ELECTRONIC RESEARCH ANNOUNCEMENTS
OF THE AMERICAN MATHEMATICAL SOCIETY
Volume 2, Number 1, August 1996

COMPRESSION AND RESTORATION OF SQUARE INTEGRABLE FUNCTIONS

RAFAIL KRICHEVSKII AND VLADIMIR POTAPOV

(Communicated by Ingrid Daubechies)

ABSTRACT. We consider classes of smooth functions on $[0,1]$ with mean square norm. We present a wavelet-based method for obtaining approximate pointwise reconstruction of every function with nearly minimal cost without substantially increasing the amount of data stored. In more detail: each function f of a class is supplied with a binary code of minimal (up to a constant factor) length, where the minimal length equals the ε-entropy of the class, $\varepsilon > 0$. Given that code of f we can calculate f, ε-precisely in L_2, at any specific $N, N \geq 1$, points of $[0,1]$ consuming $O(1+\log^*((1/\varepsilon)^{(1/\alpha)}/N))$ operations per point. If the quantity N of points is a constant, then we consume $O(\log^* 1/\varepsilon)$ operations per point. If N goes up to the ε-entropy, then the per-point time consumption goes down to a constant, which is less than the corresponding constant in the fast algorithm of Mallat [11]. Since the iterated logarithm $\log^*$ is a very slowly increasing function, we can say that our calculation method is nearly optimal.

1. INTRODUCTION

Mathematical models of signal compression have attracted attention for a long time. Many of those models have the following form. There is a set P with a norm $||\ \ ||$. A set M is an ε-net for P, $\varepsilon > 0$, if for any $x \in P$ there is an $m \in M$ such that $||x - m|| \leq \varepsilon$. The minimal log-cardinality of ε-nets is the metric ε-entropy $H_\varepsilon(P)$ of P, with log here and in the sequel binary. Points of M are encoded by different binary numbers. The code of m, $||x-m|| \leq \varepsilon$, is a compressed representation of x, with x recoverable from m with ε-precision. One wishes to achieve maximal compression, i.e., to make the length of the codes as close to $H_\varepsilon(P)$ as possible, keeping the cost of the encoding ($x \to m$) and the decoding ($m \to x$) reasonably small.

We shall restrict our consideration to the L_2 norm and the classes P_α of functions with square integrable α-th derivatives, $\alpha > 0$. P_α is defined as the set of all functions f with domain $[0,1]$ such that $\int_0^1 f(t)\,dt = 0$, $||f^{(\alpha)}|| \leq 1$, $f^{(\alpha)}$ is the α-th derivative in the sense of Hermann Weyl, i.e.,

$$f^{(\alpha)}(t) = \sum_{n=-\infty}^{\infty} c_n (2\pi i n)^\alpha e^{2\pi i n t},$$

Received by the editors October 20, 1995, and, in revised form, May 6, 1996.
1991 *Mathematics Subject Classification.* Primary 94A11; Secondary 42C10.
Key words and phrases. Wavelets, Haar functions, compression, computational complexity.

and c_n are the Fourier coefficients of f.

There are many ways of compressing a function $f \in P_\alpha$ into a bit stream with order $O(H_\varepsilon(P_\alpha))$ and $O(\varepsilon)$ distortion in L_2 (examples: transform coding in the Fourier Domain, in the Legendre Polynomial Domain, in the Wavelet Domain, differential coding of spline coefficients, etc.). The goal of the present paper is to find a method which allows the fastest decompression. By decompression we mean the ε-precise calculation of f at one or several points of $[0, 1]$ from its code. The code is kept intact allowing us to calculate repeatedly.

First of all we briefly survey methods of fast computation. We count operations in flops, an ideal unit of adds/multiples. The fast algorithm of Pan, Reif and Tate [12] enables one to compute ε-precisely the values of a J-degree polynomial at N points spending $O((N + \log\frac{1}{\varepsilon})(\log\log\frac{1}{\varepsilon})^2 + J \cdot \min(\log J, \log\frac{1}{\varepsilon}))$ flops. Their algorithm is a modification of Rokhlin's approach [13].

To compute the values of a wavelet polynomial (a linear combination of wavelets of levels below a number $\log J$) at N points one uses the fast algorithm of Mallat [11]. If $N = J$, then Mallat's algorithm consumes $O(N)$ flops. Suppose $N < J$. Partition $[0, 1]$ into N subintervals of equal length and pick up a point in each of them. Then Mallat's algorithm may consume $O(N(\log J - \log N))$ flops. The computation at a single point requires $O(J)$ flops for polynomials or $O(\log J)$ flops for wavelet polynomials.

Many papers are devoted to ε-entropy estimations. Such estimations produce compression methods.

The classes P_α were introduced by Kolmogorov and Tikhomirov in [8]. They have bounded ε-entropy as

$$C \cdot (1/\varepsilon)^{(1/\alpha)} \leq H_\varepsilon(P_\alpha) \leq C \cdot (1/\varepsilon)^{(1/\alpha)}.$$

Here and elsewhere the letter C stands for different positive constants, for simplicity sake. Kolmogorov and Tikhomirov first passed from L_2 to l_2 where P_α became an ellipsoid. Retaining the Fourier coefficients below the level $J = \lceil(\frac{1}{\varepsilon})^{1/\alpha})\rceil$ they ε-approximated the ellipsoid by a finite-dimensional one, in which they constructed a cubic lattice. It yielded the upper bound; the lower bound was obtained by the volume argument.

Besov spaces in L_p-norm are discussed in [5]. To ε-represent a function from such a space the authors used a linear combination of J wavelets, $\log J = O(\log\frac{1}{\varepsilon})$. We think that it yields some bounds on ε-entropy which may be compared with the bounds in [14].

The size of the paper does not allow us to refer to all the literature on the data compression. We mention only several papers.

Our main result is the following. We supply every function $f \in P_\alpha$ with a binary word of length $O((\frac{1}{\varepsilon})^{1/\alpha})$ bits, which is minimal up to a constant factor.

When the values of f at N points are needed, we can compute them with running time $O(1 + \log^*((1/\varepsilon)^{(1/\alpha)}/N))$ flops per point. This running time decreases from $O(\log^* 1/\varepsilon)$ to a constant, as N increases from a constant to the ε-entropy of the class. Computing f, we employ a black box subroutine of a smooth enough mother scaling function of I. Daubechies [2, 3]. If $\alpha \leq 1$, then we do not need such a subroutine and do not use multiplications. Thus, the computation of f by our method is faster than such a computation by means of trigonometric or even wavelet decompositions. Since the degrees of ε-approximating polynomials are $O((\frac{1}{\varepsilon})^{1/\alpha})$

for some $f \in P_\alpha$, which is very high, fast computation methods of polynomials do not give good results here.

We are going to develop a short and quickly computable representation of any function $f \in P_\alpha$, $\alpha > 0$. Namely, with ε-precision, f is a sum of $\log^* 1/\varepsilon$ addends. Each addend belongs to a ball lying in a finite-dimensional linear span of scaling functions. The total ε-entropy of the balls equals the ε-entropy of the class. To obtain such a representation, we start with a wavelet expansion of f. The projection of P_α on the linear span of wavelets of one and the same level is called a layer. $O(\log \frac{1}{\varepsilon})$ layers are required to ε-approximate P_α; consequently, it takes $O(\log \frac{1}{\varepsilon})$ flops to ε-compute f at a single point. It is less than $O(\frac{1}{\varepsilon})$ flops consumed by a Fourier expansion, but we want to compute still faster. The merging of all layers in a single meta-layer gives the ultimate speed of C flops per point. We pay for such a speed by the increase of the amount of data stored: the entropy of the meta-layer becomes $\log \frac{1}{\varepsilon}$-times bigger than $H_\varepsilon(P_\alpha)$. To keep the total entropy within $O(H_\varepsilon(P_\alpha))$, we merge (fine to coarse) first two layers into a meta-layer, then four, and so on in nearly explosive way $C^{C^{\cdots}}$, $C > 1$; see Lemma 5.

To show that such merging gives the required entropy and speed we estimate in Section 4 the radius of a ball containing a layer. We find in Section 5 how that radius responds to merging. We build up in Section 2 an algorithm of fast coding and decoding of balls within given precision (Salary List algorithm).

A function belonging to a meta-layer is a linear combination of scaling functions of the same level. On the other hand, it belongs to a ball. The Salary List algorithm gives a short code to such a combination allowing us to find the coefficients quickly. We use finitely many coefficients of every level to compute f at any single point $x \in [0, 1]$. It yields the Theorem (Section 5). Combining the Theorem with the fast method of Mallat allows us to compute f at any N points.

We believe that the results may be generalized to other norms and classes.

2. Salary List: fast enumeration of integer-valued vectors in simplexes and balls

A p-dimensional simplex with side N is a set of vectors with nonnegative coordinates whose sum does not exceed N. The coordinates may be thought of as salaries of employees of a firm, N being the total salary. Enumerating each integer-valued vector by the concatenation of $\lceil N+1 \rceil$-digit numbers representing the coordinates makes the length of the code equal to $p \cdot \lceil N+1 \rceil$, which is not optimal, since the volume of the simplex is $N^p/p!$. The length of the code is lowerbounded by $p \cdot \log N/p$. We present a fast algorithm attaining that bound.

Lemma 1 (Salary List code of a simplex). *Every integer-valued vector of a p-dimensional simplex with side N can be given a binary word of length $O(p \cdot (1 + \log N/p))$. Each coordinate of a vector can be found via its word in $O(\log N \cdot \log\log p + \log p)$ operations over bits.*

Proof. We define first a preliminary code of a simplex. The dimension is supposed to equal $2^r, r \geq 0$; the sum of the coordinates equals N; $l(x)$ is the length of the binary notation of a nonnegative integer x. Obviously,

$$l(x) \leq \log(x+1) + 1. \tag{1}$$

We encode an integer-valued vector $x = (x_1, \dots, x_p)$ by an r-level labelled tree. The label of its root (first level) is the $l(N)$-digit number $N_0 = x_1 + \cdots + x_{p/2}$. The left son is labelled by the $l(N_0)$-digit number $N_{00} = x_1 + \cdots + x_{p/4}$, and the right one by the $l(N_1)$-digit number $N_{10} = x_{p/2+1} + \cdots + x_{3p/4}$. Proceeding in this way we build a labelled tree. The total length of labels on its i-th level, by (1) and the arithmetic mean–geometric mean inequality, does not exceed

$$2^i + 2^{i-1} \cdot \log(1 + N/2^{i-1}).$$

Summing it over $1 \le i \le r$ and using the formulas for the sums of geometric progressions and their derivatives, we upperbound the total length of all labels as $O(p \cdot \log \frac{4(p+2N)}{p})$. A vector x can be restored through its tree. The tree itself can be restored through the concatenation (without commas) of all its labels. That concatenation is the preliminary code.

We subdivide the coordinates into $p/\log p$ equal-dimensional subvectors. The conclusive Salary List code consists, first, of the number $x_1 + \cdots + x_p$, second, the concatenation of the sums of the coordinates of those subvectors, and third, the preliminary codes of the subvectors. Such a subdividing allows us to accelerate the decoding by slightly increasing the length of the code.

Lemma 2 (Salary List code of a ball). *A p-dimensional ball of radius R can be ε-precisely encoded by binary words of length $O(p \cdot (1 + \log R/\varepsilon))$. The decoding time per coordinate is $O(\log Rp/\varepsilon \cdot \log\log p)$ operations over bits.*

Proof. Develop in the ball a cubic $\varepsilon/\sqrt{p}$-lattice. The enumeration of its nodes is reduced to the enumeration of integer-valued vectors of a simplex by the Cauchy-Schwarz inequality

$$|x_1| + \cdots + |x_p| \le \sqrt{x_1^2 + \cdots + x_p^2} \cdot \sqrt{p}.$$

3. Wavelet representation of the classes P_α

Following [2, 3] we introduce some basic notions of the wavelet theory. There is a function ψ called the mother wavelet such that its dilations and translations $\psi_{j,k}$

$$\psi_{j,k}(x) = \sqrt{2^j}\psi(2^j x - k), \qquad k, j = 0, \pm 1, \dots .$$

constitute an orthonormal basis of $L_2(-\infty, +\infty)$. We say that $\psi_{j,k}$ is a j-level wavelet function. The linear span of j-level wavelet functions is denoted by W_j; the projection onto W_j is denoted by Q_j. The projection $Q_j P_\alpha$ of the class P_α on W_j is called the j-th layer of P_α generated by the mother ψ, $j = 0, \pm 1$. The layers are orthogonal to each other.

For any mother $\psi(x)$ there is a function $\phi(x)$ called the mother scaling function. Its dilations and translations are denoted by $\phi_{j,k}$. We say that $\phi_{j,k}$ is a j-level scaling function; the linear span of j-level scaling functions is denoted by V_j. The equality $V_j \oplus W_j = V_{j+1}$ holds.

Hence, we have

Remark 1. The j-th level layer belongs to any span V_k, $k \ge j + 1$.

I. Daubechies has shown in [2, 3] that for any $r \ge 1$ there is an r-times differentiable mother scaling function $\phi(x)$. The support of ϕ is finite. The corresponding mother wavelet $\psi(x)$ has also a finite support. It is also r-times differentiable. Moreover, its first r moments vanish.

We have

Remark 2. Any point $x \in [0,1]$ belongs to supports of a finite number of wavelets and scaling functions of each level. For any $j = 0, \pm 1, \dots$ there are at most $2^j + C$ functions $\psi_{j,k}$ whose supports have nonempty intersections with $[0,1]$. Hence, the dimension of the j-th level layer is not greater than $2^j + C$, $j = 0, 1, \dots$.

The well-known Haar basis is a special case of wavelets. Indicators of binary rational intervals are scaling functions of the Haar basis.

Lemma 3. *For any class P_α of smooth functions there is a mother wavelet ψ such that the j-th level layer generated by ψ belongs to a ball whose radius is not greater than $C \cdot 2^{-j\alpha}$ and whose dimension is not greater than $C + 2^j$, where C depends on ψ only. For $j = 0, 1, 2, \dots$,*

$$\sum_{k=1}^{2^j+C} c_{j,k}^2 \le C \cdot 2^{-2j\alpha},$$

$$c_{j,k} = (f, \psi_{j,k}).$$

Proof. Consider a mother wavelet $\psi(x)$ having a finite support contained in an interval $(0, S)$ and n vanishing moments, where $n = \lfloor \alpha \rfloor$ if α is not an integer, and $n = \alpha - 1$ if α is an integer.

Taking the Taylor expansion for $f(\frac{y+k}{2^j})$ with integral remainder term at the point $\frac{k}{2^j}$, we obtain:

(2)
$$c_{j,k} = \frac{2^{-j/2}}{\Gamma(n)} \int_0^S \left(\psi(y) \int_{k/2^j}^{(y+k)/2^j} \left(f^{(n)}(\tau) - f^{(n)}\Big(\frac{k}{2^j}\Big) \right) \Big(\frac{y+k}{2^j} - \tau \Big)^{n-1} d\tau \right) dy.$$

Since $f \in P_\alpha$, we have

$$f^{(n)}(\tau) - f^{(n)}\Big(\frac{k}{2^j}\Big) = \frac{1}{\Gamma(\alpha - n)} \int_{k/2^j}^{\tau} f^{(\alpha)}(t)(\tau - t)^{\alpha - n - 1}\, dt. \tag{3}$$

Substitute (3) into (2), change the order of integration and use the definition of Γ-function:

$$c_{j,k} = \frac{1}{\Gamma(\alpha)} \int_{k/2^j}^{(S+k)/2^j} \psi_{j,k} \int_{k/2^j}^{x} f^{(\alpha)}(t)(x-t)^{\alpha-1}\, dt\, dx. \tag{4}$$

Square (4) and apply the Cauchy-Schwarz inequality. It yields

$$c_{j,k}^2 = \frac{2^j (\sup|\psi|)^2}{(2\alpha+1)\Gamma^2(\alpha+1)} \Big(\frac{S}{2^j} \Big)^{2\alpha+1} \int_{k/2^j}^{(S+k)/2^j} (f^{(\alpha)}(t))^2\, dt. \tag{5}$$

Since $f \in P^\alpha$, we have the statement of the lemma.

Corollary of Lemma 3 . *Let $\alpha > 0$, P_α be a class of smooth functions, $\varepsilon > 0$, and $J = \frac{1}{\alpha} \log \frac{1}{\varepsilon} + C$. Then there is a mother wavelet such that P_α is ε-approximated by the direct sum of layers $Q_j P_\alpha$, $-\infty < j \le J$.*

4. Layer merging

Let $a < b$ be integers. The direct sum over $a < j \le b$ of the j-th level layers of P_α is called the $(a,b]$-meta-layer and denoted by $\ell(a,b]$.

By Remark 1, $\ell(a,b] \subset V_{b+1}$. So, each element of $\ell(a,b]$ is represented by its coordinates over the basis of $b+1$-level scaling functions.

Lemma 4. *For a class P_α there is a mother wavelet such that for $a < b$ and $\varepsilon > 0$ the elements of the meta-layer $\ell(a, b]$ can be encoded by binary words of length $O(2^b \log(2^{-a\alpha}/\varepsilon))$. Those elements can be restored from their codes with ε-precision in $O(\log(2^{b-a\alpha}/\varepsilon)\log b)$ operations over bits.*

Proof. By Lemma 3, $\ell(a, b]$ belongs to a ball in V_{b+1} of radius

$$\sqrt{C \sum_{j=a-1}^{b} 2^{-2j\alpha}},$$

which is not greater than $C \cdot 2^{-a\alpha}$. The dimension of the ball is not greater than $2^{b+1} + C$ by Remark 2. Now the lemma follows from Lemma 2.

5. Approximating P_α by direct sums of meta-layers

Lemma 5 (Arithmetical). *Let $\xi_0, \xi_1, \dots$ be a sequence defined by*

$$\xi_0 = 0, \dots, \xi_{i+1} = 2^{\xi_i - i+1}, i = 1, 2, \dots .$$

Then the sequence is monotonic, and for any $R > 0$ there is a number k such that $\xi_k \geq R$ and $k \leq C \cdot \log^* R$.

Lemma 6. *For any class P_α and $\varepsilon > 0$ there is a mother wavelet such that P_α is ε-approximated by the direct sum of $k+1 = O(\log^* 1/\varepsilon)$ meta-layers. The ε-entropy of the sum equals $O(H_\varepsilon(P_\alpha))$. Each function of the sum can be ε-restored through its Salary List code of length $O(H_\varepsilon(P_\alpha))$ in C flops per coordinate (in a basis of scaling functions).*

Proof. Let

$$J = 1/\alpha \cdot \log 1/\varepsilon + C. \tag{6}$$

Take the sequence defined by Lemma 5; let k be the minimal number for which $\xi_k \geq J + 1$. According to that lemma,

$$k \leq C \cdot \log^* 1/\varepsilon. \tag{7}$$

Define a monotonic sequence

$$m_l = J - \xi_{k-l} \tag{8}$$

and a sequence of meta-layers $\ell(-\infty, m_0], \ell(m_0, m_1], \dots, \ell(m_{k-1}, m_k]$. By the Corollary of Lemma 3, P_α is ε-approximated by the direct sum of those meta-layers. For each f from the class there is a function g in the sum such that $||f - g|| < \varepsilon$. The function g is the sum of its projections on the meta-layers:

$$g = g_0 + \dots + g_k, \quad g_l \in \ell(m_{l-1}, m_l]. \tag{9}$$

Let

$$\varepsilon_l = \frac{\varepsilon \cdot 2^{m_l}}{2^{J+2}}, \qquad l = 0, 1, \dots . \tag{10}$$

Each element of the direct sum is encoded by the concatenation of codes of its projections. We encode the l-th component (9) with ε_l-precision. Then the total decoding error is $\sqrt{\sum_{l=0}^{k} \varepsilon_l^k}$, which by (8) and (10) is not greater than ε.

The component g_0 belongs to a finite-dimensional ball of radius 1, so the length of its ε_0-code is $O(\log 1/\varepsilon)$ by Lemma 2. For the total length of the code of g, by Lemma 4, (6), (10), and the monotonicity of m_l, we have:

$$|\,\mathrm{code}\ g| = C \cdot \sum_{l=1}^{k} 2^{m_l} \cdot \log \frac{2^{J(1+\alpha)+2}}{2^{m_{l-1}\cdot(1+\alpha)}} + |\,\mathrm{code}\ g_0| \tag{11}$$

Substitute (8) into (11):

$$|\,\mathrm{code}\ g| = C \cdot (\alpha+1) \cdot 2^J \cdot \sum_{l=1}^{k} \xi_{l+1}/2^{\xi_l}. \tag{12}$$

Use in (12) the definition of the sequence ξ_l :

$$|\,\mathrm{code}\ g| = C \cdot (\alpha+1) \cdot 2^J \sum_{l=1}^{k} 2^{1-l}. \tag{13}$$

From (6) and (13) we have:

$$|\,\mathrm{code}\ g| = C \cdot H_\varepsilon(P_\alpha).$$

To decode a coordinate of a projection we spend, by Lemma 4, at most

$$C \cdot (\log 2^J/\varepsilon)(\log J) = C \cdot \log 1/\varepsilon \cdot \log\log 1/\varepsilon$$

operations over bits, which is equivalent to C flops by the Shönhage-Strassen algorithm. The code of g falls into the codes of its projections. Their lengths are known. The lemma is proved.

Theorem . *Any function $f \in P_\alpha$ can be given a binary code of minimal, up to a constant factor, length $O((\frac{1}{\varepsilon})^{1/\alpha})$ through which it can be restored at any given individual point with ε-precision in nearly minimal time $\log^* \frac{1}{\varepsilon}$ flops. A precomputed table of a smooth mother scaling function is entered $\log^* \frac{1}{\varepsilon}$ times. For $\alpha \leq 1$ only adds are performed, and the table is not needed.*

Proof. By Remark 2, any point $x \in [0,1]$ belongs to a finite number of scaling functions of each level. Now the Theorem follows Lemma 6.

Corollary of the Theorem . *Any function $f \in P_\alpha$ can be given a binary code of minimal (up to a constant factor) length $O((\frac{1}{\varepsilon})^{1/\alpha})$ through which it can be restored at any given N points with ε-precision. The time consumption is $O(1 + \log^*((1/\varepsilon)^{(1/\alpha)}/N))$ flops per point, which is a constant C_1 for $N = O((1/\varepsilon)^{1/\alpha})$. The constant C_1 is less than the corresponding constant in Mallat's algorithm.*

Proof. The corollary is obtained by a straightforward combination of our Theorem with Mallat's algorithm [11]. First we apply Mallat's method taking into account expansion (9). As soon as the support of a scaling function contains only one given point, we start applying our Theorem.

6. Acknowledgement

The authors are very much indebted to the anonymous referee for the thorough reading of the manuscript and many deep remarks.

References

[1] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, *Image coding using wavelet transforms*, IEEE Trans. Image Processing, **1** (1992), 205–220.

[2] I. Daubechies, *Orthonormal bases of compactly supported wavelets*, Comm. Pure Appl. Math., **41** (1988), 909–996. MR **90m:**42039

[3] ———, *Ten lectures on wavelets*, CBMS Lecture Notes no. 61, SIAM, Philadelphia, PA, 1990. MR **93e:**42045

[4] R. A. DeVore, B. Jawerth, and B. J. Lucier, *Surface compression*, Computer-Aided Geometric Design, **9** (1992), 219–239. MR **93i:**65029

[5] ———, *Image compression through wavelet transform coding*, IEEE Trans. Info. Theory **38** (1992), 719–746. CMP 92:11

[6] R. A. DeVore, B. Jawerth, and V. Popov, *Compression of wavelet decompositions*, Amer. J. Math. **114** (1992), 737–785. MR **94a:**42045

[7] D. L. Donoho, *Unconditional bases are optimal bases for data compression and statistical estimation*, Applied and Computational Harmonic Analysis **1** (1993), 100–115. MR **94j:**94011

[8] A. N. Kolmogorov and V. M. Tikhomirov, *ε-entropy and ε-capacity of sets in function spaces*, Uspekhi Mat. Nauk **14**(1959), no. 2, 3–86. (Russian) MR **22:**2890

[9] R. Krichevsky, *Universal compression and retrieval*, Kluver Academic Publishers, 1995. MR **95g:**94006

[10] B. J. Lucier, *Wavelets and image compression*, Mathematical Methods in CAGD and Image Processing (T. Lyche and L.L. Schumaker, eds.), Academic Press, Boston, MA, 1992, pp. 1–10. MR **93d:**65022

[11] S. Mallat, *Multiresolution approximation and wavelet orthonormal bases of L^2*, Trans. Amer. Math. Soc. **315** (1989), 69–87. MR **90e:**42046

[12] V. Pan, J. Reif, and S. Tate, *The power of combining the techniques of algebraic and numerical computing: improved approximate multipoint polynomial evaluation and improved multipole algorithms* (Proc. of 33 Annual Symp. on Found. of Comp. Sc.), IEEE Comp. Soc. Press, Pittsburgh, 1992.

[13] V. Rokhlin, *A fast algorithm for the discrete Laplace transformation*, J. of Complexity **4** (1988), 12–32. MR **89b:**65309

[14] H. Triebel, *Approximation numbers and entropy numbers of embeddings of fractional Besov-Sobolev spaces in Orlicz spaces*, Proc. London Math. Soc. **66** (1993), 589–618. MR **94g:**46042

Sobolev Mathematical Institute, Novosibirsk, Russia
E-mail address: `rekri@math.nsk.su`

Sobolev Mathematical Institute, Novosibirsk, Russia
E-mail address: `rekri@math.nsk.su`