Universität Potsdam

Volker Dicken, Peter Maaß

Wavelet-Galerkin methods for
ill-posed problems

Wavelet-Galerkin methods for ill-posed problems

V. DICKEN* and P. MAAß*

Abstract — Projection methods based on wavelet functions combine optimal convergence rates with algorithmic efficiency. The proofs in this paper utilize the approximation properties of wavelets and results from the general theory of regularization methods. Moreover, adaptive strategies can be incorporated still leading to optimal convergence rates for the resulting algorithms. The so-called wavelet-vaguelette decompositions enable the realization of especially fast algorithms for certain operators.

1. INTRODUCTION

A projection method for solving the inverse problem

$$Af = g \tag{1.1}$$

where $A : X \to Y$ denotes an operator between Hilbert spaces X and Y, and $g \in Y$ is the given data is defined by the sequences of subspaces $X_h \subset X_{h'} \subset X$, $h' < h$, and $Y_h \subset Y$ growing in size with the step width h. In a realistic setting we are only given noisy data g^ε with a bounded error $\|g - g^\varepsilon\|_Y < \varepsilon$. An approximate solution $f_h \in X_h$ is then computed from the requirement

$$\langle Af_h \mid y \rangle_Y = \langle g^\varepsilon \mid y \rangle_Y, \quad \forall y \in Y_h. \tag{1.2}$$

The basic question is to estimate the quality of the approximation; i. e., $\|f - f_h\|_X$, and to determine the optimal step width h such that this quantity becomes minimal.

When choosing the function spaces X_h and Y_h opposing requirements have to be met. On the one hand X_h and respectively Y_h should be adapted to the properties of the operator A. From this perspective one would optimally choose spaces consisting of eigenfunctions or singular functions of A. In this case f_h could be computed from a diagonal linear system leading to an orthogonal expansion of f_h. But in general one does not have an explicit description of the singular function system of A.

On the other hand computational efficiency demands fast algorithms for computing the functions in X_h (respectively, Y_h) and the scalar products on the right-hand side of (1.2). Moreover, local support of the functions in X_h is needed in order to employ adaptive strategies. Finally, the function spaces need to have good approximation properties if optimal convergence rates are sought. These requirements are fulfilled, for example, by hierarchically structured finite element spaces.

*Universität Potsdam, Am Neuen Palais 10, D–14469 Potsdam, Germany.
E-mails: dicken@rz.uni-potsdam.de; maass@rz.uni-potsdam.de

A suitable compromise between these different requirements are orthogonal (or biorthogonal) wavelet spaces. They are hierarchically structured, allow fast recursive computations, and lead to some extent to the diagonalization of the operator A. In this paper we investigate the advantages and disadvantages of wavelets when used as basis functions for projection methods. The results of this preliminary investigation show that optimal convergence rates are achieved under general conditions, but full numerical efficiency is at present possible only for a special class of operators, namely, for quasi-homogeneous operators.

We first summarize the relevant properties of wavelets in Section 2. Their approximation properties directly allow to estimate the convergence rate of wavelet-projection methods, and they show that these regularization methods have the optimal order of convergence, see Section 3. The notion of a wavelet-vaguelette decomposition is introduced in Section 4. These decomposition techniques and, in particular, their relation to singular value decompositions and convergence results for inverse problems have been recently investigated in a series of papers (see, e.g., [5]). In these papers the results on optimal convergence rates were proved in terms of Bayes estimators in a stochastical setting. The results of Section 4 show that similar results hold within the standard framework for regularization methods. Moreover, optimal convergence rates can also be achieved in connection with a simple adaptive strategy. Results of a numerical simulation are presented in Section 5.

2. WAVELETS

In this section we summarize the main properties of wavelets as far as they are needed in the following sections. For a more detailed introduction to the theory of wavelets the reader is referred to [2, 8].

Definition 2.1. A function $\psi \in L^2(\mathbf{R})$ is called an orthogonal wavelet if the set

$$\left\{ \psi_{mk}(x) = 2^{m/2}\psi(2^m x - k) \;\middle|\; m, k \in \mathbf{Z} \right\} \tag{2.1}$$

forms an orthonormal basis for $L^2(\mathbf{R})$.

Orthogonal wavelets have very pleasant theoretical and practical properties, but their construction requires to solve a system of nonlinear equations. Instead one can use biorthogonal wavelets. They exhibit the same algorithmic efficiency, and they are obtained by solving a system of linear equations. Moreover, it is possible to construct symmetric biorthogonal wavelets.

Definition 2.2. A pair of functions $\psi, \tilde{\psi} \in L^2(\mathbf{R})$ is called a pair of biorthogonal wavelets if the sets $\{\psi_{mk} \mid m, k \in \mathbf{Z}\}$ and $\{\tilde{\psi}_{mk} \mid m, k \in \mathbf{Z}\}$ form the Riesz basis for $L^2(\mathbf{R})$, and if any function $f \in L^2(\mathbf{R})$ has the representation

$$f = \sum_{m \in \mathbf{Z}} \sum_{k \in \mathbf{Z}} \langle f \mid \psi_{mk} \rangle_{L^2} \cdot \tilde{\psi}_{mk}.$$

For the ease of presentation we will state the results in this paper only for orthogonal wavelets in $L^2(\mathbf{R})$. The generalization to biorthogonal wavelets and multi-dimensional applications is obvious in most cases.

A wavelet basis is labeled by two indices, the shift-index k and the scale-index m. If we collect all basis functions on the same scale in the subspace

$$W_m = \overline{\text{span} \{\psi_{mk} \mid k \in \mathbf{Z}\}}$$

then we obtain the orthogonal multi-scale decomposition

$$L^2(\mathbf{R}) = \bigoplus_{m \in \mathbf{Z}} W_m.$$

This type of multi-resolution analysis was first introduced by Y. Meyer and S. Mallat (see, e.g., [9]). The subspaces W_m are called the scales of the decomposition, and the projection

$$Q_m f = \sum_{k \in \mathbf{Z}} \langle f \mid \psi_{mk} \rangle_{L^2} \cdot \psi_{mk}$$

is said to represent the details of f on the scale m.

This hierarchical ordering of the wavelet basis shows its affinity to finite elements and multi-grid methods; it is the key to construction of efficient adaptive algorithms.

Let us denote by

$$V_m = \bigoplus_{\ell < m} W_\ell$$

the subspace of $L^2(\mathbf{R})$ which contains all details of size m or coarser. We assume the existence of a related function $\varphi \in L^2(\mathbf{R})$, called a scaling function, with the property

$$V_0 = \overline{\text{span} \{\varphi_{0k}(x) = \varphi(x - k) \mid k \in \mathbf{Z}\}}.$$

Obviously, the subspaces V_m are spanned by $\{\varphi_{mk}(x) = 2^{m/2}\varphi(2^m x - k) \mid k \in \mathbf{Z}\}$. Since $\varphi \in V_0 \subset V_1$ and $\psi \in W_0 \subset V_1$, it follows that both the wavelet ψ and the scaling function φ satisfy the scaling equations

$$\varphi(x) = \sqrt{2} \sum_{k \in \mathbf{Z}} h_k \cdot \varphi(2x - k) \tag{2.2}$$

$$\psi(x) = \sqrt{2} \sum_{k \in \mathbf{Z}} g_k \cdot \varphi(2x - k) \tag{2.3}$$

with real scaling coefficients h_k, g_k. For example, the Haar function $\varphi(x) = \chi_{[0,1]}(x)$ and the Haar wavelet $\psi(x) = \chi_{[0,1/2]}(x) - \chi_{[1/2,1]}(x)$ satisfy the above conditions with

$$\varphi(x) = \varphi(2x) + \varphi(2x - 1)$$

$$\psi(x) = \varphi(2x) - \varphi(2x - 1).$$

The Haar wavelet is the only obvious compactly supported orthogonal wavelet. It was the accomplishment of I. Daubechies [2] to construct a family $\{\psi_N\}$ of orthogonal wavelets with $\text{supp}(\psi_N) = [0, 2N - 1]$ and linearly increasing regularity. Here ψ_1 equals the Haar wavelet, and ψ_2 is the solution of (2.2), (2.3) with coefficients

$$h_0 = \frac{1 + \sqrt{3}}{8}, \qquad h_1 = \frac{3 + \sqrt{3}}{8}, \qquad h_2 = \frac{3 - \sqrt{3}}{8}, \qquad h_3 = \frac{1 - \sqrt{3}}{8}.$$

The corresponding filter coefficients g_k in (2.3) are given by

$$g_k = (-1)^k h_{2N-1-k}, \quad k = 0, \ldots, 2N - 1.$$

For actual computations one is only interested in wavelets and scaling functions which obey such finite scaling equations. In this case the computation of the wavelet decomposition; i. e., the computation of the scalar products

$$d_k^m = \langle f \mid \psi_{mk} \rangle_{L^2} \tag{2.4}$$

is particularly easy and fast. Assume that we know the values of the scalar products

$$c_k^L = \langle f \mid \varphi_{Lk} \rangle_{L^2} \tag{2.5}$$

on some fine scale L. Then, employing the scaling equations (2.2), (2.3), we obtain recursively

$$c_k^m = \sum_{\ell \in \mathbf{Z}} h_{\ell-2k} c_k^{m+1} \tag{2.6}$$

$$d_k^m = \sum_{\ell \in \mathbf{Z}} g_{\ell-2k} c_k^{m+1}. \tag{2.7}$$

It remains to compute the values of $\{c_k^L\}$ on a fine scale L. In a realistic situation we are rather given discrete values f_k, $k \in \mathbf{Z}$, of f than the function itself. Not all scaling functions have the interpolating property like the Haar function. But they necessarily have a non-vanishing mean value [10], $\int \varphi(x)\,\mathrm{d}x \neq 0$, hence $2^L \varphi(2^L x)$ approximates the δ-distribution, and we simply put

$$f_k = c_k^L$$

on an appropriate scale L. For a given discrete sequence $\{f_k\}$ of length n and a wavelet with a finite scaling equation the complexity of computing the discrete wavelet decomposition is then of the order $O(n)$.

So far we have indicated the algorithmic advantages of wavelet bases such as the hierarchical structure and recursive computations. For the proofs of convergence results in the following sections the key ingredient is the approximation property of wavelets. Any wavelet ψ has a vanishing mean value, hence its Fourier transform has a zero at the origin. Therefore the Fourier transform of the wavelet basis

$$\hat{\psi}_{mk}(\omega) = 2^{-m/2} \mathrm{e}^{-2\pi\mathrm{i}\cdot 2^m k\omega}\, \hat{\psi}(2^{-m}\omega)$$

picks up different frequency bands as m varies. It is not surprising that the knowledge of the wavelet coefficients

$$d_k^m = \langle f \mid \psi_{mk} \rangle_{L^2} = \int \hat{f}(\omega) \cdot \hat{\psi}_{mk}(\omega)\,\mathrm{d}\omega$$

allows us to estimate the Sobolev regularity of f itself [3, 10]. The Sobolev space of the order s is defined as

$$H^s(\mathbf{R}) = \left\{ f \mid (1 + |\omega|^2)^{s/2} \hat{f}(\omega) \in L^2(\mathbf{R}) \right\}$$

with the norm

$$\|f\|_s := \int_{\mathbf{R}} \left| (1 + |\omega|^2)^{s/2} \hat{f}(\omega) \right|^2 \mathrm{d}\omega.$$

Theorem 2.1. *Given a wavelet $\psi \in H^r(\mathbf{R})$, then for every s, $0 \leq s < r$, there exist constants C_1, C_2, $0 < C_1, C_2 < \infty$, such that for every $f \in H^s(\mathbf{R})$ we have*

$$C_1 \|f\|_s \leq \left\{ \sum_{m \in \mathbf{Z}} (1 + 2^{2ms}) \left(\sum_{k \in \mathbf{Z}} |d_k^m|^2 \right) \right\}^{1/2} \leq C_2 \|f\|_s.$$

By duality this theorem remains valid for negative s as long as $|s| < r$. In full generality this result extends to Besov spaces $B_p^{s,q}$ [3].

We need a slightly different version of this theorem. By the definition of V_m and W_m, any function $f \in L^2(\mathbf{R})$ has the following decomposition

$$f = \sum_{m \geq 0} \sum_{k \in \mathbf{Z}} d_k^m \psi_{mk} + \sum_{k \in \mathbf{Z}} \langle f \mid \varphi_{0k} \rangle_{L^2} \cdot \varphi_{0k}. \tag{2.8}$$

Corollary 2.1. *Assume that* $\psi, \varphi \in H^r(\mathbf{R})$, *then for* $|s| < r$ *the Sobolev norm is equivalent to*

$$\|f\|_s^2 \sim \sum_{m \geq 0} 2^{2ms} \left(\sum_{k \in \mathbf{Z}} |d_k^m|^2 \right) + \sum_{k \in \mathbf{Z}} \left| \langle f \mid \varphi_{0k} \rangle_{L^2} \right|^2.$$

Now it is easy to obtain the approximation result for the truncated wavelet expansion. This will be useful in proving the convergence of wavelet-Galerkin methods.

Lemma 2.1. *Let a wavelet* $\psi \in H^r(\mathbf{R})$ *and a function* $f \in H^\nu(\mathbf{R})$, $0 \leq \nu < r$, *be given. Let* f_M, $M \geq 0$, *denote the truncated wavelet expansion of* f

$$f_M = \sum_{m=0}^M \sum_{k \in \mathbf{Z}} d_k^m \psi_{mk} + \sum_{k \in \mathbf{Z}} \langle f \mid \varphi_{0k} \rangle_{L^2} \cdot \varphi_{0k}.$$

Then for every real number $s \leq \nu$, $|s| < r$, *the truncation error is bounded*

$$\|f - f_M\|_s \leq C \cdot 2^{-M(\nu - s)} \|f\|_\nu$$

where the constant C *is independent of* f *and* s.

Proof. The above-stated wavelet decomposition and Theorem 2.1 give

$$
\begin{aligned}
\|f - f_M\|_s^2 &\leq C_2 \sum_{m > M} 2^{2ms} \left(\sum_{k \in \mathbf{Z}} |d_k^m|^2 \right) \\
&= C_2 \sum_{m > M} 2^{2m(s-\nu)} 2^{2m\nu} \left(\sum_{k \in \mathbf{Z}} |d_k^m|^2 \right) \\
&\leq (C_2/C_1) \cdot 2^{-2M(\nu - s)} \|f\|_\nu^2. \qquad \square
\end{aligned}
$$

If one interprets the factor 2^{-M} which equals the shift parameter on the scale M as a step size or discretization parameter h, then Lemma 2.1 states that wavelets have the same approximation properties as, for example, the standard finite element spaces. In addition, the truncated wavelet expansion also satisfies the inverse property. The proof proceeds as in Lemma 2.1.

Lemma 2.2. *Let a wavelet* $\psi \in H^r(\mathbf{R})$ *be given, and let for any function* $f \in H^\nu(\mathbf{R})$, $0 \leq \nu < r$, *the symbol* f_M, $M > 0$, *denote the truncated wavelet-expansion of* f

$$f_M = \sum_{m=0}^M \sum_{k \in \mathbf{Z}} d_k^m \psi_{mk} + \sum_{k \in \mathbf{Z}} \langle f \mid \varphi_{0k} \rangle_{L^2} \cdot \varphi_{0k}.$$

Then there exists a constant C such that for every real number $s \leq \nu$, $|s| < r$, one has

$$\|f_M\|_\nu \leq C \cdot 2^{M(\nu-s)} \|f_M\|_s.$$

3. PROJECTION METHODS

Let $A : X \to Y$ be a linear bounded operator between Hilbert spaces X and Y. Consider the linear equation

$$Af = g.$$

If $AX \subset Y$ is not closed, then the problem of reconstructing f from a given right-hand side g is ill-posed; i.e., the set $f_\varepsilon := \{\tilde{f} \mid \|Af - A\tilde{f}\|_Y < \varepsilon\}$ is unbounded. Assume that we are given perturbed data g^ε with

$$\|g - g^\varepsilon\|_Y < \varepsilon.$$

The aim is to construct an approximation f_h by some regularization method and to estimate the difference $\|f - f_h\|_X$.

A generalized Galerkin or projection method is defined by families of subspaces

$$\{X_h \subset X\} \quad \text{and} \quad \{Y_h \subset Y\}$$

which are usually ordered so that $X_h \subset X_{h'}$ if $h' < h$. We determine the approximate solution f_h in the subspace X_h by solving

$$\langle Af_h \mid v \rangle_Y = \langle g^\varepsilon \mid v \rangle_Y, \quad \forall v \in Y_h. \tag{3.1}$$

Fix a basis for X_h and Y_h

$$X_h = \text{span}\{\psi_j \mid j \in I_h\}$$
$$Y_h = \text{span}\{v_j \mid j \in I_h\}.$$

Then (3.1) transforms into the following system of linear equations for the expansion coefficients

$$f_h = \sum_{j \in I_h} x_j \psi_j$$
$$A_h x = y$$

where $x = \{x_j \mid j \in I_h\}$ and where the entries of the matrix and the right-hand side y are given by

$$(A_h)_{jk} = \langle A\psi_k \mid v_j \rangle_Y, \quad j,k \in I_h \tag{3.2}$$

$$y_j = \langle g^\varepsilon \mid v_j \rangle_Y. \tag{3.3}$$

We assume that the subspaces X_h, Y_h are chosen such that A is injective on X_h, hence the linear system (3.2) has a unique solution.

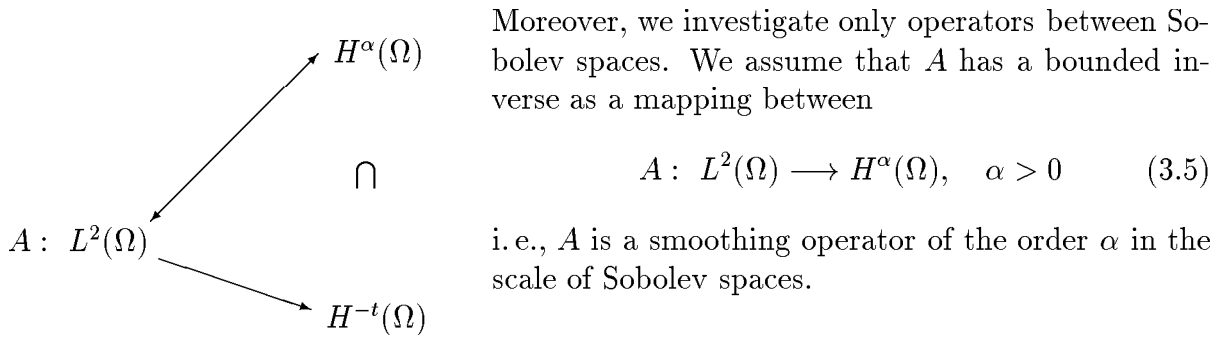We obtain the corresponding regularization operator T_h

$$T_h \; : \; Y \; \longrightarrow \; X$$

$$g \; \longmapsto \; T_h g = f_h.$$

Now we fix X_h to be the space of wavelet functions

$$X_h = \overline{\operatorname{span} \{\varphi_{0k}, \psi_{mk} \mid k \in \mathbf{Z}, \; 0 \le m \le M = -\log_2 h\}}. \tag{3.4}$$

For ease of notation we write $\psi_{-1,k}$ for $\phi_{0,k}$ from now on instead of $\psi(x/2 - k)$. Below we list some convergence results for wavelet projection methods which are obtained from the approximation properties stated in Lemmas 2.1 and 2.2 and the general theory of projection methods.

Projection methods have been studied in numerous publications, see, e. g., [6, 7] and the references there. We apply the results of [11] where projection methods were studied in terms of robustness and quasi-optimality. In order to give a short summary we restrict our attention to the least-squares method; i. e., $Y_h = A X_h$.

Moreover, we investigate only operators between Sobolev spaces. We assume that A has a bounded inverse as a mapping between

$$H^\alpha(\Omega)$$
$$\cap$$
$$A: \; L^2(\Omega)$$
$$H^{-t}(\Omega)$$

$$A: \; L^2(\Omega) \longrightarrow H^\alpha(\Omega), \quad \alpha > 0 \tag{3.5}$$

i. e., A is a smoothing operator of the order α in the scale of Sobolev spaces.

Examples of operators of this type are the Radon transform ($\alpha = 1/2$), Symm's operator ($\alpha = 1$), or more fashionable pseudo-differential operators with the symbol

$$\sigma(x, \omega) \sim (1 + |\omega|)^{-\alpha}.$$

The stability of wavelet projection methods for pseudo-differential operators has been studied in [1], but there the estimates were proved for the exactly given right-hand side g, and the influence of data errors was not investigated. However this is the crucial part when dealing with ill-posed problems.

Remark 3.1. To be precise one should distinguish a smoothing order and a smoothing potential in the following way. Let A be a convolution operator with a kernel $k \in \mathcal{S}'(\mathbf{R}^n)$

$$A f(x) = f * k(x) = \mathcal{F}^{-1}(\hat{f} \cdot \hat{k})(x) = \int_{\mathbf{R}^n} e^{2\pi i x \cdot \omega} \cdot \hat{f}(\omega) \cdot \hat{k}(\omega) \, d\omega.$$

Let $\mathcal{D}(A) = \{f \in L^2(\mathbf{R}^n) \mid Af \in L^2(\mathbf{R}^n)\}$. The operator A is injective on $\mathcal{D}(A)$ if there exists a function $\hat{p} \in \mathcal{S}'(\mathbf{R}^n)$ such that $\hat{p} \cdot \hat{k} = 1$ a. e. Usually A is referred to as being of *smoothing order* α if $\hat{k} = O(|\omega|^{-\alpha})$ as $|\omega| \to \infty$. We say that A has the *smoothing potential* α if $\hat{p} = O(r^\alpha)$. These two definitions agree for common operators with $\hat{k}(\omega) \sim |\omega|^{-\alpha}$ near ∞ but differ for instance if

$$\hat{k}(\omega) = \begin{cases} n^{-\alpha}, & \text{if } |\omega| \in [2n, 2n+1) \\ n^{-\beta}, & \text{if } |\omega| \in [2n+1, 2n+2) \end{cases}$$

which defines the operator of smoothing order α, but with the smoothing potential β. The relevant quantity in our applications is the smoothing potential.

In order to incorporate data errors we consider A as the mapping

$$A: \; L^2(\Omega) \longrightarrow H^{-t}(\Omega), \quad t \geq 0$$

and assume that

$$\|g - g^\varepsilon\|_{-t} \leq \varepsilon. \tag{3.6}$$

Typical values for t are $t = 0$, which allows to deal with L^2-errors, and $t > n/2$ which reflects the white noise over an n-dimensional space.

The last Sobolev index we need is ν. It indicates an *a priori* information about the smoothness of the exact solution f

$$f \in H^\nu(\Omega), \qquad \|f\|_\nu \leq \varrho. \tag{3.7}$$

These assumptions describe the standard setup for ill-posed problems. In this context a regularization method is called optimal if there exists a regularization parameter

$$h_{\mathrm{opt}} = h(\varepsilon, \varrho)$$

and a constant C independent of f and g such that the reconstruction error of $T_{h_{\mathrm{opt}}} g^\varepsilon = f_{h_{\mathrm{opt}}}$ can be bounded for $\varepsilon \to 0$ by

$$\|f - f_{h_{\mathrm{opt}}}\|_{L^2} \leq C(\varrho)\varepsilon^{\nu/(\nu+\alpha+t)}. \tag{3.8}$$

The properties of a particular projection method obviously depend crucially on the choice of X_h, Y_h. The common choice $Y_h = AX_h$ leads to least-squares methods. They have been studied in detail (see, e.g., [6, 11, 12]). There convergence properties are governed by the quantity

$$\beta_h = \sup \left\{ \|u\|_{L^2} \; \big| \; u \in X_h, \; \|Au\|_t \leq 1 \right\}.$$

Theorem 3.1. *Consider the least-squares method defined by X_h, $Y_h = AX_h$. Suppose that for every $f \in X$ there exists $u \in X_h$ such that*

$$\|f - u\|_{L^2} + \beta_h \|A(f - u)\|_t \leq C\|f\|_{L^2}$$

where C is independent of f. Then

$$\|f - f_h\|_{L^2} \leq C \left(\inf \left\{ \|f - u\|_{L^2} \; \big| \; u \in X_h \right\} + \beta_h \varepsilon \right).$$

This result follows from the standard theory (see, e.g., [6, 11]) with minor modifications, since we measure the data error in H^t instead of L^2.

Now we can state the main result of this section.

Theorem 3.2. *Let $\psi, \varphi \in H^r(\mathbf{R})$ be given. Let X_h be defined according to (3.4), and let Y_h be chosen as $Y_h = AX_h$ (wavelet least-squares method). Assume that (3.5)–(3.7)*

hold for $0 < \nu < r$, $|\alpha + t| < r$. Then the least-squares method is optimal; i.e., if h_{opt} is chosen according to

$$h_{\mathrm{opt}} = C \cdot \left(\frac{\varepsilon}{\varrho}\right)^{1/(\nu+\alpha+t)}$$

then the reconstruction error is asymptotically bounded by

$$\|f - f_{h_{\mathrm{opt}}}\|_{L^2} \le C(\varrho) \cdot \varepsilon^{\nu/(\nu+\alpha+t)}.$$

Proof. Following the outline given by Theorem 3.1 we first need an estimate for

$$\beta_h = \sup\left\{\|u\|_{L^2} \,\middle|\, u \in X_h, \; \|Au\|_t \le 1\right\}.$$

The condition (3.5) gives

$$\|Au\|_t \ge C_1 \|u\|_{t-\alpha}.$$

We apply Lemma 2.2 with $\nu = 0$, $s = t - \alpha$, and $h = 2^{-M}$

$$\|u\|_{L^2} \le C_2 h^{t-\alpha} \|u\|_{t-\alpha} \le (C_2/C_1) h^{t-\alpha} \|Au\|_t, \quad \forall u \in X_h.$$

Hence

$$\beta_h \le C h^{t-\alpha}. \tag{3.9}$$

Then for a fixed $f \in L^2(\mathbf{R})$ we estimate the value of

$$\gamma_h = \inf\left\{\|A(f - u)\|_t \,\middle|\, u \in X_h\right\}.$$

Here we use

$$\gamma_h \le \|A(f - f_M)\|_t \sim \|f - f_M\|_{t-\alpha}$$

where $f_M \in X_h$ denotes the truncated wavelet expansion of f. Now apply Lemma 2.1 with $s = t - \alpha$ and $\nu \equiv 0$

$$\gamma_h \le C h^{\alpha-t} \|f\|_{L^2}. \tag{3.10}$$

Hence the assumptions of Theorem 3.1 hold in the given situation, and we deduce the following estimate for $f_h = T_h g^\varepsilon$

$$\|f - f_h\|_{L^2} \le C\left(\inf\left\{\|f - u\|_{L^2} \,\middle|\, u \in X_h\right\} + \beta_h \varepsilon\right).$$

Since $\|f\|_\nu \le \varrho$, we again apply Lemma 2.1 and obtain

$$\|f - f_h\|_{L^2} \le C\left\{h^\nu \varrho + \varepsilon h^{t-\alpha}\right\}.$$

This quantity becomes minimal for the choice

$$h_{\mathrm{opt}} = C' \cdot \left(\frac{\varepsilon}{\varrho}\right)^{1/(\nu+\alpha+t)}$$

where C' depends on α, t, and ν. With this h we finally obtain the desired estimate

$$\|f - f_h\|_{L^2} \le C \cdot \varepsilon^{\nu/(\nu+\alpha+t)} \cdot \varrho^{(\alpha+t)/(\nu+\alpha+t)}. \qquad \square$$

4. WAVELET-VAGUELETTE DECOMPOSITION

Still the question remains whether to use wavelets at all. The rate of convergence proved in the previous section is achieved by many other regularization methods and various function systems like splines, finite elements, or trigonometric functions.

Wavelet-Galerkin methods will have practical significance if we are able to exploit the main algorithmic advantages of wavelets, namely, hierarchical structure and recursive computations. The first property will allow us to prove optimal convergence rates for an adaptive strategy. But since fast algorithms depend on an efficient evaluation of the scalar products $\langle g^\varepsilon \mid v_j \rangle_Y$, therefore, the second property is only useful if both function systems of the Galerkin method; i. e., $\{\psi_j\}$ and $\{v_j\}$, can be computed recursively. Below we will study the so-called wavelet-vaguelette decomposition and the related Galerkin methods.

The starting point for this particular projection method is the observation that the linear system (3.2) reduces to a diagonal system if $\{\psi_j\}$ form an orthogonal basis for X_h, and if the test functionals v_j are chosen to be the solutions of

$$A^* v_j = \kappa_j \cdot \psi_j$$

where A^* denotes the adjoint operator of A, and κ_j is defined by the requirement $\|v_j\| = 1$; i. e., we consider a dual least-squares method (cf. [11]). This construction requires at least that ψ_j is in the range of A^*; i. e., since A usually is a smoothing operator, this demands some regularity of X_h.

The name wavelet-vaguelette decomposition refers to the special case where X_h is given by (3.4) and where the test functionals are defined by

$$Y_h = \mathrm{span}\left\{ v_{mk} \;\middle|\; A^* v_{mk} = \kappa_{mk} \psi_{mk}, \; \|v_{mk}\|_Y = 1 \right\}.$$

The functions v_{mk} are called (A, ψ)-vaguelettes if they satisfy the following additional stability requirement. We require that two equivalent norms on $Y_h = \mathrm{span}\{v_j \mid j \in I_h\} \subset Y$ are given by $\| \cdot \|_Y$ and

$$\left(\sum_{j \in I_h} \left| \langle \cdot \mid v_j \rangle_Y \right|^2 \right)^{1/2}$$

or, slightly more general, we assume that the functions v_{mk} belong to Y^*, $v_{mk} \in Y^*$, and that the mapping

$$Y \ni g \longrightarrow \left(\sum_{j \in I_h} \left| (1 + 2^m)^t \langle g \mid v_j \rangle_{Y \times Y^*} \right|^2 \right)^{1/2} \tag{4.1}$$

is continuous for some t with respect to $\| \cdot \|_Y$. We use the latter assumption for example to deal with the white noise model $Y = H^{-n/2}(\mathbf{R}^n)$, $Y^* = H^{n/2}(\mathbf{R}^n)$. In the latter case κ_m ought to be chosen such that

$$\langle v_{mk} \mid v_{mk} \rangle_{Y \times Y^*} \asymp 1. \tag{4.2}$$

The triplet $\{\psi_{mk}, v_{mk}, \kappa_{mk}\}$ has properties very similar to the singular value decomposition (SVD) of A, and, indeed, it may be used as an easily computable substitute for SVD in various applications. Similarly one defines the scaling vaguelettes via $A^* u_{mk} = \lambda_{mk} \varphi_{mk}$. The concept of vaguelette decompositions goes back to an idea by P. Tchamitchian; for a more detailed introduction see [10].

The resulting projection method has been studied in a series of papers by D. Donoho (see, e. g., [5]) in a stochastic framework. There a white noise model for the data error

was used, and the convergence results were given in terms of Bayes estimators. Those results have their counterparts in the Hilbert space framework used in this paper. To this end we introduce the following simple (nonlinear) adaptive regularization operator.

The wavelet-Galerkin method studied in the previous section was based on a truncated wavelet expansion with a fixed truncation level $M = -\log_2 h$. An adaptive strategy should allow us to add local refinements if some criterion indicates small but relevant structures at this point.

As a local criterion we use the absolute value of $|\langle g^\varepsilon \mid v_{mk}\rangle_Y|$. We define

$$T_{h,\delta} \;:\; Y \;\longrightarrow\; X$$

$$g \;\longmapsto\; T_{h,\delta} g \;:=\; f_{h,\delta}$$

$$:= \sum_{-1 \leq m \leq M} \sum_{k \in \mathbf{Z}} \eta_m \Big(|\langle g \mid v_{mk}\rangle_Y| \Big) \langle g \mid v_{mk}\rangle_Y \cdot \kappa_{mk}^{-1} \psi_{mk}$$

where

$$\eta_m(t) = \begin{cases} 0, & \text{for } m > M \text{ or } |t| \leq \delta(m) \\ 1, & \text{otherwise.} \end{cases}$$

If we use compactly supported wavelets, then a large value of $|\langle g \mid v_{mk}\rangle_Y|$ will indeed lead to a local change of $f_{h,\delta}$. The support of the refinement gets smaller as m gets larger. The cutoff levels $\delta(m)$ may change from scale to scale; here we use a simple cutoff which increases as m tends to infinity. For some $\beta > 0$ and $\mu \in [0, M)$ we choose

$$\delta(m) \;=\; \begin{cases} \delta \cdot 2^{m\beta}, & \text{if } m \geq M - \mu \\ 0, & \text{otherwise.} \end{cases} \tag{4.3}$$

Now we have to optimize the regularization parameters: the number M of considered scales (or h), the number μ of scales with adaptive filtering, and δ, β, so that the value

$$\|f - T_{h,\delta} g^\varepsilon\|_X$$

becomes minimal.

As in the previous section we consider the special class of operators. We replace (3.5), which demands that high frequencies are damped by the operator A, by

$$C_1 \cdot 2^{-\alpha m} \leq |\kappa_{mk}|, \quad 0 < C_1 < \infty. \tag{4.4}$$

Theorem 4.1. *Let the operator A have a wavelet-vaguelette decomposition $\{\psi_{mk}, v_{mk}, \kappa_{mk}\}$ satisfying (4.4). Let (3.6), (3.7) hold for $0 < \nu < r$ and $\alpha - t < r$ with t as in (4.1). Then the adaptive wavelet-vaguelette regularization method $T_{h,\delta}$ with truncation levels (4.3) is optimal; i.e., for every $\delta > 0$ and $h = C \cdot (\varepsilon/\varrho)^{1/(\nu+\alpha+t)}$, respectively, $M = \mathrm{const}(\varrho) \cdot \log(\varepsilon)/(\nu + \alpha + t)$, the truncation error is bounded by*

$$\|f - f_{h,\delta}\|_X \leq C(\varrho) \cdot \varepsilon^{\nu/(\nu+\alpha+t)}.$$

Proof. We split the error into the data error and the regularization error

$$\|f - f_{h,\delta}\|_X \leq \|f_h - f_{h,\delta}\|_X + \|f - f_h\|_X.$$

Here f_h is defined, using the convention $\eta_m(t) = 0$ if $m > M$, as

$$f_h := \sum_{m \in \mathbf{Z}, k \in \mathbf{Z}} \eta_m\left(\langle g^\varepsilon \mid v_{mk}\rangle_{Y \times Y^*}\right) \cdot \langle g \mid v_{mk}\rangle_{Y \times Y^*} \cdot \kappa_m^{-1} \cdot \psi_{mk}.$$

Observe that this is different from $T_{h,\delta}(g)$, i.e., the reconstruction obtained if the data were unperturbed, which is the quantity usually used for defining data and regularization errors.

The data error is bounded by virtue of the conditions $\eta_m(t) = 0$ if $m > M$, $|\eta_m(t)| \le 1$, and formulas (4.1), (4.2), and (4.4)

$$
\begin{aligned}
\|f_h - f_{h,\delta}\|_X^2 &= \sum_{m \in \mathbf{Z}, k \in \mathbf{Z}} \left|\langle f_h - f_{h,\delta} \mid \psi_{mk}\rangle_X\right|^2 \\
&\le \sum_{m \in \mathbf{Z}, k \in \mathbf{Z}} \left|\eta_m\left(\langle g^\varepsilon \mid v_{mk}\rangle_{Y \times Y^*}\right) \cdot \kappa_m^{-1} \cdot \langle g - g^\varepsilon \mid v_{mk}\rangle_{Y \times Y^*}\right|^2 \\
&\le \text{const} \cdot \sum_{m < M} \sum_{k \in \mathbf{Z}} 4^{(\alpha-t)\cdot m} \cdot \left|\langle g - g^\varepsilon \mid v_{mk}\rangle_{Y \times Y^*}\right|^2 \\
&\le C_d^2 \cdot 4^{(\alpha+t)\cdot M}\|g - g^\varepsilon\|_Y^2 \ \le\ C_d^2 \cdot 4^{(\alpha+t)M} \cdot \varepsilon^2.
\end{aligned}
$$

Using

$$f = \sum_{m \in \mathbf{Z}, k \in \mathbf{Z}} \langle f \mid \psi_{mk}\rangle \cdot \psi_{mk} = \sum_{m \in \mathbf{Z}, k \in \mathbf{Z}} \langle g \mid v_{mk}\rangle_{Y \times Y^*} \cdot \kappa_m^{-1} \cdot \psi_{mk}$$

we see that the regularization error is bounded because of the requirement $\eta_m(t) = 1$ for all $m < M - \mu$ and the regularity assumption (3.7)

$$
\begin{aligned}
\|f - f_h\|_X^2 &= \sum_{m \in \mathbf{Z}, k \in \mathbf{Z}} \left|\left(1 - \eta_m(\langle g^\varepsilon \mid v_{mk}\rangle_{Y \times Y^*})\right) \cdot \kappa_m^{-1} \cdot \langle g \mid v_{mk}\rangle_{Y \times Y^*}\right|^2 \\
&\le \sum_{m > M - \mu} \sum_{k \in \mathbf{Z}} \left|\langle f \mid \psi_{mk}\rangle_X\right|^2 \ \le\ C_r^2 \cdot 4^{-\nu(M-\mu)} \cdot \|f\|_{X^\nu}^2.
\end{aligned}
$$

Thus we get

$$\|f - f_{h,\delta}\|_X \le C_r \cdot 2^{\nu \cdot \mu} \cdot 2^{-\nu \cdot M} \cdot \varrho + C_d \cdot 2^{(\alpha+t)\cdot M} \cdot \varepsilon.$$

Choosing M as above for a fixed μ we get the desired result. $\square$

The estimates in this proof are quite crude for the scales with $m \in [M - \mu, M]$, and in practice we can expect better results through tuning the parameters δ and β such that the regularization error will not increase with μ as the above estimation suggests. In particular, in the case when the error consists mostly of white noise better results may be achieved.

5. EXAMPLES

The existence of a wavelet-vaguelette decomposition and the requirement that $\{v_{mk}\}$ can be computed efficiently impose some restrictions on the operator A. These conditions are met for example by convolution operators as long as they fulfil the quasi-homogeneity condition and partly by some pseudo-differential operators.

Definition 5.1. We call a convolution operator $A : f \mapsto f * k = \mathcal{F}^{-1}(\hat{f} \cdot \hat{k})$ quasi-homogeneous of degree α if its Fourier transformed kernel $\hat{k}$ is continuously differentiable over $\mathbf{R}^n \setminus \{0\}$, and if there exists a function $\hat{p} \in \mathcal{S}'$ such that $\hat{p} \cdot \hat{k} = 1$, and for some continuous function $\Omega : S^1 \to \mathbf{R}$ and $r(x) := \|x\|_2$ the estimates $\Omega(x/r) \cdot r^\alpha \cdot \hat{k} = 1 + o(1)$ near ∞ and $|\hat{k}| > 1/C_0 > 0$ near zero are satisfied.

Simple examples of quasi-homogeneous convolution operators of degree α are negative powers of the Laplacian $\Delta^{-\alpha/2}$ with $\hat{k} = r^{-\alpha}$ or the operators generated by $\hat{k}(\omega) = (1 + |\omega|^2)^{-\alpha/2}$.

Below we show that for such an operator there always exists a wavelet-vaguelette decomposition. Moreover, in order to calculate the scalar products $\langle g^\varepsilon \mid v_{mk} \rangle_{Y \times Y*}$ from given sample values of the perturbed data one only has to calculate one vaguelette or one scaling vaguelette per scale m; the others are obtained by integer shifts. These functions can be calculated in advance with high precision using fast Fourier transforms. Theoretically one could use Mallats algorithm to calculate all scalar products $\langle g \mid v_{mk} \rangle$ from the scalar products with the scaling vaguelette on the finest scale (cf. Proposition 5.2). However, this seems to be numerically unstable in most cases.

In case of powers of the Laplacian one needs even less pre-calculations: one has to calculate only a single vaguelette/scaling vaguelette using the homogeneity of the Laplacian. After that one computes all scalar products $\langle g \mid v_{mk} \rangle$ and the corresponding approximate solution $f_{h,\delta}$ recursively, hence, the number of the required operations grows only linearly with the number of sample values. This means that for the above-mentioned operators this method could be applicable in real-time applications. Moreover, all computations are local (provided the vaguelettes decay sufficiently fast), and the recursive computations can be carried out faster than in the traditional approach using fast Fourier transforms.

To be more precise we give some propositions.

Proposition 5.1. *Let A be a quasi-homogeneous convolution operator of the order α. Let ψ and ϕ be the wavelet and its scaling function of an orthonormal wavelet basis over $\mathbf{R}^n$ with $\hat{\phi}$ and $\hat{\psi}$ decaying of the order $\alpha + n/2 + t$ near ∞. Then A is injective, the scaled and shifted wavelets ψ_{mk}, ϕ_{mk} are elements of $\mathcal{R}(A^*)$, and the functions v_{mk} (vaguelettes) with $A^* v_{mk} = 2^{-\alpha m} \cdot \psi_{mk}$ and u_{mk} (scaling vaguelettes), $A^* u_{mk} = 2^{-\alpha m} \cdot \phi_{mk}$, satisfy for some C, $0 < C < \infty$, and $\forall M \in \mathbf{Z}$, $\forall g \in H^t$ the estimate*

$$\|g\|_t^2 \leq C \cdot \sum_{k \in \mathbf{Z}, m > M} \left| (1 + 2^m)^t \langle g \mid v_{mk} \rangle_{H^t \times H^{-t}} \right|^2 + \left(1 + 2^M \right)^{2t} \sum_{k \in \mathbf{Z}} \left| \langle g \mid u_{M,k} \rangle_{H^t \times H^{-t}} \right|^2. \quad (5.1)$$

If, moreover, A is continuous we also have

$$\|g\|_t^2 \geq c \cdot \sum_{k \in \mathbf{Z}, m > M} \left| (1 + 2^m)^t \langle g \mid v_{mk} \rangle_{H^t \times H^{-t}} \right|^2 + \left(1 + 2^M \right)^{2t} \sum_{k \in \mathbf{Z}} \left| \langle g \mid u_{M,k} \rangle_{H^t \times H^{-t}} \right|^2 \quad (5.2)$$

for some $c > 0$.

Proof. The proof is simple in the case when the wavelet has sufficiently many (more than $\alpha + t$) vanishing moments, and the operator A is a homeomorphism of the order α in the scale of Sobolev spaces; i.e., A and A^{-1} are continuous operators $\forall s \in \mathbf{R}$

$$A \ : \ H^s \longmapsto H^{s+\alpha}$$

$$A^{-1} \ : \ H^{s+\alpha} \longmapsto H^s.$$

In this case we have $g = A \cdot f$ for some $f \in H^{t-\alpha}$, and the results follow from the well-known estimate [10]

$$\|f\|_s \sim \sum_{m \in \mathbf{Z}, k \in \mathbf{Z}} (1 + 2^m)^s \langle f \mid \psi_{mk} \rangle_{H^s \times H^{-s}}$$

for $f \in H^s$, $|s| < \operatorname{Reg} \psi$ where $\operatorname{Reg} \psi$ is the regularity of ψ. We have $\operatorname{Reg} \psi > \alpha + t$ by our assumptions. In that case we have

$$
\begin{aligned}
\|g\|_t &\sim \|f\|_{t-\alpha} \\
&\sim \sum_{m \in \mathbf{Z}, k \in \mathbf{Z}} (1 + 2^m)^{t-\alpha} \langle f \mid \psi_{mk} \rangle_{H^{t-\alpha} \times H^{-t+\alpha}} \\
&= \sum_{m \in \mathbf{Z}, k \in \mathbf{Z}} 2^{\alpha m} \cdot (1 + 2^m)^{t-\alpha} \langle g \mid v_{mk} \rangle_{H^t \times H^{-t}} \\
&\sim \sum_{k \in \mathbf{Z}, m > M} \left| (1 + 2^m)^t \langle g \mid v_{mk} \rangle_{H^t \times H^{-t}} \right|^2 + \left(1 + 2^M\right)^{2t} \sum_{k \in \mathbf{Z}} \left| \langle g \mid u_{M,k} \rangle_{H^t \times H^{-t}} \right|^2.
\end{aligned}
$$

A full proof of this proposition is rather technical, therefore, we only outline it here. The difficulties arise because $\hat{k}$ may be unbounded near 0, and, for example, $\hat{k}$ may oscillate between $r^{-\alpha}$ and $r^{-\beta}$ near ∞. A complete proof may be found in [4, Chapter 3.5].

Let $\hat{k}$ denote the Fourier transformed kernel of A. We use the notation $\hat{p}$ for the function with $\hat{k} \cdot \hat{p} = 1$. We then observe that the Fourier transform of the vaguelettes v_{mk} and u_{mk} is given by

$$
\begin{aligned}
\hat{v}_{mk}(\omega) &= 2^{-(\alpha+1/2)m} \cdot e^{-2\pi i \cdot 2^m k \cdot \omega} \cdot \overline{\hat{p}} \cdot \hat{\psi}(2^{-m}\omega) \\
\hat{u}_{mk}(\omega) &= 2^{-(\alpha+1/2)m} \cdot e^{-2\pi i \cdot 2^m k \cdot \omega} \cdot \overline{\hat{p}} \cdot \hat{\phi}(2^{-m}\omega).
\end{aligned}
$$

Because the wavelets decay at ∞ and always have some vanishing moments

$$\int_{\mathbf{R}^n} x^\beta \cdot \psi(x) \, dx = 0, \quad \text{for } |\beta| \le m_0$$

therefore, there exist constants $a, b, c \in (0, \infty)$ such that

$$|\hat{v}_{mk}(2^{-m}\omega)| \le \min\left(a \cdot r^{t-n/2}, b, c \cdot r^{m_0}\right) \tag{5.3}$$

$$|\hat{u}_{mk}(2^{-m}\omega)| \le \min\left(a \cdot r^{t-n/2}, b\right). \tag{5.4}$$

Here we used the fact that $|\hat{p}| \le C_0 + C_\infty \cdot r^\alpha$ which follows from the quasi-homogeneity of A (with $r = \|\cdot\|_2$).

Applying extensions of techniques used in [2, Chapter 3.3], standard estimates on the geometric series, and the definition of the Sobolev spaces we conclude that any family generated by shifting any one vaguelette per scale (and shifting a single scaling vaguelette on the coarsest scale) leads to the estimate (5.1), provided that (5.3), (5.4) are satisfied.

The estimate (5.2) follows from considering the dual family $\tilde{v}_{mk} := 2^{+\alpha m} A \psi_{mk}$, $\tilde{u}_{mk} := 2^{+\alpha m} A \phi_{mk}$ which exists if A is continuous, and which also satisfies equations (5.3), (5.4) with suitable constants and with $2\alpha - t - n/2$ substituted for $t - n/2$. We observe further that

$$
\begin{aligned}
\langle \tilde{v}_{mk} \mid v_{m',k'} \rangle &= \langle \psi_{mk} \mid \psi_{m',k'} \rangle = \delta_{m,m'} \cdot \delta_{k,k'} \\
\langle \tilde{v}_{mk} \mid u_{m',k'} \rangle &= \langle \psi_{mk} \mid \phi_{m',k'} \rangle = \delta_{m,m'} \cdot \delta_{k,k'}
\end{aligned}
$$

if $m \geq m'$. Thus the families $\{v_{mk}, u_{Mk}\}$ and $\{\tilde{v}_{mk}, \tilde{u}_{Mk}\}$ are biorthogonal. Standard results on biorthogonal families then give the estimate (5.2). $\square$

The next proposition gives the key to the fast evaluation of the operator $T_{\delta,h}$.

Proposition 5.2. *With the assumptions of the previous proposition and the scaling equations (2.2), (2.3) for the wavelet and its scaling function we have the following scaling equations for vaguelettes*

$$u_{mk} = 2^{\alpha} \cdot \sum_{l \in \mathbf{Z}} h_l \cdot u_{m+1,2k+l}$$

$$v_{mk} = 2^{\alpha} \cdot \sum_{l \in \mathbf{Z}} g_l \cdot u_{m+1,2k+l}.$$

Proof. The scaling equations (2.2), (2.3) are equivalent to $\hat{\phi}(\omega) = \hat{h} \cdot \hat{\phi}(2\omega)$ and $\hat{\psi}_e(\omega) = \hat{g} \cdot \hat{\phi}(2\omega)$ for some trigonometric polynomials

$$h_n(\xi) = \sum h_n e^{\pi i n \xi} \qquad \text{and} \qquad \hat{g}(\xi) = \sum g_n e^{\pi i n \xi}.$$

Let $\hat{k}$ denote the Fourier transformed kernel of A. Employing the notation $\hat{p}$ for the function with $\hat{k} \cdot \hat{p} = 1$ we have the following formulas for the scaling vaguelettes u_m with $u_{m,k} = T_{2^{-m}k} u_m$ and $A^* u_{m,k} = \phi_{m,k}$

$$
\begin{aligned}
u_m(x) &= 2^{\alpha} \cdot \mathcal{F}^{-1}\left(\overline{\hat{p}} \cdot \hat{h}\left(\frac{\omega}{2^m}\right) \hat{\phi}_{m+1}\right) \\
&= 2^{\alpha} \cdot \mathcal{F}^{-1}\left(\hat{h}\left(\frac{\omega}{2^m}\right) \left(\overline{\hat{p}} \cdot \hat{\phi}_{m+1}\right)\right) \\
&= 2^{\alpha} \cdot \sum_l h_l \cdot T_{2^{-m+1}l} \cdot \mathcal{F}^{-1}\left(\overline{\hat{p}} \cdot \hat{\phi}_{m+1}\right) \\
&= 2^{\alpha} \cdot \sum_l h_l \cdot u_{m+1,l}
\end{aligned}
$$

if the sequences h_l, g_l decay sufficiently fast to have absolutely converging sums, which is always the case in applications.

From $T_{2^{-m}k} u_{m+1,n} = u_{m+1,2k+n}$ we derive the desired result for $u_{m,k}$. An analogous proof yields the formula for $v_{m,k}$. $\square$

With the above scaling equations at hand we only need to compute scalar products $\langle g^{\varepsilon} \mid u_{Mk} \rangle$ on the finest scale which can be easily done, since the vaguelettes u_{Mk} on the finest scale have small (effective) support. All the other scalar products are then calculated using a scaled version of Mallats algorithm.

Propositions 5.1 and 5.2 yield the following theorem.

Theorem 5.1. *Let A be a quasi-homogeneous convolution operator of the order α. Let ψ and ϕ be a wavelet and its scaling function of an orthonormal wavelet basis over $\mathbf{R}^n$ with $\hat{\phi}$ and $\hat{\psi}$ decaying of the order $\alpha + n/2 + t$ near ∞. Then the adaptive method for solving the ill-posed problem of reconstructing a function $f \in H^{\nu}$ with $\|f\|_{\nu} \leq \varrho$ from*

noisy data $g^\varepsilon = Af + z$ with $\|z\|_{-t} \le \varepsilon$ described above can be applied at computational costs that are linear in the number of sample values of the data (if pre-computations are not taken into account) and has the optimal order of convergence $\nu/(\alpha + t + \nu)$ for $\varepsilon \to 0$.

6. NUMERICAL RESULTS

We conclude our paper with some illustrations. There we applied the wavelet-vaguelette method to a homogeneous standard problem

$$Af = \Delta^{-1}f = \mathcal{F}^{-1}\left(\hat{f} \cdot r^{-2}\right) = \int_{\mathbf{R}} e^{2\pi i \omega \cdot x} \hat{f}(\omega) \cdot |\omega|^2 \, d\omega$$

where the ill-posed inverse problem consists in evaluating the Laplacian of perturbed data.

As data we used some test function constructed of splines of various degrees. We added an ε-multiple of an evenly distributed random variable to the sample values.

We performed the cutoff as well as the adaptive regularization as described above. Both algorithms seem to give similarly good results for this example.

The figures show:

1. The data and their Fourier transform.

2. The true solution and its Fourier transform.

3. The approximate solution with cutoff regularization and M chosen too large $(M = 4)$.

4. The approximate solution with cutoff regularization and well chosen M $(M = 2)$.

5. The approximate solution with adaptive regularization and well chosen parameters.

6. The wavelets used (orthogonal Daubechies wavelets with 10 vanishing moments; we denote them as Dau10) and their Fourier transform.

7. The vaguelettes used (calculated via FFT from Dau10 for the Laplace example) and their Fourier transform.
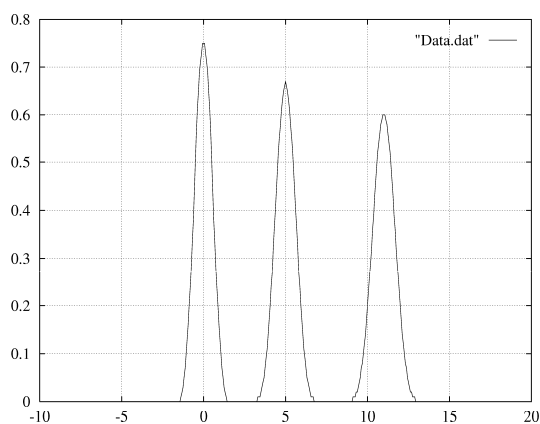
Fig. 1a. Unperturbed data: the sum of shifted splines of orders 3, 4, and 5 respectively.
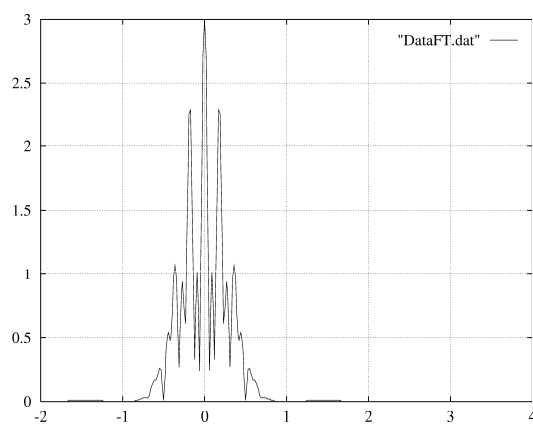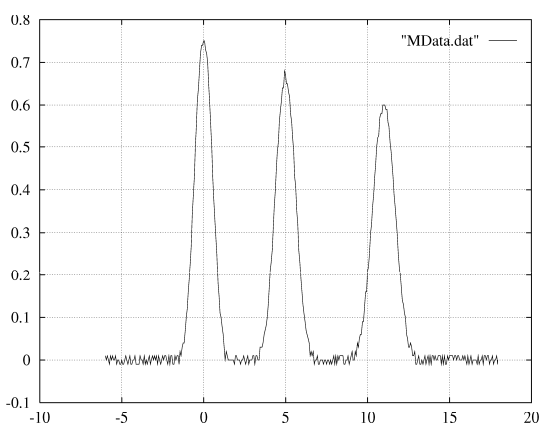

Fig. 1b. The Fourier transformed data.

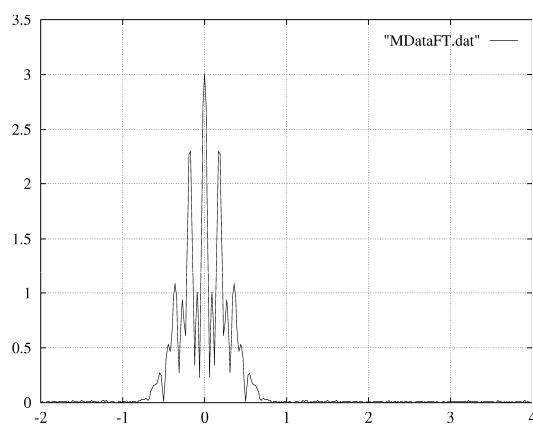Fig. 1c. The "measured" data: our data with 2.5% additive random noise.

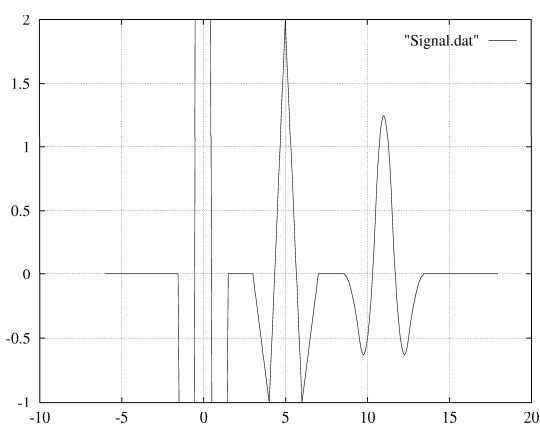Fig. 1d. The Fourier transformed noisy data.

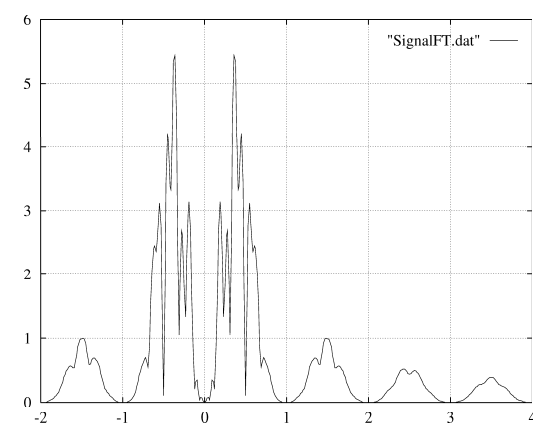Fig. 2a. The solution: the second derivative of the sum of shifted splines.

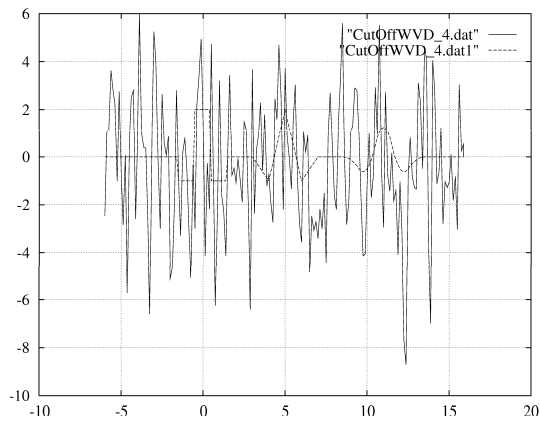Fig. 2b. The Fourier transform of the solution.

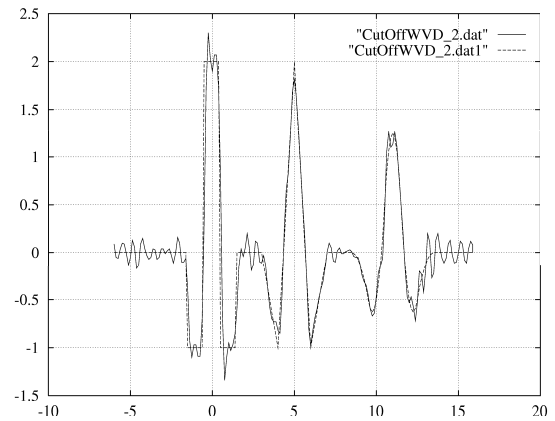Fig. 3. The result of applying the cutoff algorithm to the perturbed data, $M = 4$ (too large).

Fig. 4. The result of applying the cutoff algorithm to the perturbed data, $M = 2$ (optimal).
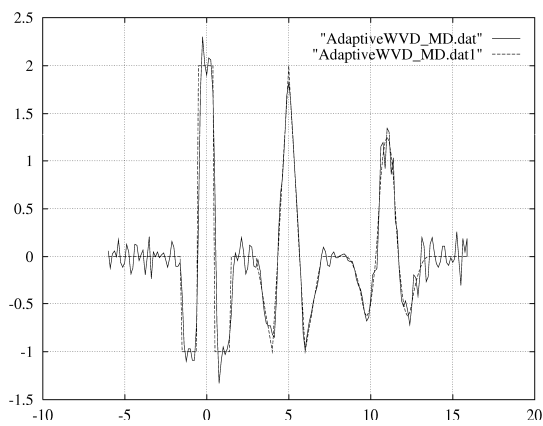
Fig. 5. The result of applying the adaptive algorithm to the perturbed data.
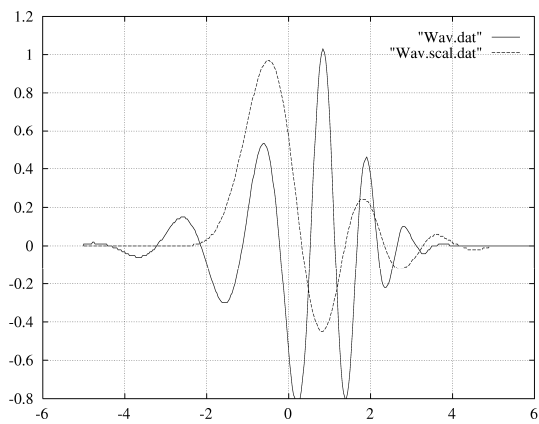

Fig. 6a. The wavelets. Here we used orthogonal Daubechies wavelets with 10 vanishing moments.
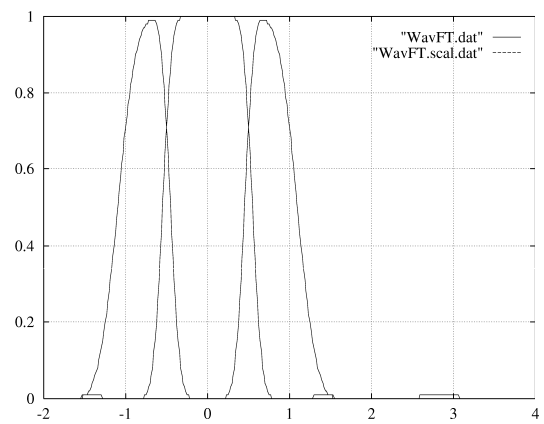

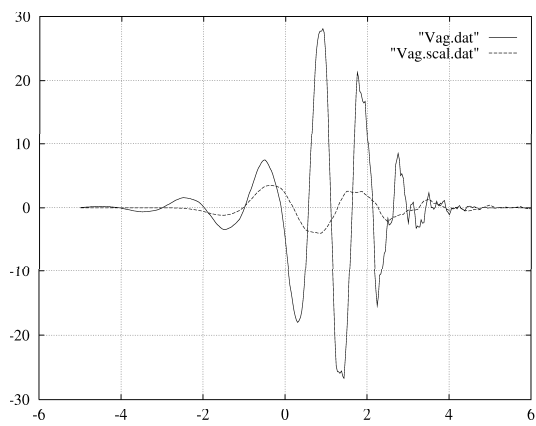Fig. 6b. The Fourier transform of the wavelets used.

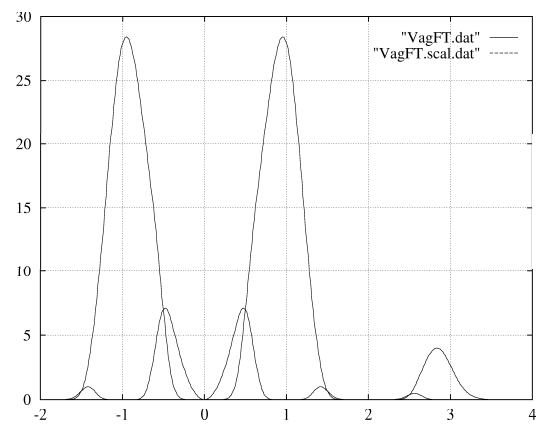Fig. 7a. The vaguelettes calculated via FFT from the Dau10 wavelets and the Fourier transformed kernel.

Fig. 7b. The Fourier transform of the vaguelettes.

REFERENCES

1. W. Dahmen, S. Prößdorf, and R. Schneider, Wavelet approximation methods for pseudo-differential equations. I, II. *Advances in Comp. Math.* (1993) **1**, 259–335.

2. I. Daubechies, Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.* (1988) **41**, 909–996.

3. R. DeVore and V. Popov, Interpolation of Besov spaces. *Trans. Amer. Math. Soc.* (1988) **305**(1), 397–414.

4. V. Dicken, Die Behandlung inverser Probleme durch Wavelet-Zerlegungen. *Diploma work.* Philipps University, Marburg, 1994.

5. D. L. Donoho, Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Technical Report.* Stanford University, Department of Statistics, 1992.

6. A. K. Louis, *Inverse und Schlecht Gestellte Probleme.* Teubner, Stuttgart, 1989.

7. A. K. Louis and P. Maaß, Smoothed projection methods for the moment problem. *Numer. Math.* (1990) **59**, 295–310.

8. A. K. Louis, P. Maaß, and A. Rieder, *Wavelets: Eine Einführung in Theorie und Anwendungen.* Teubner Verlag, Stuttgart, 1994.

9. S. Mallat, Multi-resolution approximation and wavelets. *Trans. Amer. Math. Soc.* (1988) **315**, 69–88.

10. Y. Meyer, *Ondelettes et Operateurs. Vol. 1, 2.* Herrmann, Paris, 1990.

11. F. Natterer, Regularisierung schlecht gestellter Probleme durch Projektionsverfahren. *Numer. Math.* (1977) **28**, 511–522.

12. R. Plato and G. Vainikko, On the regularization of projection methods for solving ill-posed problems. *Numer. Math.* (1990) **57**, 63–79.