{\mathrm{M}}(t) = \frac{\log_2 \underset{\sim}{\mathrm{R}}(t)}{\log_2 \underset{\sim}{\mathrm{Q}}_{\mathrm{M}}(t)}. \tag{18}$$

(This notion may depend on the tensor and not only on the support.)

▶ **Proposition 7.** *Let s, t and u be tensors.*
(i) $\omega_{\mathrm{M}}(t,t) = 1$.
(ii) $\omega_{\mathrm{M}}(s,t)\, \omega_{\mathrm{M}}(t,u) \geq \omega_{\mathrm{M}}(s,u)$ (triangle inequality).
(iii) $\omega_{\mathrm{M}}(s,t) \geq \omega(s,t)$.
(iv) $\mathbf{i}_{\mathrm{M}}(t) \geq \mathbf{i}(t)$.

▶ **Definition 8.** *Let t be a tensor.*
- *If $\mathbf{i}_{\mathrm{M}}(t) = 1$, then we say that t is* monomially reversible.
- *If $\mathbf{i}_{\mathrm{M}}(t) > 1$, then we say that t is* monomially irreversible.

There exist tensors that are reversible and monomially irreversible. For example, let C be the structure tensor of the algebra $\mathbb{C}[\mathbb{Z}/3\mathbb{Z}]$ in the natural basis,

$$C = e_{0,0,0} + e_{0,1,1} + e_{1,0,1} + e_{2,0,2} + e_{0,2,2} + e_{1,1,2} + e_{1,2,0} + e_{2,1,0} + e_{2,2,1}. \tag{19}$$

Then we have $\underset{\sim}{\mathrm{R}}(C) = 3$, $\underset{\sim}{\mathrm{Q}}(C) = 3$ and $\underset{\sim}{\mathrm{Q}}_{\mathrm{M}}(C) = 2.75..$ (this is proven in [14, 24], see also [9] for the connection to [23]), so that $\mathbf{i}(C) = 1$ and $\mathbf{i}_{\mathrm{M}}(C) = 1.08..$ We note that C is equivalent to the diagonal tensor $\langle 3\rangle$, see e.g. [9] for the basis transformation that shows this.

With regards to matrix multiplication, the standard construction for (13) in fact shows that

$$\omega_{\mathrm{M}}(\langle 2,2,2\rangle, \langle 2\rangle) = \tfrac{1}{2}. \tag{20}$$

2.5 Balanced tensors

We finish this section with a general comment on upper bounds on irreversibility. A tensor $t \in V_1 \otimes V_2 \otimes V_3$ with $\dim(V_1) = \dim(V_2) = \dim(V_3)$ is called *balanced* if the corresponding maps $t_1 : V_1 \to V_2 \otimes V_3$, $t_2 : V_2 \to V_1 \otimes V_3$ and $t_3 : V_3 \to V_1 \otimes V_2$ (called flattenings) are full-rank and for each $i \in [3]$ there is an element $v \in V_i$ such that $t_i(v)$ has full-rank [22, page 121]. For any tensor space with cubic format $\mathbb{F}^n \otimes \mathbb{F}^n \otimes \mathbb{F}^n$ over an algebraically closed field $\mathbb{F}$, being balanced is a generic condition, i.e. almost all elements in such a space are balanced. Balanced tensors are called 1-generic tensors in [16]. Let $t \in \mathbb{F}^n \otimes \mathbb{F}^n \otimes \mathbb{F}^n$ be balanced. Then [22, Proposition 3.6]

$$\underset{\sim}{\mathrm{R}}(t) \leq n^{\frac{2}{3}\omega} \tag{21}$$

$$\underset{\sim}{\mathrm{Q}}(t) \geq n^{\frac{2}{3}} \tag{22}$$

and so

$$\mathbf{i}(t) \leq \omega. \tag{23}$$

If moreover $\underset{\sim}{\mathrm{R}}(t) = n$, then

$$\mathbf{i}(t) \leq 3/2. \tag{24}$$

3 Barriers through irreversibility

With the new notion of irreversibility available, we present a barrier for approaches to upper bound ω via an intermediate tensor t.

3.1 The irreversibility barrier

For any tensor t the inequality

$$\omega(\langle 2\rangle, t)\,\omega(t, \langle 2,2,2\rangle) \geq \omega \tag{25}$$

holds by the triangle inequality. Any such approach to upper bound ω respects the following barrier in terms of the irreversibility $\mathbf{i}(t)$ of t.

▶ **Theorem 9.** *For any tensor t holds*

$$\omega(\langle 2\rangle, t)\,\omega(t, \langle 2,2,2\rangle) \geq 2\,\mathbf{i}(t). \tag{26}$$

Proof. By the triangle inequality (Proposition 3),

$$\omega(\langle 2\rangle, t)\,\omega(t, \langle 2,2,2\rangle)\,\omega(\langle 2,2,2\rangle, \langle 2\rangle) \geq \omega(\langle 2\rangle, t)\,\omega(t, \langle 2\rangle) = \mathbf{i}(t). \tag{27}$$

Therefore, using the fact $\omega(\langle 2,2,2\rangle, \langle 2\rangle) = \frac{1}{2}$ from (13), we have

$$\omega(\langle 2\rangle, t)\,\omega(t, \langle 2,2,2\rangle) \geq \frac{\mathbf{i}(t)}{\omega(\langle 2,2,2\rangle, \langle 2\rangle)} = 2\,\mathbf{i}(t). \tag{28}$$

This proves the claim. ◀

Theorem 9, in particular, implies that if $\mathbf{i}(t) > 1$, then $\omega(\langle 2\rangle, t)\,\omega(t, \langle 2,2,2\rangle) > 2$, i.e. one cannot prove $\omega = 2$ via any fixed irreversible intermediate tensor. (Of course one can consider sequences of intermediate tensors with irreversibility converging to 1.)

3.2 Better barriers through more structure

Naturally, we should expect that imposing more structure on the approach to upper bound ω leads to stronger barriers. In this section we impose that the final step of the approach is an application of the Schönhage τ-theorem. The Schönhage τ-theorem (Strassen's general version [22]) says that

$$\underset{\sim}{\mathrm{R}}\Big(\bigoplus_{i=1}^{q}\langle a_i, b_i, c_i\rangle\Big) \geq \sum_{i=1}^{q}(a_i b_i c_i)^{\omega/3}. \tag{29}$$

In particular holds

$$\underset{\sim}{\mathrm{R}}\big(\langle q\rangle \otimes \langle a,a,a\rangle\big) \geq q a^{\omega}. \tag{30}$$

Therefore, in the language of rates, for any $\alpha, \beta \in \mathbb{N}$ holds that

$$\omega(\langle 2\rangle, \langle 2\rangle^{\alpha}\langle 2,2,2\rangle^{\beta}) \geq \alpha + \beta\omega \tag{31}$$

that is

$$\frac{\omega(\langle 2\rangle, \langle 2\rangle^{\alpha}\langle 2,2,2\rangle^{\beta}) - \alpha}{\beta} \geq \omega. \tag{32}$$

(Here α corresponds to $\log_2 q$ and β corresponds to $\log_2 a$. For simplicity and concreteness we will consider only integer α and β.) Thus for any tensor t and for any $\alpha, \beta \in \mathbb{N}$ holds that

$$\frac{\omega(\langle 2\rangle, t)\, \omega(t, \langle 2\rangle^{\alpha} \langle 2,2,2\rangle^{\beta}) - \alpha}{\beta} \geq \omega. \tag{33}$$

We prove the following barrier in terms of α, β and the irreversibility $\mathbf{i}(t)$ of t.

▶ **Theorem 10.** *For any tensor t and $\alpha, \beta \in \mathbb{N}$ holds*

$$\frac{\omega(\langle 2\rangle, t)\, \omega(t, \langle 2\rangle^{\alpha} \langle 2,2,2\rangle^{\beta}) - \alpha}{\beta} \geq 2\,\mathbf{i}(t) + \frac{\alpha}{\beta}\big(\mathbf{i}(t) - 1\big) \geq 2\,\mathbf{i}(t). \tag{34}$$

Proof. By the triangle inequality,

$$\omega(\langle 2\rangle, t)\, \omega(t, \langle 2\rangle^{\alpha} \langle 2,2,2\rangle^{\beta})\, \omega(\langle 2\rangle^{\alpha} \langle 2,2,2\rangle^{\beta}, \langle 2\rangle) \geq \omega(\langle 2\rangle, t)\, \omega(t, \langle 2\rangle) = \mathbf{i}(t). \tag{35}$$

Therefore,

$$\omega(\langle 2\rangle, t)\, \omega(t, \langle 2\rangle^{\alpha} \langle 2,2,2\rangle^{\beta}) \geq \frac{\mathbf{i}(t)}{\omega(\langle 2\rangle^{\alpha} \langle 2,2,2\rangle^{\beta}, \langle 2\rangle)} = (\alpha + 2\beta)\,\mathbf{i}(t). \tag{36}$$

Subtracting α, dividing by β and using that $\mathbf{i}(t) - 1 \geq 0$ (Proposition 5) gives the barrier

$$\frac{\omega(\langle 2\rangle, t)\, \omega(t, \langle 2\rangle^{\alpha} \langle 2,2,2\rangle^{\beta}) - \alpha}{\beta} \geq \frac{(\alpha + 2\beta)\,\mathbf{i}(t) - \alpha}{\beta} = 2\,\mathbf{i}(t) + \frac{\alpha}{\beta}(\mathbf{i}(t) - 1) \geq 2\,\mathbf{i}(t). \tag{37}$$

This proves the claim. ◀

As a corollary of the above theorem we present a barrier on any approach of the following form. The Schönhage τ-theorem implies that for any $a, b, c \in \mathbb{N}_{\geq 1}$ and any tensor t holds

$$\frac{\omega(\langle 2\rangle, t)\, \omega(t, \langle 2\rangle^{\alpha} \langle a,b,c\rangle) - \alpha}{\frac{1}{3}\log_2(abc)} \geq \omega. \tag{38}$$

We prove the following barrier in terms of a, b, c, α and the irreversibility of the cyclically symmetrized $\mathrm{cyc}(t) \coloneqq t \otimes ((1,2,3)\cdot t) \otimes ((1,2,3)^2 \cdot t)$.

▶ **Corollary 11.** *For any tensor t and $\alpha \in \mathbb{N}$ and $a, b, c \in \mathbb{N}_{\geq 1}$ holds*

$$\begin{aligned} \frac{\omega(\langle 2\rangle, t)\, \omega(t, \langle 2\rangle^{\alpha} \langle a,b,c\rangle) - \alpha}{\frac{1}{3}\log_2(abc)} &\geq 2\,\mathbf{i}(\mathrm{cyc}(t)) + \frac{\alpha}{\frac{1}{3}\log_2(abc)}(\mathbf{i}(\mathrm{cyc}(t)) - 1) && (39)\\ &\geq 2\,\mathbf{i}(\mathrm{cyc}(t)). && (40) \end{aligned}$$

One verifies that $\mathbf{i}(t) \geq \mathbf{i}(\mathrm{cyc}(t))$. If t is cyclically symmetric, then $\mathrm{cyc}(t) = t^{\otimes 3}$ and we have the equality $\mathbf{i}(t) = \mathbf{i}(\mathrm{cyc}(t))$.

Proof. One verifies directly that $\omega(\langle 2\rangle, t) \geq \omega(\langle 2\rangle, \mathrm{cyc}(t)^{\frac{1}{3}})$ and

$$\omega(t, \langle 2\rangle^{\alpha} \langle a,b,c\rangle) \geq \omega(\mathrm{cyc}(t)^{\frac{1}{3}}, \langle 2\rangle^{\alpha} \langle 2,2,2\rangle^{\frac{1}{3}\log_2(abc)}). \tag{41}$$

Note that we are using real powers here, which is justified by taking powers of the relevant tensors and taking a limit. Using both inequalities and then applying Theorem 10 gives

$$\begin{aligned} \frac{\omega(\langle 2\rangle, t)\, \omega(t, \langle 2\rangle^{\alpha} \langle a,b,c\rangle) - \alpha}{\frac{1}{3}\log_2(abc)} &\geq \frac{\omega(\langle 2\rangle, \mathrm{cyc}(t))\, \omega(\mathrm{cyc}(t), \langle 2\rangle^{\alpha} \langle 2,2,2\rangle^{\frac{1}{3}\log_2(abc)}) - \alpha}{\frac{1}{3}\log_2(abc)} && (42)\\ &\geq 2\,\mathbf{i}(\mathrm{cyc}(t)). && (43) \end{aligned}$$

This proves the statement of the theorem. ◀

▶ Remark 12. For cyclically symmetric tensors t our Corollary 11 implies the lower bound

$$\omega_g(t) \geq \omega_u(t) \geq 2\,\mathbf{i}(t), \tag{44}$$

on the parameter ω_g (and the "universal" version ω_u) studied in [3], which is a significant improvement over the barrier

$$\omega_g(t) \geq \frac{3}{\frac{1}{2\,\mathbf{i}(t)} + 1} \tag{45}$$

proven in [3, Theorem IV.1].

3.3 Better barriers through monomial irreversibility

Finally, we impose as an extra constraint that the transformation from the intermediate tensor t to the matrix multiplication tensor happens via monomial restriction (Section 2.4), i.e. we consider the approach

$$\omega(\langle 2\rangle, t)\,\omega_{\mathrm{M}}(t, \langle 2,2,2\rangle) \geq \omega \tag{46}$$

and the more structured approaches

$$\frac{\omega(\langle 2\rangle, t)\,\omega_{\mathrm{M}}(t, \langle 2\rangle^{\alpha}\langle 2,2,2\rangle^{\beta}) - \alpha}{\beta} \geq \omega \tag{47}$$

and

$$\frac{\omega(\langle 2\rangle, t)\,\omega_{\mathrm{M}}(t, \langle 2\rangle^{\alpha}\langle a,b,c\rangle) - \alpha}{\frac{1}{3}\log_2(abc)} \geq \omega. \tag{48}$$

The proofs in the previous sections can be directly adapted to prove:

▶ **Theorem 13.** *For any tensor t holds*

$$\omega(\langle 2\rangle, t)\,\omega_{\mathrm{M}}(t, \langle 2,2,2\rangle) \geq 2\,\mathbf{i}_{\mathrm{M}}(t). \tag{49}$$

▶ **Theorem 14.** *For any tensor t and $\alpha, \beta \in \mathbb{N}$ holds*

$$\frac{\omega(\langle 2\rangle, t)\,\omega_{\mathrm{M}}(t, \langle 2\rangle^{\alpha}\langle 2,2,2\rangle^{\beta}) - \alpha}{\beta} \geq 2\,\mathbf{i}_{\mathrm{M}}(t) + \frac{\alpha}{\beta}\big(\mathbf{i}_{\mathrm{M}}(t) - 1\big) \geq 2\,\mathbf{i}_{\mathrm{M}}(t). \tag{50}$$

▶ **Corollary 15.** *For any tensor t and $\alpha \in \mathbb{N}$ and $a, b, c \in \mathbb{N}_{\geq 1}$ holds*

$$\frac{\omega(\langle 2\rangle, t)\,\omega_{\mathrm{M}}(t, \langle 2\rangle^{\alpha}\langle a,b,c\rangle) - \alpha}{\frac{1}{3}\log_2(abc)} \geq 2\,\mathbf{i}_{\mathrm{M}}(\mathrm{cyc}(t)) + \frac{\alpha}{\frac{1}{3}\log_2(abc)}(\mathbf{i}_{\mathrm{M}}(\mathrm{cyc}(t)) - 1) \tag{51}$$

$$\geq 2\,\mathbf{i}_{\mathrm{M}}(\mathrm{cyc}(t)). \tag{52}$$

4 Explicit irreversibility lower bounds

We have seen how barriers arise from lower bounds on irreversibility. In this section we compute lower bounds on the irreversibility of two well-known intermediate tensors that play a crucial role in the best upper bounds on ω: the small and big Coppersmith–Winograd tensors.

4.1 Irreversibility and the asymptotic spectrum of tensors

We begin with a general discussion of how to compute irreversibility. The asymptotic spectrum of tensors is the set of $\leq$-monotone semiring homomorphisms from the semiring of tensors (with tensor product and direct sum as multiplication and addition) to the nonnegative reals,

$$\Delta = \{F \in \mathrm{Hom}(\{\text{tensors}\}, \mathbb{R}_{\geq 0}) : a \leq b \Rightarrow F(a) \leq F(b)\}. \tag{53}$$

Strassen proves in [22] that $\underaccent{\sim}{\mathrm{Q}}(t) = \min_{F\in\Delta} F(t)$ and $\underaccent{\sim}{\mathrm{R}}(t) = \max_{F\in\Delta} F(t)$ and he also proves (implicitly) that $\omega(s,t) = \max_{F\in\Delta} \log_2 F(t)/\log_2 F(s)$. From this we directly obtain:

▶ **Proposition 16.** *Let t be a tensor. Then*

$$\mathbf{i}(t) = \frac{\max_{F\in\Delta} \log F(t)}{\min_{F\in\Delta} \log F(t)}. \tag{54}$$

In an ideal world we would know Δ and use it to compute $\mathbf{i}(t)$ (or better, we would use it to compute ω). In practice we currently only have partial knowledge of Δ. This partial knowledge is easiest to describe in terms of the best known lower bounds on $\underaccent{\sim}{\mathrm{R}}(t)$ and the best known upper bounds on $\underaccent{\sim}{\mathrm{Q}}(t)$. The best known lower bounds on $\underaccent{\sim}{\mathrm{R}}(t)$ are simply the matrix ranks of each of the three flattenings t_1, t_2, t_3 of t as described in Section 2.5. For arbitrary fields, the best general upper bounds on $\underaccent{\sim}{\mathrm{Q}}(t)$ that we are aware of are the Strassen upper support functionals ζ^θ from [23], which we will define and use in the next section. They relate asymptotically to slice rank via [9]

$$\underaccent{\sim}{\mathrm{Q}}(t) \leq \limsup_n \operatorname{slicerank}(t^{\otimes n})^{1/n} \leq \min_\theta \zeta^\theta(t). \tag{55}$$

We are not aware of any example for which any of the inequalities in (55) is strict. For oblique tensors [1] the right inequality is an equality [9] and for tight tensors [2] both inequalities are equalities [23]. We thus have:

▶ **Proposition 17.** *Let t be a tensor. Then*

$$\mathbf{i}(t) \geq \frac{\max_i \log_2 \mathrm{R}(t_i)}{\min_\theta \log_2 \zeta^\theta(t)}. \tag{56}$$

For tensors over the complex numbers (i.e. $\mathbb{F} = \mathbb{C}$) we have a deeper understanding of the theory of upper bounds on the asymptotic subrank, via the quantum functionals F^θ introduced in [9]. The quantum functionals satisfy $F^\theta \leq \zeta^\theta$ and their minimum equals the asymptotic slice rank [9], i.e.

$$\underaccent{\sim}{\mathrm{Q}}(t) \leq \limsup_n \operatorname{slicerank}(t^{\otimes n})^{1/n} = \min_\theta F^\theta(t) \leq \min_\theta \zeta^\theta(t). \tag{57}$$

For free tensors[3] the right inequality in (57) is an equality [9]. We thus have:

▶ **Proposition 18.** *Let t be a tensor over the complex numbers. Then*

$$\mathbf{i}(t) \geq \frac{\max_i \log_2 \mathrm{R}(t_i)}{\min_\theta \log_2 F^\theta(t)}. \tag{58}$$

[1] a tensor $t \in \mathbb{F}^{n_1} \otimes \mathbb{F}^{n_2} \otimes \mathbb{F}^{n_3}$ is called oblique if the support $\mathrm{supp}(t) \in [n_1] \times [n_2] \times [n_3]$ in some basis is an antichain in the product of the natural orders on the $[n_i]$

[2] a tensor t is called tight if for some choice of basis there are injective maps $\alpha_1, \alpha_2, \alpha_3$ such that for every $a \in \mathrm{supp}(t)$ holds $\alpha_1(a_1) + \alpha_2(a_2) + \alpha_3(a_3) = 0$

[3] a tensor t is called free if in some basis any two different $a, b \in \mathrm{supp}(t)$ differ in at least two entries

4.2 Irreversibility of Coppersmith–Winograd tensors

We now compute lower bounds for the irreversibility of the Coppersmith–Winograd tensors. As mentioned, we will use the support functionals of Strassen [23] in our computation to upper bound the asymptotic subrank. For any $\theta \in \mathbb{R}^3_{\geq 0}$ with $\theta_1 + \theta_2 + \theta_3 = 1$ the upper support functional ζ^θ is defined as

$$\zeta^\theta(t) \coloneqq 2^{\rho^\theta(t)} \tag{59}$$

$$\rho^\theta(t) \coloneqq \min_{s \cong t} \max_{P \in \mathcal{P}(\mathrm{supp}(s))} \sum_{i=1}^{3} \theta_i H(P_i), \tag{60}$$

where the minimum is over all tensors s isomorphic to t, the maximum is over all probability distributions on the support of s in the standard basis, and $H(P_i)$ denotes the Shannon entropy of the ith marginal of P. Strassen proves in [23] that $1/\omega(t, \langle 2 \rangle) = \log_2 \underset{\sim}{\mathrm{Q}}(t) \leq \rho^\theta(t)$.

(Besides from the Strassen support functionals, upper bound on the asymptotic subrank of complex tensors may be obtained from the quantum functionals. For the tensors in Theorem 19 and Theorem 22, however, this will give the same bound, since these tensors are free [9, Section 4.3].)

▶ **Theorem 19** (Small Coppersmith–Winograd tensors [12, Section 6]). *For the small Coppersmith–Winograd tensor*

$$\mathrm{cw}_q \coloneqq \sum_{i=1}^{q} e_{0,i,i} + e_{i,0,i} + e_{i,i,0} \tag{61}$$

the lower bound

$$2\,\mathbf{i}(\mathrm{cw}_q) \geq \frac{2\log_2(q+1)}{\log_2 3 - \frac{2}{3} + \frac{2}{3}\log_2 q} \tag{62}$$

holds.

Proof. The rank of each flattening of cw_q equals $q+1$. Therefore, $\underset{\sim}{\mathrm{R}}(\mathrm{cw}_q) \geq q+1$. To upper bound the asymptotic subrank $\underset{\sim}{\mathrm{Q}}(\mathrm{cw}_q)$ one can upper bound the Strassen upper support functional with $\theta = (1/3, 1/3, 1/3)$ as in [9, Example 4.22] by

$$\rho^\theta(\mathrm{cw}_q) \leq \log_2 3 - \frac{2}{3} + \frac{2}{3}\log_2 q. \tag{63}$$

We find that

$$\mathbf{i}(\mathrm{cw}_q) \geq \frac{\log_2(q+1)}{\log_2 3 - \frac{2}{3} + \frac{2}{3}\log_2 q}. \tag{64}$$

This proves the theorem. ◀

▶ Remark 20. If $q > 2$, then the right-hand side of (62) is at least 2.02.. See the table in Section 1 for more values. If $q = 2$, however, then the right-hand side of (62) equals 2. Theorem 19 thus does not rule out using cw_2 to prove that $\omega = 2$. Indeed, as observed in [12, Section 11]), if $\omega(\langle 2 \rangle, \mathrm{cw}_2) = \log_2 3$, then $\omega = 2$.

Currently, the best upper bound we have on $\omega(\langle 2 \rangle, \mathrm{cw}_q)$ is $\log_2(q+2)$. If $\omega(\langle 2 \rangle, \mathrm{cw}_q) = \log_2(q+2)$, then instead of (62) we get the better barrier

$$2\,\mathbf{i}(\mathrm{cw}_q) \geq \frac{2\log_2(q+2)}{\log_2 3 - \frac{2}{3} + \frac{2}{3}\log_2 q}. \tag{65}$$

The right-hand side of (65) has a minimum value of

$$\frac{18}{5\log_2 3} = 2.27.. \tag{66}$$

attained at $q = 6$.

▶ Remark 21. The following computation serves as a sanity check for our barrier. Namely we see in an example how by putting some extra assumption the barrier becomes tight. Coppersmith and Winograd in [12] used cw_q as an intermediate tensor in combination with the laser method and a certain "outer structure", see also [6, Section 9]. When we impose that we apply the laser method on cw_q with this outer structure to upper bound ω we get the following better barrier via Theorem 10/Corollary 11:

$$2\,\mathbf{i}(\mathrm{cw}_q) + \frac{h(1/3)}{\frac{1}{3}\log_2(q)}(\mathbf{i}(\mathrm{cw}_q) - 1) \tag{67}$$

where

$$\mathbf{i}(\mathrm{cw}_q) \geq \frac{\log_2(q+1)}{\log_2(3) - \frac{2}{3} + \frac{2}{3}\log_2(q)}. \tag{68}$$

Some values of (67) are:

q	
2	2
3	2.04..
4	2.10..
5	2.15..
6	2.19..
7	2.22..

If in addition we assume that $\omega(\langle 2\rangle, \mathrm{cw}_q) = \log_2(q+2)$, then we obtain the barrier

$$2\,\mathbf{i}(\mathrm{cw}_q) + \frac{h(1/3)}{\frac{1}{3}\log_2(q)}(\mathbf{i}(\mathrm{cw}_q) - 1) \tag{69}$$

where

$$\mathbf{i}(\mathrm{cw}_q) \geq \frac{\log_2(q+2)}{\log_2(3) - \frac{2}{3} + \frac{2}{3}\log_2(q)}. \tag{70}$$

Some values of (69) are:

q	
2	3.24..
3	2.65..
4	2.50..
5	2.44..
6	2.41..
7	2.40..
8	2.40..
9	2.40..
10	2.40..
11	2.41..

with minimum value of 2.40... These barriers in fact match the upper bound

$$\omega \le \log_q \frac{4(q+2)^3}{27} \tag{71}$$

that was obtained by Coppersmith and Winograd by applying the laser method in the way described above. Other intermediate tensors with a given outer structure may be analyzed similarly.

▶ **Theorem 22** (Big Coppersmith–Winograd tensors [12, Section 7]). *For the big Coppersmith–Winograd tensor*

$$\mathrm{CW}_q \coloneqq e_{0,0,q+1} + e_{0,q+1,0} + e_{q+1,0,0} + \sum_{i=1}^{q} e_{0,i,i} + e_{i,0,i} + e_{i,i,0} \tag{72}$$

the lower bound

$$2\,\mathbf{i}(\mathrm{CW}_q) \ge \begin{cases} \dfrac{2\log_2(3)}{f(\frac{1}{18}(\sqrt{33}-3))} = 2.16.. & q = 1 \\ \dfrac{2\log_2(4)}{f(\frac{1}{9})} = 2.17.. & q = 2 \\ \dfrac{2\log_2(q+2)}{f\big(\frac{3q-\sqrt{32+q^2}}{6(q^2-4)}\big)} & q \ge 3 \end{cases} \tag{73}$$

holds, where

$$f(x) \coloneqq -\Big(\frac{2}{3} - qx\Big)\log_2\Big(\frac{2}{3} - qx\Big) - q2x\log_2(2x) - \Big(\frac{1}{3} - qx\Big)\log_2\Big(\frac{1}{3} - qx\Big). \tag{74}$$

Proof. The rank of each flattening of CW_q equals $q+2$, which coincides with the well-known border rank upper bound $\underline{\mathrm{R}}(\mathrm{CW}_q) \le q+2$. Therefore, $\underset{\sim}{\mathrm{R}}(\mathrm{CW}_q) = q+2$.

To upper bound the asymptotic subrank $\underset{\sim}{\mathrm{Q}}(\mathrm{CW}_q)$ we use the Strassen upper support functional with $\theta = (1/3, 1/3, 1/3)$. In the standard basis, the support of CW_q is the set

$$\{(0,i,i),(i,0,i),(i,i,0) : i \in [q]\} \cup \{(0,0,q+1),(0,q+1,0),(q+1,0,0)\}. \tag{75}$$

The symmetry implies that we can assign probability x to each of $(0,i,i)$, $(i,0,i)$ and $(0,i,i)$, and $\frac{1}{3} - qx$ to $(0,0,q+1)$, $(0,q+1,0)$ and $(q+1,0,0)$. This leads to an average marginal entropy of $f(x)$ as defined in the theorem statement. The maximum of $f(x)$ is attained at

$$x = \begin{cases} \frac{1}{18}(\sqrt{33}-3) & q = 1 \\ \frac{1}{9} & q = 2 \\ \dfrac{3q - \sqrt{32+q^2}}{6(q^2-4)} & q \ge 3. \end{cases} \tag{76}$$

This proves the theorem. ◀

▶ Remark 23. The lowest value of the right-hand side of (73) is 2.16.. attained at $q = 1$. See the table in Section 1 for more values.

▶ Remark 24. A lower bound on the irreversibility of the tensors t_n mentioned in the introduction follows directly from the results in [23, Theorem 6.7].

4.3 Monomial irreversibility of structure tensors of finite group algebras

We now discuss irreversibility and monomial irreversibility in the context of the group-theoretic approach developed in [10]. This approach produces upper bounds on ω via intermediate tensors that are structure tensors of complex group algebras of finite groups. Let $\langle G\rangle$ denote the structure tensor of the complex group algebra $\mathbb{C}[G]$ of the finite group G, in the standard basis. The group-theoretic approach (in particular [10, Theorem 4.1]) produces an inequality of the form

$$\langle G\rangle \geq_{\mathrm{M}} \langle a, b, c\rangle \tag{77}$$

which ultimately (see [10, Eq. (1)]) leads to the bound

$$\frac{\omega(\langle 2\rangle, \langle G\rangle)\, \omega_{\mathrm{M}}(\langle G\rangle, \langle a, b, c\rangle)}{\frac{1}{3}\log_2(abc)} \geq \omega \tag{78}$$

where $\geq_{\mathrm{M}}$ and ω_{M} are the monomial restriction and monomial relative exponent defined in Section 2.4.

Now the monomial irreversibility barrier from Section 3.3 comes into play. Upper bounds on the monomial asymptotic subrank of $\langle G\rangle$ have (using different terminology) been obtained in [7, 8, 18]. Those upper bounds imply that $\langle G\rangle$ is monomially irreversible for every nontrivial finite group G. Together with our results in Section 3.3 and the fact that the tensor $\langle G\rangle$ is symmetric up to a permutation of the basis of one of the tensor legs, this directly leads to nontrivial barriers for the left-hand side of (78) for any fixed nontrivial group G, thus putting the work of [7, 8, 18] in a broader context. We have not tried to numerically optimise the monomial irreversibility barriers for group algebras.

Finally we mention that the irreversibility barrier (rather than the monomial irreversibility barrier) does not rule out obtaining $\omega = 2$ via $\langle G\rangle$. Namely, $\langle G\rangle$ is isomorphic to a direct sum of matrix multiplication tensors, $\langle G\rangle \cong \bigoplus_i \langle d_i, d_i, d_i\rangle$ and, therefore, we have $\mathbf{i}(\langle G\rangle) = (\log_2 \sum_i d_i^{\omega})/(\log_2 \sum_i d_i^2)$. Thus, if $\omega = 2$, then $\langle G\rangle$ is reversible.

References

1. Josh Alman. Limits on the Universal Method for Matrix Multiplication. *arXiv*, 2018. `arXiv:1812.08731`.
2. Josh Alman and Virginia Vassilevska Williams. Further Limitations of the Known Approaches for Matrix Multiplication. In Anna R. Karlin, editor, *9th Innovations in Theoretical Computer Science Conference (ITCS 2018)*, volume 94 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 25:1–25:15, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. `doi:10.4230/LIPIcs.ITCS.2018.25`.
3. Josh Alman and Virginia Vassilevska Williams. Limits on All Known (and Some Unknown) Approaches to Matrix Multiplication. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 580–591, October 2018. `doi:10.1109/FOCS.2018.00061`.
4. Andris Ambainis, Yuval Filmus, and François Le Gall. Fast matrix multiplication: limitations of the Coppersmith-Winograd method (extended abstract). In *STOC'15—Proceedings of the 2015 ACM Symposium on Theory of Computing*, pages 585–593. ACM, New York, 2015. `doi:10.1145/2746539.2746554`.
5. Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and Ashish V. Thapliyal. Exact and asymptotic measures of multipartite pure-state entanglement. *Phys. Rev. A*, 63(1):012307, 2000. `doi:10.1103/PhysRevA.63.012307`.
6. Markus Bläser. *Fast Matrix Multiplication.* Number 5 in Graduate Surveys. Theory of Computing Library, 2013. `doi:10.4086/toc.gs.2013.005`.

7 Jonah Blasiak, Thomas Church, Henry Cohn, Joshua A. Grochow, Eric Naslund, William F. Sawin, and Chris Umans. On cap sets and the group-theoretic approach to matrix multiplication. *Discrete Anal.*, 2017. doi:10.19086/da.1245.

8 Jonah Blasiak, Thomas Church, Henry Cohn, Joshua A Grochow, and Chris Umans. Which groups are amenable to proving exponent two for matrix multiplication? *arXiv*, 2017. arXiv:1712.02302.

9 Matthias Christandl, Péter Vrana, and Jeroen Zuiddam. Universal points in the asymptotic spectrum of tensors. In *Proceedings of 50th Annual ACM SIGACT Symposium on the Theory of Computing (STOC'18)*. ACM, New York, 2018. doi:10.1145/3188745.3188766.

10 Henry Cohn and Christopher Umans. A group-theoretic approach to fast matrix multiplication. In *Foundations of Computer Science, 2003. Proceedings. 44th Annual IEEE Symposium on*, pages 438–449. IEEE, 2003. doi:10.1109/SFCS.2003.1238217.

11 Henry Cohn and Christopher Umans. Fast matrix multiplication using coherent configurations. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1074–1086. SIAM, 2013. URL: http://dl.acm.org/citation.cfm?id=2627817.2627894.

12 Don Coppersmith and Shmuel Winograd. Matrix multiplication via arithmetic progressions. *J. Symbolic Comput.*, 9(3):251–280, 1990. doi:10.1016/S0747-7171(08)80013-2.

13 Wolfgang Dür, Guivre Vidal, and Juan Ignacio Cirac. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A (3)*, 62(6):062314, 12, 2000. doi:10.1103/PhysRevA.62.062314.

14 Jordan S. Ellenberg and Dion Gijswijt. On large subsets of $\mathbf{F}_q^n$ with no three-term arithmetic progression. *Ann. of Math. (2)*, 185(1):339–343, 2017. doi:10.4007/annals.2017.185.1.8.

15 Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Modern Phys.*, 81(2):865–942, 2009. doi:10.1103/RevModPhys.81.865.

16 J.M. Landsberg and Mateusz Michałek. Abelian tensors. *Journal de Mathématiques Pures et Appliquées*, 108(3):333–371, 2017. doi:10.1016/j.matpur.2016.11.004.

17 François Le Gall. Powers of tensors and fast matrix multiplication. In *ISSAC 2014—Proceedings of the 39th International Symposium on Symbolic and Algebraic Computation*, pages 296–303. ACM, New York, 2014. doi:10.1145/2608628.2608664.

18 Will Sawin. Bounds for matchings in nonabelian groups. *arXiv*, 2017. arXiv:1702.00905.

19 Andrew James Stothers. *On the complexity of matrix multiplication*. PhD thesis, University of Edinburgh, 2010. URL: http://hdl.handle.net/1842/4734.

20 Volker Strassen. Gaussian elimination is not optimal. *Numer. Math.*, 13(4):354–356, 1969. doi:10.1007/BF02165411.

21 Volker Strassen. Relative bilinear complexity and matrix multiplication. *J. Reine Angew. Math.*, 375/376:406–443, 1987. doi:10.1515/crll.1987.375-376.406.

22 Volker Strassen. The asymptotic spectrum of tensors. *J. Reine Angew. Math.*, 384:102–152, 1988. doi:10.1515/crll.1988.384.102.

23 Volker Strassen. Degeneration and complexity of bilinear maps: some asymptotic spectra. *J. Reine Angew. Math.*, 413:127–180, 1991. doi:10.1515/crll.1991.413.127.

24 Terence Tao. A symmetric formulation of the Croot–Lev–Pach–Ellenberg–Gijswijt capset bound, 2016. URL: https://terrytao.wordpress.com.

25 F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde. Four qubits can be entangled in nine different ways. *Phys. Rev. A (3)*, 65(5, part A):052112, 5, 2002. doi:10.1103/PhysRevA.65.052112.

26 Péter Vrana and Matthias Christandl. Asymptotic entanglement transformation between W and GHZ states. *J. Math. Phys.*, 56(2):022204, 12, 2015. doi:10.1063/1.4908106.

27 Virginia Vassilevska Williams. Multiplying matrices faster than Coppersmith-Winograd. Extended abstract. In *STOC'12—Proceedings of the 2012 ACM Symposium on Theory of Computing*, pages 887–898. ACM, New York, 2012. doi:10.1145/2213977.2214056.

28 Nengkun Yu, Cheng Guo, and Runyao Duan. Obtaining a W state from a Greenberger-Horne-Zeilinger state via stochastic local operations and classical communication with a rate approaching unity. *Physical review letters*, 112(16):160401, 2014. doi:10.1103/PhysRevLett.112.160401.