ON A REVERSE ℓ_2–INEQUALITY FOR SPARSE CIRCULAR CONVOLUTIONS

Philipp Walk

Institute of Theoretical Information Technology, Technical University of Munich (TUM)

philipp.walk@tum.de

Peter Jung

Heinrich-Hertz-Chair for Information Theory and Theoretical Information Technology, Technical University of Berlin

peter.jung@mk.tu-berlin.de

ABSTRACT

In this paper we show that convolutions of sufficiently sparse signals always admit a non–zero lower bound in energy if oversampling of its Fourier transform is employed. This bound is independent of the signals and the ambient dimension and is determined only be the sparsity of both input signals. This result has several implications for blind system and signal identification and detection, noncoherent communication of sporadic and short–message type user data and strategies for its compressive reception. Furthermore, we give some first insights into the combinatorial nature of this problem, its scaling behavior and present numerical results as well.

***Index Terms*—** Circular Convolution, Sparsity, Young Inequality, Discrete Uncertainty Principle

1. INTRODUCTION

Starting with the first fundamental results in compressed sensing [1] it is known nowadays that the geometry of compressible signals can be used for an impressive reduction of the sampling rate down to the order of its information content. This includes sparse signals and signals with certain power law decay in its magnitude–ordered components.

A new objective is the characterization of operations between compressible signals and the determination of the complexity of the resulting output set [2, 3, 4, 5]. This situation has been considered for *bilinear mappings* in the paper [2] and an important case here is the convolution of two signals both being sparse in the canonical basis. The rate of non–adaptive compression of the output set depends strongly on its geometry and this in turn is determined by the bilinear mapping itself. A peculiar property that implies a sampling complexity being additive in the sparsity of both inputs is that the ℓ_2–norm of the output can be related to the product of the ℓ_2–norms of the inputs[1]. For convolutions this refers to a reversed version of the Young inequality which is known only for positive signals [6]. The goal of this paper is to indicate such an input–output energy equivalence for sufficiently sparse inputs. This is a surprising result, since this equivalence does not exists for arbitrary input signals.

Independently of the question of compressive reception, this input–output–property has also further important and practical applications. A typical situation in wireless communication, for example, is that one contribution is some possibly random but known probe signal [7] or even an unknown sparse signal containing a short message and the second contribution might be a time–invariant channel with a quite small number of taps. From the discrete uncertainty principle [8] one could argue heuristically that with decreasing support of both signals its Fourier transforms will get more and more overlapping supports. Hence, with improved input sparsity of the signals its convolution will become increasingly observable in additive noise only by energy. For example, in blind system identification and signal detection both signals have to be identified based on observations of its convolution [9]. This noncoherent communication approach will become more important since in many new wireless applications (like for example "internet of things" and "machine type communication") only short messages like status updates are transmitted through an unknown but sparse channel and should be decoded without performing complex channel estimation, feedback and channel–aware transmitter strategies [10].

The paper is organized as follows: In Section 2 we reformulate sparse convolutions as certain linear mappings on tensor products. We state our main results in Section 3 which includes a reversed ℓ_2–inequality, combinatorial statements and insights from discrete uncertainty principles. Furthermore, we show that the optimal constants are given as the minima of bi–quadratic optimization problems, which are usually NP–hard. In Section 4 we sketch the proof of the inequality and discuss the scaling of the optimal constants in terms of sparsity. Finally, we present in Section 5 a numerical algorithm that approximates the constants up to a desired accuracy but with exponential complexity.

2. SPARSE CIRCULAR CONVOLUTION

Let $\mathbf{F}$ be the Fourier matrix having elements $[\mathbf{F}]_{lk} = \frac{1}{\sqrt{n}} e^{-2\pi\imath\frac{lk}{n}}$ for $l, k \in N$ with $N := \{0, \dots, n-1\}$. The circular convolution $\mathbf{x} \circledast \mathbf{y} \in \mathbb{R}^n$ of two (real) vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is given [11] by:

$$\mathbf{x} \circledast \mathbf{y} := \sqrt{n}\mathbf{F}^* \text{diag}(\mathbf{F}\mathbf{x})\mathbf{F}\mathbf{y} =: \mathbf{X}\mathbf{y}. \tag{1}$$

As well–known, the circulant matrix $\mathbf{X}$ can be diagonalized by the discrete Fourier transform $\mathbf{F}$ and is given in terms of the (right) $n \times n$ shift matrices $\mathbf{S}^i$ with elements $(\mathbf{S}^i)_{kl} = \delta_{k,l\oplus i}$ as:

$$\mathbf{X} = \sum_{i=0}^{n-1} x_i \mathbf{S}^i. \tag{2}$$

$\oplus$ denotes the addition modulo n and $\delta_{i,j} = 0$ for $i \neq j$ and 1 else.

On the other hand, since $\mathbf{x} \circledast \mathbf{y}$ is bilinear in both inputs, the circular convolution can be described by standard lifting as a linear mapping $\mathbf{S}^N : \mathbb{R}^{n \otimes n} \to \mathbb{R}^n$ from tensor products $\mathbf{x} \otimes \mathbf{y}$ or, equivalently, from rank–one matrices $\mathbf{x}\mathbf{y}^*$ to $\mathbb{R}^n$. In the canonical basis $\{\mathbf{e}_k \otimes \mathbf{e}_l\}_{k,l=0}^{n-1}$, where $[\mathbf{e}_k]_l := \delta_{kl}$, the mapping is represented by

[1] In [2] we have called this "restricted norm multiplicativity"(RNMP)

the $n \times n^2$–matrix:

$$\mathbf{S}^N = [\mathbb{1}_n\ \mathbf{S}^1\ \cdots\ \mathbf{S}^{n-1}] \in \mathbb{R}^{n \times n^2}. \tag{3}$$

Consider now two subsets $I, J \subset N$ and the corresponding canonical subspaces $X := \operatorname{span}\{\mathbf{e}_i\}_{i \in I}$ and $Y := \operatorname{span}\{\mathbf{e}_j\}_{j \in J}$ of dimensions $\dim(X) = |I| =: s$ and $\dim(Y) = |J| =: f$. Let be $I = \{i_0, \dots, i_{s-1}\}$. Then the action of $\mathbf{C}$ restricted on $X \otimes Y$ is then determined by the stacked $n \times sf$–matrix:

$$\mathbf{S}_J^I = [\mathbf{S}_J^{i_0}\ \mathbf{S}_J^{i_1}\ \cdots\ \mathbf{S}_J^{i_{s-1}}] \in \mathbb{R}^{n \times sf} \tag{4}$$

where $\mathbf{S}_J^i$ is the $n \times f$–submatrix of $\mathbf{S}^i$ with column indices in J.

3. MAIN RESULT

In this section, we will show for $s \le f \le n$, that all $(s, f)-$sparse circular convolutions embedded in $\mathbb{R}^{\tilde{n}}$ with $\tilde{n} = 2n - 1$ have an universal ℓ^2-norm lower bound. Notation: The sets

$$\Sigma_s = \bigcup_{\substack{I \subset N \\ |I| = s}} \operatorname{span}\{\mathbf{e}_i\}_I \quad , \quad \Sigma_f = \bigcup_{\substack{J \subset N \\ |J| = f}} \operatorname{span}\{\mathbf{e}_i\}_J$$

denote all s-sparse resp. f-sparse vectors in $\mathbb{R}^n$. The support of $\mathbf{x}$ is denoted by $\operatorname{supp} \mathbf{x}$. The ℓ_2-norm is given by $\|\mathbf{x}\| := \sqrt{\sum_i x_i^2}$. We set $\mathbf{0} := (0, \dots, 0)^T \in \mathbb{R}^{n-1}$ and $\tilde{\Sigma}_s := (\Sigma_s, \mathbf{0}) \subset \mathbb{R}^{\tilde{n}}$.

Theorem 1. *Let s,f and n be natural numbers with $1 \le s \le f \le n$. Then there exist a constant $a_m > 0$ with $m = \min\{sf, n\}$, such that for all $\tilde{\mathbf{x}} \in \tilde{\Sigma}_s$ and $\tilde{\mathbf{y}} \in \tilde{\Sigma}_f$ it holds:*

$$a_m \|\tilde{\mathbf{x}}\|^2 \|\tilde{\mathbf{y}}\|^2 \le \|\tilde{\mathbf{x}} \circledast \tilde{\mathbf{y}}\|^2 \le s \|\tilde{\mathbf{x}}\|^2 \|\tilde{\mathbf{y}}\|^2. \tag{5}$$

Moreover, a_m is a strictly decreasing sequence.

Related work: As already noted in the introduction, the lower bound in the theorem refers to a variant of a reverse Young inequality [6]. be a new and fundamental result for sparse convolutions. The condition of appending $n - 1$ zeros to $\mathbf{x}$ and $\mathbf{y}$ seems to be necessary: for $s = f = 2$, $\tilde{n} = 2n - 2$ and every $n \ge 4$ it follows that $\|\tilde{\mathbf{x}} \circledast \tilde{\mathbf{y}}\| = 0$ for $\tilde{\mathbf{x}}, \tilde{\mathbf{y}} \in \mathbb{R}^{\tilde{n}}$ when the non-zero components are $\tilde{x}_0 = \tilde{x}_{n-1} = y_0 = \sqrt{1/2}$ and $\tilde{y}_{n-1} = -\tilde{y}_0$ and therefore $a = 0$. Our result establishes a *restricted norm multiplicativity* (RNMP) [12] for all (s, f)–sparse convolutions on the specific subset $\tilde{\Sigma}_s \times \tilde{\Sigma}_f$. This condition in turn can be used to establish a *restricted isometry property* (RIP) on the output set $\tilde{\Sigma}_s \circledast \tilde{\Sigma}_f$. However, the concrete improvement of the sampling rate for the output signals, as suggested in [3], is still an open question.

3.1. Discrete Uncertainty Principle

Here, we use results on the discrete uncertainty principle of DONOHO-STARK [13] and a refined version of TAO [8] and CHEBOTARËV [14] to motivate our result for groups of prime order without an extra zero padding. Donoho and Stark could show in [13] the following discrete Uncertainty Principle for any $\mathbf{x} \in \mathbb{R}^n$

$$\|\mathbf{x}\|_0 \|\mathbf{F}\mathbf{x}\|_0 \ge n \tag{6}$$

$$\|\mathbf{x}\|_0 + \|\mathbf{F}\mathbf{x}\|_0 \ge 2\sqrt{n} \tag{7}$$

Since $\operatorname{diag}(\mathbf{F}\mathbf{x})\mathbf{F}\mathbf{y} = \mathbf{F}\mathbf{x} \odot \mathbf{F}\mathbf{y}$, where $\odot$ denotes the pointwise product, we can see by definition (1), that

$$\|\mathbf{x} \circledast \mathbf{y}\| = \sqrt{n}\, \|\mathbf{F}\mathbf{x} \odot \mathbf{F}\mathbf{y}\|. \tag{8}$$

Hence, we get the following implication

$$\begin{aligned}(\mathbf{x} \ne \mathbf{0} \ne \mathbf{y}, \|\mathbf{F}\mathbf{x}\|_0 + \|\mathbf{F}\mathbf{y}\|_0 > n) &\Rightarrow (\mathbf{F}\mathbf{x} \odot \mathbf{F}\mathbf{y} \ne 0) \\ &\Leftrightarrow (\mathbf{0} \ne \mathbf{x} \circledast \mathbf{y}) \\ &\Leftrightarrow (0 \ne \|\mathbf{x} \circledast \mathbf{y}\|)\end{aligned} \tag{9}$$

By the DONOHO-STARK inequality (6) and with $\|\mathbf{x}\|_0 \le s$ and $\|\mathbf{y}\|_0 \le f$ in Theorem 1 we get

$$\|\mathbf{F}\mathbf{x}\|_0 + \|\mathbf{F}\mathbf{y}\|_0 \ge \frac{n}{s} + \frac{n}{f} = n\frac{s+f}{sf} \overset{!}{>} n. \tag{10}$$

This is only possible if $s + f > sf$, which holds if and only if $s = 1$ or $f = 1$. But this is trivial. So Donoho-Stark can not provide the existence of $a > 0$ in Theorem 1 for all cases. In fact, if n is prime, the TAO inequality [8]

$$\|\mathbf{x}\|_0 + \|\mathbf{F}\mathbf{x}\|_0 \ge n + 1 \Leftrightarrow \|\mathbf{F}\mathbf{x}\|_0 \ge n + 1 - s \tag{11}$$

yields with the assumption $n \ge s + f - 1$

$$\|\mathbf{F}\mathbf{x}\|_0 + \|\mathbf{F}\mathbf{y}\|_0 \ge 2n + 2 - s - f \ge n + 1 > n. \tag{12}$$

Hence, whenever $1 \le s + f - 1 \le n$ and n is prime, we have for all $\mathbf{x} \in \Sigma_s, \mathbf{y} \in \Sigma_f$ that $\mathbf{x} \circledast \mathbf{y} \ne \mathbf{0} \Leftrightarrow \mathbf{x} \ne \mathbf{0} \ne \mathbf{y}$. Due to the upper bound in (5), which was shown in [12], the map $\|\mathbf{x} \circledast \mathbf{y}\|$ is continuous and hence the infimum a is attained and larger than zero. Nevertheless, we will see in the next section, that solving a is an NP-hard problem.

3.2. Bi-quadratic Optimization and NP–hardness

The optimal lower bound a_{opt} in (5) can be formalized as a non-convex optimization problem. For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ the objective function can be calculated as:

$$\begin{aligned} b(\mathbf{x}, \mathbf{y}) :&= \left\| \sum_{i=0}^{n-1} x_i \mathbf{S}^i \mathbf{y} \right\|^2 = \sum_{m=0}^{n-1} \left(\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} [\mathbf{S}^i]_{mi} x_i y_j \right)^2 \\ &= \sum_{m=0}^{n-1} \left(\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} [\mathbf{S}^i]_{mj} x_i y_j \sum_{i'=0}^{n-1} \sum_{j'=0}^{n-1} [\mathbf{S}^{i'}]_{mj'} x_{i'} y_{j'} \right) \\ &= \sum_{i,i'} \sum_{j,j'} \delta_{i \oplus j, i' \oplus j'} x_i y_j x_{i'} y_{j'} \end{aligned}$$

where we define the tensor $B = (\delta_{i \oplus j, i' \oplus j'})$ of fourth order by

$$\delta_{i \oplus j, i' \oplus j'} := \sum_m [\mathbf{S}^i]_{mj} [\mathbf{S}^{i'}]_{mj'} = \sum_m \delta_{m, i \oplus j} \cdot \delta_{m, i' \oplus j'} \tag{13}$$

Since B is not partially symmetric, we symmetrize by

$$b_{iji'j'} := \frac{1}{2} \left(\delta_{i \oplus j, i' \oplus j'} + \delta_{i' \oplus j, i \oplus j'} \right) \tag{14}$$

using the property $\delta_{i,j} = \delta_{j,i}$. Hence B satisfy

$$\sum_{i,i} \sum_{j,j'} \delta_{i \oplus j, i' \oplus j'} x_i y_j x_{i'} y_{j'} = \sum_{i,i} \sum_{j,j'} b_{iji'j'} x_i \tilde{y}_j x_{i'} y_{j'}$$

with the partially symmetries

$$b_{iji'j'} = b_{i'jij'} = b_{ij'i'j} \quad , \quad 0 \le i, j, i', j' \le n - 1. \tag{15}$$

For $I \subset N$ we define the embedding of $\hat{\mathbf{x}} \in \mathbb{R}^s$ in $\mathbb{R}^n$ by $\mathbf{x}_\alpha := [\mathbf{P}_I^n \hat{\mathbf{x}}]_j := \sum_{\alpha=0}^{s-1} \delta_{j_\alpha, j} \hat{x}_\alpha$ for $j \in N$. We denote for fixed I, J by

$a_{I,J}$ the solution of the *bi-quadratic optimization problem over the bi-sphere* given by

$$(P) \qquad \min_{(\hat{\mathbf{x}},\hat{\mathbf{y}})\in S^{s-1}\times S^{f-1}} b(\mathbf{P}_I^n\hat{\mathbf{x}},\mathbf{P}_J^n\hat{\mathbf{y}}), \tag{16}$$

then the optimal lower bound is given by $a_{\text{opt}} := \min_{I,J} a_{I,J}$. For fixed B, I, J, this problem was well studied in [15, (1.1)] and is by Theorem 2.2. NP-hard. Furthermore the authors in [15] could show that (P) can be written as a *bilinear SDP relaxation* problem, which obtains the same optimal solution by Theorem 2.4. Hence, this relaxation is also NP-hard. An alternative approach is to fix $\mathbf{y}$ and calculate the $n\times n$ matrix $\mathbf{B_y}$, by

$$b_{i,i'}(\mathbf{y}) := \sum_{j,j'} b_{iji'j'}y_jy_{j'} = \sum_{j,j'}\delta_{i\oplus j,i'\oplus j'}y_jy_{j'} \tag{17}$$

$$= \sum_j y_jy_{j\oplus(i\ominus i')} = b_{i-i'}(\mathbf{y}), \tag{18}$$

which defines a symmetric Toeplitz matrix [16],[11]. Moreover

$$\min_{\|\mathbf{x}\|=1} b(\mathbf{x},\mathbf{y}) = \min_{\|\mathbf{x}\|=1}\langle \mathbf{x},\mathbf{B_y}\mathbf{x}\rangle = \lambda(\mathbf{B_y}) \tag{19}$$

is the smallest eigenvalue of $\mathbf{B_y}$ [15, (1.3)], which is non-negative and efficiently solvable, see Section 5.

4. SKETCH OF PROOF

Embedding $\mathbb{R}^n$ in $\mathbb{R}^{\tilde{n}}$ with $\tilde{n} = 2n-1$ by adding $n-1$ zeros to each vector, we can replace the addition modulus $\tilde{n}$ with the usual addition, since for all index sums in (18) it holds $j+k \leq 2n-2$ for $j,k\in N$. Moreover, using the support property $\tilde{\mathbf{x}} = (\mathbf{x},\mathbf{0})$ equation (19) defines the smallest eigenvalue of the $n\times n$ principal submatrix in $\tilde{\mathbf{B}}_{\tilde{\mathbf{y}}}$ given by $[\mathbf{B}_{\tilde{\mathbf{y}}}]_{ii'} = b_{i-i'} = b_k$ for $i,i'\in N$. The first row is then given for $k\in N$ as

$$b_k(\tilde{\mathbf{y}}) = \sum_{j=0}^{\tilde{n}-1}\tilde{y}_j\tilde{y}_{j\oplus k} = \sum_{j=0}^{n-1} y_j\tilde{y}_{j+k} = \sum_{j=0}^{n-1-k} y_jy_{j+k} = b_k(\mathbf{y}) \tag{20}$$

and $\mathbf{B}_{\tilde{\mathbf{y}}} = \mathbf{B_y}$. The aperiodic autocorrelation vector $\mathbf{b}(\mathbf{y})$ in (20), can be written as $\mathbf{b} = \mathbf{Y}\mathbf{y}$, where $\mathbf{Y}$ is an $n\times n$ skew-symmetric Toeplitz matrix with elements $[\mathbf{Y}]_{jk} = y_{j+k}$, which has due to the removed periodicity a triangular structure. The shift k moves the support out of J, Hence the number of non-zero coefficients per rows decreases. If we restrict the support $\mathbf{x},\mathbf{y}$ to I resp J, then we cut out a symmetric $s\times s$ Toeplitz matrix $\mathbf{B}_{\mathbf{y}}^I$ with coefficients $k\in I\ominus i_0 = I - i_0 = \{k_0,\dots,k_{s-1}\}$ Assume we can represent (20) for each I, J and $\hat{\mathbf{y}}$ by a s samples $\bar{k}_i$ of the autocorrelation of a vector $\bar{\mathbf{y}} = \bar{\mathbf{y}}(\hat{\mathbf{y}})\in\mathbb{R}^m$, i.e. for all $i\in\{0,\dots,s-1\}$ we set

$$b_i^I(\mathbf{y}) \coloneqq b_{k_i}(\mathbf{y}) = \sum_{j=0}^{m}\bar{y}_j\bar{y}_{j+\bar{k}_i} =: b_{\bar{k}_i}(\bar{\mathbf{y}}), \tag{21}$$

which generates the $m\times m$ Toeplitz matrix

$$\mathbf{B}_{\bar{\mathbf{y}}}^m = \begin{pmatrix} b_0(\bar{\mathbf{y}}) & \dots & b_{m-1}(\bar{\mathbf{y}})\\ \vdots & \ddots & \vdots\\ b_{m-1}(\bar{\mathbf{y}}) & \dots & b_0(\bar{\mathbf{y}})\end{pmatrix} \tag{22}$$

with symbol

$$b(\bar{\mathbf{y}},\omega) = 1 + \sum_{l=0}^{m-1} b_l(\bar{\mathbf{y}})\cos(l\omega) \quad , \quad \omega\in[0,2\pi). \tag{23}$$

Then the s coefficients in (21), define the principal submatrix $\mathbf{B}_{\mathbf{y}}^I$ of $\mathbf{B}_{\bar{\mathbf{y}}}^m$ and by CAUCHYS Interleacing Theorem, the eigenvalues are bounded from below by the smallest eigenvalue of $\mathbf{B}_{\bar{\mathbf{y}}}^m$, i.e.

$$\min_{\hat{\mathbf{y}}\in S^{f-1}}\lambda(\mathbf{B}_{\mathbf{P}_J^n\hat{\mathbf{y}}}^I) \geq \min_{\mathbf{x},\bar{\mathbf{y}}\in S^{m-1}}\langle\mathbf{x},\mathbf{B}_{\bar{\mathbf{y}}}^m\mathbf{x}\rangle = \min_{\bar{\mathbf{y}}}\lambda(\mathbf{B}_{\bar{\mathbf{y}}}^m), \tag{24}$$

where the right hand side is independent of I and J. Hence, $b(\bar{\mathbf{y}},\omega)$ with b_l given in (21) is by the SZEGÖ Theorem a *nonnegative cosine polynomial* of order $m-1$ [17, Thm.4], i.e. we have to consider the minimum over all Szegö type polynomial of order $m-1$. For each normalized $\bar{\mathbf{y}}\in\mathbb{R}^m$ we get $0\leq\min_\omega b(\bar{\mathbf{y}},\omega)$ and by Böttcher in [18, (10.2)] we have $\lambda(\mathbf{B}_{\bar{\mathbf{y}}}^m) > 0$. Then $\mathbf{B}_{\bar{\mathbf{y}}}^m$ is invertible and the determinant $|\det(\mathbf{B}_{\bar{\mathbf{y}}}^m)| > 0$. Using

$$\frac{1}{\lambda(\mathbf{B}_{\bar{\mathbf{y}}}^m)} = \left\|\mathbf{B}_{\bar{\mathbf{y}}}^m\right\|_2 \tag{25}$$

[18, p.59], we can estimate the smallest eigenvalue (singular value) by [18, Thm. 4.2] to the determinant as

$$\lambda(\mathbf{B}_{\bar{\mathbf{y}}}^m) \geq |\det(\mathbf{B}_{\bar{\mathbf{y}}}^m)|\frac{1}{\sqrt{m}(\sum_l|b_l(\bar{\mathbf{y}})|^2)^{(m-1)/2}}. \tag{26}$$

where the ℓ^2-norm of the sequence b_k can be upper bounded by

$$\sum_l|b_l(\bar{\mathbf{y}})|^2 \leq 1+2\sum_{l=1}^{m-1}|\sum_{j=0}^{m-1}\bar{y}_j\bar{y}_{j+l}|^2 \leq 1+2m-2\leq 3m, \tag{27}$$

which is independent of $\bar{\mathbf{y}}\in S^{m-1}$! Since the determinant is a continuous function in $\mathbf{y}$ over a compact set, the minimum is attained and is denoted by $c_m := \min_{\bar{\mathbf{y}}}|\det(\mathbf{B}_{\bar{\mathbf{y}}}^m)|$. Note, that c_m is a decreasing sequence, since we extend the minimum to a larger set by increasing m. Hence we get

$$\min_{\bar{\mathbf{y}}}\left(|\det(\mathbf{B}_{\bar{\mathbf{y}}}^m)|\frac{1}{\sqrt{m}(3m)^{(m-1)/2}}\right) = \frac{\sqrt{3}}{(3m)^{m/2}}c_m \tag{28}$$

This is a valid lower bound by (26) for the smallest eigenvalue of all possible $\mathbf{B}_{\bar{\mathbf{y}}}^m$. Hence we have shown

$$a_{\text{opt}} \geq \min_{\substack{\tilde{\mathbf{y}}\in\tilde{\Sigma}_f\\ \|\tilde{\mathbf{y}}\|=1}}\min_{\substack{\tilde{\mathbf{x}}\in\tilde{\Sigma}_s\\ \|\tilde{\mathbf{x}}\|=1}} b(\tilde{\mathbf{x}},\tilde{\mathbf{y}}) \geq \sqrt{3}(3m)^{-\frac{m}{2}}c_m =: a(m). \tag{29}$$

The upper-bound in (5) was shown by the authors in [2, (32)].

4.1. Combinatoric via Construction of an Algorithm

Let us set $S := \{0,\dots,s-1\}$ and $F := \{0,\dots,f-1\}$ and assume that $sf\leq n$. We will now show that then the dimension m of the Toeplitz matrices can chosen to be at most $m = sf$. To this end we define the vector $\bar{\mathbf{y}}\in\mathbb{R}^m$ by $\bar{\mathbf{y}}_j = \sum_{\alpha\in F}\delta_{\bar{j}_\bullet,j}\hat{y}_\alpha$ and we denote its support by the set $\bar{J} = \{\bar{j}_\alpha \mid \alpha\in F\}$. According to (21) we have to show that there exist f resp. s indices $\{\bar{j}_\alpha\}_{\alpha\in F}$ and $\{\bar{k}_i\}_{i\in S}$ such that for all $i\in S$ it holds:

$$\sum_{\alpha,\beta}\left[\sum_{j=0}^{n-1-\bar{k}_i}\delta_{j,\bar{j}_\bullet}\delta_{j+\bar{k}_i,\bar{j}_\beta} - \sum_{j=0}^{n-1-k_i}\delta_{j,j_\bullet}\delta_{j+k_i,j_\beta}\right]\hat{y}_\alpha\hat{y}_\beta \tag{30}$$

$$=: \langle\hat{\mathbf{y}},\mathbf{C}^{(i)}\hat{\mathbf{y}}\rangle = 0 \quad \text{for all} \quad \hat{\mathbf{y}}\in\mathbb{R}^f$$

Thus, all s matrices $\mathbf{C}^{(i)}\in\mathbb{R}^{f\times f}$ must identically zero. In other words, the dimension m must be large enough, such that the $f^2\cdot s$ equations:

$$\delta_{\bar{j}_\alpha+\bar{k}_i,\bar{j}_\beta} = \delta_{j_\alpha+k_i,j_\beta} \quad \text{for all} \quad i\in S \tag{31}$$

can be fulfilled for some $\{\bar{j}_\alpha\}_{\alpha\in F}, \{\bar{k}_i\}_{i\in S} \subset \{0,\dots,m-1\}$.
In the Algorithm 1 below we show an inductive strategy to construct such index sets. The algorithm starts with dimension $m = f$ and

---

Algorithm 1 "Inserting Zeros"

Set $\bar{j}_\alpha = \alpha$ for all $\alpha \in F \quad \Rightarrow \quad m = f$
Set $\bar{k}_i = i$ for all $i \in S$
for $i = 1$ to $s - 1$ **do**
  for $\alpha = 0$ to $f - 1$ **do**
    $\Omega \Leftarrow \{j_\alpha + k_i\} \cap \{j_{\beta_\bullet + \alpha}, \dots, j_{f-1}\}$, note that $|\Omega| \le 1$
    if $|\Omega| = 1$ **then**
      $\{j_{\beta_\bullet + 1}\} := \Omega$
      $\bar{k}_{i'} \Leftarrow \bar{k}_{i'} + \beta_{\alpha+1} - 1$ for $i \le i' < s$
    else
      $\bar{j}_{\alpha'} \Leftarrow \bar{j}_{\alpha'} + 1$ for $\alpha < \alpha' < f$
    end if
    [verification, see **(*)**]
  end for
end for

---

successively increases the dimension $m \Rightarrow m + 1$. It contains two nested loops $i = 1 \dots s - 1$ and $\alpha = 0 \dots f - 1$ such that the total number of added dimensions is $(s - 1)f$. Hence, the algorithms finishes with at most $m = f + (s - 1)f = sf$. Furthermore, it holds $\bar{k}_{s-1} \le s - 1 + (s - 1)(f - 2) \le sf - 1$.
Let us emphasize some special sparse models:
If I, J are arithmetic progressions with same distance, then the algorithm only enters the case $|\Omega| = 1$ and generates $\bar{k}_i = i$ and hence we have $m = f$, which is the smallest possible embedding. Moreover, the inequality (24) becomes an equality such that the smallest eigenvalue of all symmetric positive Toeplitz matrices corresponds to the optimal lower bound of the convolution.
If I, J are maximal separated, i.e. for each $k \in K$ we have $J + k \cap J = \emptyset$, then all Toeplitz matrices equals the Identity. Hence $a = \lambda = 1$. Moreover, we have even equality in (5), see [12].

(*) Finally, we verify now that after step (i, α) the corresponding requirement (31) already achieved in steps $i' < i$ are still satisfied for all α' and β'. However, in this step only $\bar{j}_{\alpha'}$ for $\alpha < \alpha' \le f - 1$ are redefined. Not changing the value of $\delta_{j_{\bullet'} + k_{i'}, j_{\beta'}}$ means that the following cases for each fixed i' and α':

$$|\Omega| = 1 \quad \Leftrightarrow \quad \bar{j}_{\alpha'} + \bar{k}_{i'} = \bar{j}_{\beta'} \tag{32}$$

$$|\Omega| = 0 \quad \Leftrightarrow \quad \forall \beta' : \bar{j}_{\alpha'} + \bar{k}_{i'} \neq \bar{j}_{\beta'} \tag{33}$$

should be remain unchanged. If the algorithm reassigns $\bar{j}_{\alpha'} \Rightarrow \bar{j}_{\alpha'} + 1$ we have to satisfy:

$$|\Omega| = 1 \quad \Leftrightarrow \quad \bar{j}_{\alpha'} + 1 + \bar{k}_{i'} = \bar{j}_{\beta'} + 1 \tag{34}$$

since $\beta' > \alpha$. This is sufficient due to $\beta' \le \alpha < \alpha'$, J is ordered and $\bar{k}_{i'} \ge 0$. On the other hand:

$$|\Omega| = 0 \quad \Leftrightarrow \quad \forall \beta' : \bar{j}_{\alpha'} + 1 + \bar{k}_{i'} \neq \begin{cases} \bar{j}_{\beta'} + 1 & \beta' > \alpha \\ \bar{j}_{\beta'} & \beta' \le \alpha \end{cases} \tag{35}$$

If $\beta' > \alpha$ then from $\bar{j}_{\alpha'} + \bar{k}_{i'} \neq \bar{j}_{\beta'}$ it follows also $\bar{j}_{\alpha'} + 1 + \bar{k}_{i'} \neq \bar{j}_{\beta'} + 1$. If $\beta' \le \alpha$ then $\bar{j}_{\alpha'} + 1 + \bar{k}_{i'} > \bar{j}_{\beta'}$ since $\bar{k}_{i'} \ge 0$. Hence, also here the step (i, α) possibly changing $\bar{j}_{\alpha'} \Rightarrow \bar{j}_{\alpha'} + 1$ is consistent with previous steps.

5. ALGORITHMIC IMPLEMENTATION

The Problem in (19) can be approximated by discretization of $\mathbf{y}$ in $D = \{0, \sqrt{1/d}, \dots, \sqrt{d/d}\}$ with $d \in \mathbb{N}$. Hence D^m is a m-dimensional uniform grid of the cube. For each fixed $\mathbf{y}_d \in D^m$ with $\|\mathbf{y}_d\| = 1$ we get $\mathbf{B}_{\mathbf{y}_d}$ and obtain the approximate solution

$$a_m^d = \min_{\mathbf{y}_d \in D^m, \|\mathbf{y}_d\| = 1} \lambda_1(\mathbf{B}_{\mathbf{y}_d}), \tag{36}$$

which is an $(1 - 1/d)$-approximation solution to a, [15],i.e.

$$a_m \ge a_m^{\text{low},d} := a_m^d - \frac{m}{d}. \tag{37}$$

The price, is the size of the cube grid: the number of possible grid points $\mathbf{y}_d$ are of the order $|D|^s = (d + 1)^m < m^m$ and hence sub-exponential. We could establish in Fig. 1 with MATLAB global lower bounds, drawn as doted green lines, for a_m. For $m > 6$ the computational time was to large to establish a global lower bound.

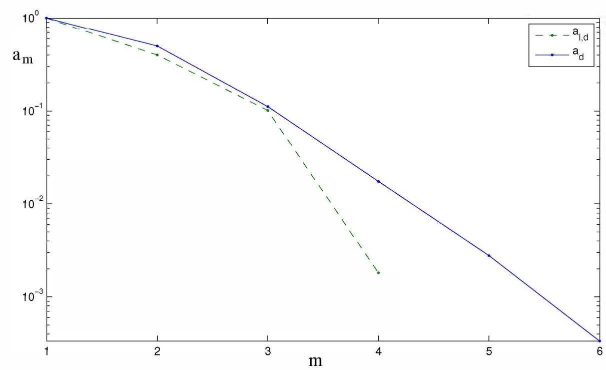

Fig. 1: Approximation results of the lower bound a_m.

6. CONCLUSIONS

We could show a non-zero ℓ^2-norm lower bound for highly sparse circular convolutions and establish a sub-exponential scaling depending only on the input sparsity. This is a new and surprising results, which offers new insight in digital signal processing.

7. ACKNOWLEDGEMENTS

We would like to thank Holger Boche and Gisbert Janssen for helpful discussions. A special thank goes to Götz Pfander for pointing out the zero cases of the convolution. This work was partly supported by the Deutsche Forschungsgemeinschaft (DFG) grant JU 2795/1-1&2.

8. REFERENCES

[1] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[2] P. Walk and P. Jung, "Compressed sensing on the image of bilinear maps," in *Information Theory Proceedings, 2012 IEEE International Symposium on*, July 2012, pp. 1291 – 1295.

[3] C. Hegde and R. Baraniuk, "Signal recovery on incoherent manifolds," Submitted.

[4] R. Baraniuk and M. Wakin, "Random projections of smooth manifolds," *Foundations of Computational Mathematics*, vol. 9, pp. 51–77, 2009.

[5] R. Baraniuk, V. Cevher, and M. Wakin, "Low-dimensional models for dimensionality reduction and signal recovery: A geometric perspective," *Proceedings of the IEEE*, vol. 98, pp. 959–971, 2010.

[6] Franck Barthe, "Optimal young's inequality and its converse:a simple proof," *Geom. Funct. Anal.*, vol. 8, pp. 234–242, 1998.

[7] J. Romberg and Ramesh Neelamani, "Sparse Channel Separation using Random Probes," *Inverse Problems*, vol. 26, 2010.

[8] T. Tao, "An uncertainty principle for cyclic groups of prime order," *Math. Research Letters.*, vol. 12, no. 1, pp. 121–127, 2005.

[9] K. Abed-Meraim, Wanzhi Qiu, and Yingbo Hua, "Blind system identification," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1310 –1322, aug 1997.

[10] J. Haupt and W. Bajwa, "Toeplitz compressed sensing matrices with applications to sparse channel estimation," *Information Theory, IEEE ...*, 2010.

[11] Philip J. Davis, *Circulant matrices*, Wiley, 1979.

[12] P. Walk and P. Jung, "Approximation of Löwdin Orthogonalization to a Spectrally Efficient Orthogonal Overlapping PPM Design for UWB Impulse Radio," *EURASIP Journal on Applied Signal Processing*, vol. 92, pp. 649–666, 2012.

[13] D. Donoho and Philip B. Stark, "Uncertainty principles and signal recovery," *SIAM J. Appl. Math.*, vol. 49, pp. 906–931, 1989.

[14] P. Stevenhagen and H. W. Lenstra, "Chebotarëv and his density theorem," *The Mathematical Intelligencer*, vol. 18, pp. 26–37, 1996.

[15] Chen Ling, Jiawang Nie, Liqun Qi, and Yinyu Ye, "Biquadratic optimization over unit spheres and semidefinite programming relaxations," *SIAM J. Optim.*, vol. 20, pp. 1286–1310, 2009.

[16] Robert M. Gray, *Toeplitz And Circulant Matrices*, vol. 2, Now Publishers Inc, 2006.

[17] Dimitar K. Dimitrov, *Approximation Theory: A volume dedicated to Blagovest Sendov*, Marin Drinov Academic Publishing House, Sofia, 2004.

[18] Albrecht Böttcher and Sergei M. Grudsky, *Spectral Properties of Banded Toeplitz Matrices*, SIAM, 2005.