Sparse Approximate Solution of Partial Differential Equations*

Sadegh Jokar† Volker Mehrmann† Marc Pfetsch‡

Harry Yserentant†

03/25/2008

Abstract

A new concept is introduced for the adaptive finite element discretization of partial differential equations that have a sparsely representable solution. Motivated by recent work on compressed sensing, a recursive mesh refinement procedure is presented that uses linear programming to find a good approximation to the sparse solution on a given refinement level. Then only those parts of the mesh are refined that belong to nonzero expansion coefficients. Error estimates for this procedure are refined and the behavior of the procedure is demonstrated via some simple elliptic model problems.

Keywords partial differential equation, sparse solution, dictionary, compressed sensing, restricted isometry property, mutual incoherence, hierarchical basis, linear programming

AMS subject classification. 65N50, 65K05, 65F20, 65F50

1 Introduction

The sparse representation of functions via a linear combination of a small number of basic functions has recently received a lot of attention in several mathematical fields such as approximation theory [15, 28, 31, 32] as well as signal and image processing [4, 5, 6, 7, 8, 9, 11, 17, 18, 19, 20, 21, 22, 23]. In terms of representations of functions, we can describe the problem as follows. Consider a linearly dependent set of n functions ϕ_i, $i = 1, 2, \ldots, n$, (*a dictionary* [12]) and a function f represented as

$$f = \sum_{i=1}^{n} x_i \, \phi_i.$$

Since the set of functions is not linearly independent, this representation is not unique and we may want to determine the sparsest representation, i.e., a representation with a maximal number of vanishing coefficients among $x_1, \ldots, x_n$.

*Supported by the *Deutsche Forschungsgemeinschaft* through the DFG Research Center MATHEON *Mathematics for key technologies* in Berlin.

†Institut für Mathematik, TU Berlin, Str. des 17. Juni 136, 10623 Berlin, Germany. {jokar,mehrmann,yserenta}@math.tu-berlin.de.

‡Zuse Institute Berlin, Takustr. 7, 14195 Berlin, Germany. pfetsch@zib.de

In the setting of numerical linear algebra, this problem can be formulated as follows. Consider a linear system

$$\Phi x = b, \tag{1}$$

with $\Phi \in \mathbb{R}^{m,n}$, where $m \leq n$ and $b \in \mathbb{R}^m$. The columns of the matrix Φ and the right hand side b represent the functions ϕ_i and the function f, respectively, with respect to some basis of the relevant function space. The problem is then to find the sparsest possible solution x, i.e., x has as many zero components as possible. This optimization problem is in general NP-hard [24, 29]. Starting from the work of [11], however, a still growing number of articles have developed sufficient conditions that guarantee that an (approximate) sparse solution $\hat{x}$ to (1) can be obtained by solving the linear program

$$\min \|x\|_1, \ \ s.t. \ \Phi x = b,$$

which can be done in polynomial time [26, 27]. We will give a brief survey of this theory in Section 2.2.

In the literature, the development has mostly focused on the construction of appropriate coding matrices Φ that allow for the sparse representation of a large class of functions (signals or images). Furthermore, properties of the columns of the matrix (or the dictionary) have been investigated, which guarantee that the computation of the sparse solution can be done efficiently via a linear programming approach, see, for instance, [9]. Often the term *compressed sensing* is used for this approach.

In this paper we consider a related but different problem. We are interested in the numerical solution of partial differential equations

$$Lu = f,$$

with a differential operator L, to be solved in a domain $\Omega \subset \mathbb{R}^d$ with smooth boundary Γ and appropriate boundary conditions given on Γ.

Considering a classical Galerkin or Petrov-Galerkin finite element approach, see e.g. [3], one seeks a solution u in some function space $\mathbb{U}$ (which is spanned by $\phi_1, \ldots, \phi_n$), represented as

$$u = \sum_{i=1}^{n} u_i \, \phi_i. \tag{2}$$

Again we are interested in sparse representations with a maximal number of vanishing coefficients u_i. In contrast to the cases discussed before, here we would like to construct the space $\mathbb{U}$ and the basis functions ϕ_i in the finite element discretization in such a way that first of all a sparse representation of the solution to (2) exists and second that it can be determined efficiently. Furthermore, it would be ideal if the functions ϕ_i could be constructed in a multilevel or adaptive way.

The usual approach to achieve this goal is to use local a posteriori error estimation to determine where a refinement, i.e., the addition of further basis functions is necessary. For example, in the dual weighted residual approach [2] this is done by solving an optimization problem for the error.

In this article, we examine the possibility to use similar approaches as those used in compressed sensing, i.e., to use ℓ_1-minimization and linear programming to perform the adaptive refinement in the finite element method in such a way that the solution is sparsely represented by a linear combination of basis functions. In order to achieve this goal, we propose the following framework.

We determine $u \in \mathbb{U}$ as the solution of the weak formulation

$$(v, Lu - f) = 0 \;\text{ for all }\; v \in \mathbb{V}.$$

Here, $\mathbb{V}$ is a space of test functions and $(\cdot,\cdot)$ is an appropriate inner product. In the simplest version of a two-level approach, we construct finite dimensional spaces of coarse and fine basis functions $\mathbb{U}_1^n \subset \mathbb{U}_1^N \subset \mathbb{U}$ and corresponding spaces for coarse and fine test functions $\mathbb{V}_1^n \subset \mathbb{V}_1^N \subset \mathbb{V}$. Then we determine the sparsest solution in $\mathbb{U}_1^N$, such that

$$(v, Lu - f) = 0 \;\text{ for all } v \in \mathbb{V}_N^1 \setminus \mathbb{V}_n^1$$

via the solution of an underdetermined system of the form (1). Based on the sparse solution, we determine new coarse and fine spaces $\mathbb{U}_2^n \subset \mathbb{U}_2^N \subset \mathbb{U}$, $\mathbb{V}_2^n \subset \mathbb{V}_2^N \subset \mathbb{V}$, and iterate this procedure.

This framework combines the ideas developed in compressed sensing with well-known concepts arising in adaptive and multilevel finite element methods. But instead of using local and global error estimates to obtain error indicators by which the grid refinement is controlled, here the solution of the ℓ_1-minimization is used to control the grid refinement and adaptivity.

Many issues of this approach have, however, not yet been resolved, in particular, the theoretical analysis of this approach (see Section 4). We see the following potential advantages and disadvantages of this framework. On the positive side, the ℓ_1-minimization approach allows for an easy automation. We will demonstrate this with some numerical examples in Section 5. On the downside, the analysis of the approach seems to be hard even for classical elliptic problems, see Section 4 and due to the potentially high complexity of the linear programming methods this approach will only be successful, if the procedure needs only a few levels and a small sparse representation of the solution exists, see Section 5.

2 Notation and Preliminaries

2.1 Notation

For $m, n \in \mathbb{N} = \{1, 2, \dots\}$, we denote by $\mathbb{R}^{m,n}$ the set of real $m \times n$ matrices, and by I_n the $n \times n$ identity matrix. Furthermore, we denote the Euclidean inner product on $\mathbb{R}^n$ by $\langle \cdot, \cdot \rangle$, i.e., for $x, y \in \mathbb{R}^n$,

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j \, y_j.$$

For $1 \le p \le \infty$, the ℓ_p-norm of $x \in \mathbb{R}^n$ is defined by

$$\|x\|_p := \big(\sum_{j=1}^{n} |x_j|^p \big)^{\frac{1}{p}},$$

with the special case

$$\|x\|_\infty := \max_{j\in\{1,\dots,n\}} |x_j|,$$

if $p = \infty$.

The definition of $\|\cdot\|_p$ can also be formally extended to the case that $0 \le p < 1$. For $0 < p < 1$, $\|\cdot\|_p$ is only a quasi-norm, since the triangle inequality is not satisfied, but still a *generalized triangle inequality* holds, i.e., for every $x, y \in \mathbb{R}^n$ one has

$$\|x + y\|_p^p \le \|x\|_p^p + \|y\|_p^p.$$

Finally, for $p = 0$ and $x \in \mathbb{R}^n$, we introduce the notation

$$\|x\|_0 := \#\operatorname{supp}(x),$$

where $\operatorname{supp}(x) := \{j \in \{1, \dots, n\} : x_j \ne 0\}$ is the *support* of x. Hence, $\|x\|_0$ counts the number of nonzero entries of x. Note that in this case even the homogeneity is violated, since for $\alpha \ne 0$ we have $\|\alpha x\|_0 = \|x\|_0$.

For a symmetric positive definite matrix $A = A^{\mathrm{T}} \in \mathbb{R}^{n,n}$, we introduce the *energy inner product*

$$(u, v)_A := \langle u, Av\rangle$$

and the induced *energy norm*

$$\|x\|_A := \sqrt{(x, x)_A}.$$

Every symmetric positive definite matrix $A \in \mathbb{R}^{n,n}$ has a unique symmetric positive definite *square root* $B := A^{\frac{1}{2}}$, with $A = B^2 = B^{\mathrm{T}}B$ satisfying the relation [25]:

$$\|x\|_A = \|Bx\|_2.$$

2.2 Sparse Representation and Compressed Sensing

In this part we survey some recent results on sparse representations of functions based on the solution of underdetermined linear systems via ℓ_1-minimization. We also discuss the recently introduced concept of compressed sensing.

Definition 2.1 ([7])**.** *Let $\Phi \in \mathbb{R}^{m,n}$ with $m \le n$ and $k \in \{1, \dots, n\}$. The* k-restricted isometry constant *is the smallest number $\delta_k \ge 0$, such that*

$$(1 - \delta_k)\|x\|_2^2 \le \|\Phi x\|_2^2 \le (1 + \delta_k)\|x\|_2^2 \tag{3}$$

for all $x \in \mathbb{R}^n$ with $\|x\|_0 \le k$.

If Φ in Definition 2.1 is orthonormal, then clearly $\delta_k = 0$ for all k. Conversely, if the constant δ_k is close to 0 for some matrix Φ, every set of columns of Φ of cardinality less than or equal to k behaves like an orthonormal system. In the case that $0 \le \delta_k < 1$, we say that the matrix Φ has the *restricted isometry property.*

For $\Phi \in \mathbb{R}^{m,n}$ with $m \le n$, a vector of the form $b = \Phi x$ *represents* (encodes) the vector x in terms of the columns of Φ. To extract the information about x that b contains, we use a *decoder* $\Delta : \mathbb{R}^m \to \mathbb{R}^n$ which is a (not necessarily linear) mapping. Then $y := \Delta(b) = \Delta(\Phi x)$ is our approximation to x from the information given in b. In general, for a given b and matrix Φ, $\Delta(b)$ may not be unique and it could be a set of vectors. But here for simplicity we take one of them and deal with this vector only.

Let $\Sigma_k := \{z \in \mathbb{R}^n : \|z\|_0 \le k\}$ denote the vectors in $\mathbb{R}^n$ of support less than or equal to k. In the following we use the classical ℓ_p-norm, but also other norms are possible, see Theorem 2.2 below. We introduce the distance

$$\sigma_k(x)_p := \min_{z \in \Sigma_k} \|x - z\|_p,$$

and observe that for $x, z \in \mathbb{R}^n$ the following inequality holds:

$$\sigma_{2k}(x + z)_p \le \sigma_k(x)_p + \sigma_k(z)_p. \tag{4}$$

We have the following theorem.

Theorem 2.2 ([15])**.** *Consider a matrix $\Phi \in \mathbb{R}^{m,n}$ with $m \le n$, a value $k \in \{1, \dots, n\}$, and let $\mathcal{N} = \ker(\Phi)$. If there exists a constant C_0 independent of k and n such that*

$$\|\eta\|_p \le \tfrac{C_0}{2}\, \sigma_{2k}(\eta)_p, \quad \textit{for all } \eta \in \mathcal{N}, \tag{5}$$

then there exists a decoder Δ such that

$$\|x - \Delta(\Phi x)\|_p \le C_0\, \sigma_k(x)_p, \quad \textit{for all } x \in \mathbb{R}^n. \tag{6}$$

Conversely, if there exists a decoder Δ such that (6) *holds, then*

$$\|\eta\|_p \le C_0\, \sigma_{2k}(\eta)_p, \quad \textit{for all } \eta \in \mathcal{N}. \tag{7}$$

If we combine Theorem 2.2 for $p = 1$ with the restricted isometry property (3), then we have the following.

Theorem 2.3 ([15])**.** *Let $\Phi \in \mathbb{R}^{m,n}$, $m \le n$ and $k \in \{1, \dots, n\}$. Assume that Φ satisfies*

$$(1 - \delta_{3k})\|x\|_2^2 \le \|\Phi x\|_2^2 \le (1 + \delta_{3k})\|x\|_2^2$$

for all x with $\|x\|_0 \le 3k$, such that

$$\delta_{3k} \le \delta < \frac{(\sqrt{2} - 1)^2}{3}.$$

Define a decoder Δ for Φ via

$$\Delta(b) := \operatorname{argmin}\{\|x\|_1 : b = \Phi x\}.$$

Then

$$\|x - \Delta(\Phi x)\|_1 \le C_0\, \sigma_k(x)_1,$$

where

$$C_0 = \frac{2\sqrt{2} + 2 - (2\sqrt{2} - 2)\delta}{\sqrt{2} - 1 - (\sqrt{2} + 1)\delta}.$$

Theorem 2.3 shows that the ℓ_1-norm solution can be as good as the best k-term approximation. An analogous result is the following.

Theorem 2.4 ([7])**.** *Let $\Phi \in \mathbb{R}^{m,n}$, $m \le n$ and $k \in \{1,\dots,n\}$. Assume that Φ satisfies the restricted isometry property* (3) *with $3k$ and $4k$, such that $\delta_{3k} + 3\delta_{4k} < 2$. If*

$$\Delta(b) = \operatorname{argmin}\{\|x\|_1 \,:\, b = \Phi x\}.$$

then

$$\|x - \Delta(\Phi x)\|_2 \le C_k \, \frac{\sigma_k(x)_1}{\sqrt{k}},$$

where C_k is a constant only depending on δ_{4k}.

Besides the k-restricted isometry constant δ_k, a second quantity plays an important role in compressed sensing [22, 33, 34].

Definition 2.5. *Let $\Phi \in \mathbb{R}^{m,n}$ with $m \le n$ have unit norm columns, i.e., $\Phi = [\phi_1 \cdots \phi_n]$ with $\|\phi_i\|_2 = 1$, for $i = 1,\dots,n$. Then the* mutual incoherence *of the matrix Φ is defined by*

$$\mathcal{M}(\Phi) := \max_{i \ne j} |\langle \phi_i, \phi_j \rangle|.$$

The mutual incoherence $\mathcal{M}(\Phi)$ of a matrix Φ is related to the k-restricted isometry constant via

$$\delta_k \le (k-1)\mathcal{M}(\Phi).$$

The following Lemma shows how the mutual incoherence may be used to bound the norm of the encoded vector $b = \Phi x$.

Lemma 2.6. *Let $\Phi = [\phi_1 \cdots \phi_n] \in \mathbb{R}^{m,n}$ with $m \le n$ have unit norm columns. Then for every $x \in \mathbb{R}^n$ the inequality*

$$\|\Phi x\|_2^2 \le \big(1 - \mathcal{M}(\Phi)\big)\|x\|_2^2 + \mathcal{M}(\Phi)\|x\|_1^2.$$

holds.

Proof. The proof follows by the following (in)equalities.

$$\begin{aligned}\|\Phi x\|_2^2 &= \sum_{i=1}^n \sum_{j=1}^n x_i \, x_j \langle \phi_i, \phi_j \rangle = \|x\|_2^2 + \sum_{i \ne j} x_i \, x_j \langle \phi_i, \phi_j \rangle \\ &\le \|x\|_2^2 + \mathcal{M}(\Phi) \sum_{i \ne j} |x_i| \, |x_j| = \|x\|_2^2 + \mathcal{M}(\Phi)\big(\|x\|_1^2 - \|x\|_2^2\big).\end{aligned}$$

□

Lemma 2.6 states that $\|\Phi x\|_2^2$ is bounded by a convex combination of $\|x\|_2^2$ and $\|x\|_1^2$ with the mutual incoherence as a parameter.

Compressed Sensing (Compressive Sampling) refers to a problem of "efficient" recovery of an unknown vector $x \in \mathbb{R}^n$ from the *partial* information provided by linear measurements $\langle x, \phi_j \rangle, \phi_j \in \mathbb{R}^n, j = 1,\dots,m$. The goal in compressed sensing is to design an algorithm that approximates x from the information $b = (\langle x, \phi_1 \rangle, \dots, \langle x, \phi_m \rangle) \in \mathbb{R}^m$. Clearly the most important case is when the number of measurements m is much smaller than n. The crucial step for this to work, is to build a set of sensing vectors $\phi_j \in \mathbb{R}^n$, $j = 1,\dots,m$ that is "good" for the approximation of all vectors $x \in \mathbb{R}^n$. Clearly, the terms "economical" and "good" should be clarified in a mathematical setting of the problem.

A natural variant of this setting, and this is the approach that is discussed here, uses the concept of sparsity. The problem can then be stated as follows. For given integers $m \leq n$ we want to determine the largest sparsity $k(m,n)$ such that there exists a set of vectors $\phi_j \in \mathbb{R}^n, j = 1,\dots,m$ and an efficient decoder Δ, mapping b into $\mathbb{R}^n$ in such a way that for any x of sparsity $k(m,n)$ one has exact recovery $\Delta(x) = x$, (see [17]).

3 Sparse Representations of Solutions of PDEs

As discussed in the introduction, we want to use similar ideas as those used in compressed sensing in the context of the solution of partial differential equations.

3.1 General Setup

For a Hilbert space of functions $\mathbb{H} = \mathbb{H}(\Omega)$ on a domain $\Omega \subset \mathbb{R}^d$ with smooth boundary Γ, we denote by (f,g) the inner product of $f, g \in \mathbb{H}$ and by $\|f\|_{\mathbb{H}} := \sqrt{(f,f)}$ the induced norm.

A *generating system* or *dictionary* for $\mathbb{H}$ is a family $\{\phi_i\}_{i=1}^{\infty}$ of unit norm elements (i.e., $\|\phi_i\|_{\mathbb{H}} = 1$) in $\mathbb{H}$, such that finite linear combinations of the elements $\{\phi_i\}$ are dense in $\mathbb{H}$. A smallest possible dictionary is a *basis* of $\mathbb{H}$, while the other dictionaries are *redundant* families of elements. Elements of $\mathbb{H}$ do not have unique representations as a linear sums of redundant dictionary elements.

We will consider *elliptic boundary value problems*, see, e.g., [3], and we want to find the solution of

$$Lu = f \quad \text{in } \Omega,$$

for a differential operator L and homogeneous Dirichlet boundary conditions $u = 0$ on the boundary Γ of Ω.

To fix notation and explain our approach, we will lay out a finite element approach, where we assume that test and solution space $\mathbb{U} \subset \mathbb{H}$ are the same. In order to get a closer analogy between the linear algebra formulation and the function space formulation, we assume that we have a redundant dictionary $\mathcal{D} = \{\phi_i\}_{i=1}^{\infty}$, such that

$$\operatorname{span}\{\phi_j\}_{j=1}^{\infty} = \mathbb{U}.$$

The corresponding weak formulation for the PDE problem is to find $u \in \mathbb{U}$ such that

$$a(u,v) := (Lu, v) = (f, v), \quad \text{for all } v \in \mathbb{U}, \tag{8}$$

where $a(\cdot,\cdot)$ is a bilinear form.

Finitely expressing u, v in terms of the dictionary as

$$u = \sum_{i=1}^{\infty} u_i \, \phi_i \quad \text{and} \quad v = \sum_{i=1}^{\infty} v_i \, \phi_i,$$

we can write the problem in terms of infinite vectors and matrices as

$$\sum_{i,j=1}^{\infty} v_j \, a_{ij} \, u_i = (v^{\infty})^{\mathrm{T}} A^{\infty} u^{\infty} = (v^{\infty})^{\mathrm{T}} b^{\infty}, \tag{9}$$

Here, $A^\infty := [a_{ij}^\infty]$ is the *stiffness matrix*, with $a_{ij}^\infty := (L\phi_i, \phi_j)$, $i, j \in \mathbb{N}$, the right hand side b^∞ is defined by $b_i^\infty := (f, \phi_i)$, for $i \in \mathbb{N}$, and the coordinate vectors are

$$u^\infty := \begin{bmatrix} u_1 \\ u_2 \\ \vdots \end{bmatrix}, \quad v^\infty := \begin{bmatrix} v_1 \\ v_2 \\ \vdots \end{bmatrix}. \tag{10}$$

The weak solution then can be formulated as the (infinite) linear system

$$A^\infty u^\infty = b^\infty. \tag{11}$$

Note that if the ϕ_j's are not linearly independent, the infinite linear system (9) and hence (11) may not be solvable or may have many solutions.

3.2 Algorithmic Framework

Let us now consider finite dimensional subproblems of (11) by assuming that the dictionary $\mathcal{D}$ subsumes a *refinement* procedure, i.e., for each basis function ϕ_j there exists a set of refined basis functions in $\mathcal{D}$. In particular, we assume that there is a mapping of each basis function ϕ_i to the index set $\mathrm{ref}(i) \subset \mathbb{N}$ of refined basis functions. In a *hierarchical* refinement, every ϕ_i can be written as a linear combination of $\{\phi_j\}_{j\in\mathrm{ref}(i)}$.

In many practical applications, the refinement will arise from a geometric refinement, for instance, by subdividing some triangulation and corresponding basis functions. Furthermore, the refinement usually satisfies $\mathrm{ref}(j)\cap\mathrm{ref}(\ell) = \varnothing$ for $j \neq \ell$. For notational convenience, we define

$$\mathrm{ref}(S) := \bigcup_{j\in S} \mathrm{ref}(j),$$

for $S \subset \mathbb{N}$. With this notation, we obtain a sequence of index sets

$$T^0 := \{1\},\ T^1 = \mathrm{ref}(T^0),\ T^2 = \mathrm{ref}(T^1), \dots$$

Define $S^k := T^0 \cup \dots \cup T^k$ and denote the corresponding nested subspaces as

$$\mathbb{U}^0 \subset \mathbb{U}^1 \subset \mathbb{U}^2 \subset \dots \subset \mathbb{U} \qquad \text{with} \qquad \mathbb{U}^k := \mathrm{span}\{\phi_j\}_{j\in S^k}.$$

We will appropriately select subsets $R^k \subseteq C^k \subseteq S^k$, where R^k and C^k are seen as subsets of the rows and columns of A^∞, respectively. The corresponding submatrix is defined as follows.

$$A^k := A^\infty[R^k, C^k] := [(L\phi_i, \phi_j)]_{i\in R^k, j\in C^k}.$$

The corresponding right hand side is

$$b^k := b^\infty[R^k] := [(f, \phi_j)]_{j\in R^k}.$$

We thus arrive at the finite dimensional subsystem

$$A^k x^k = b^k, \tag{12}$$

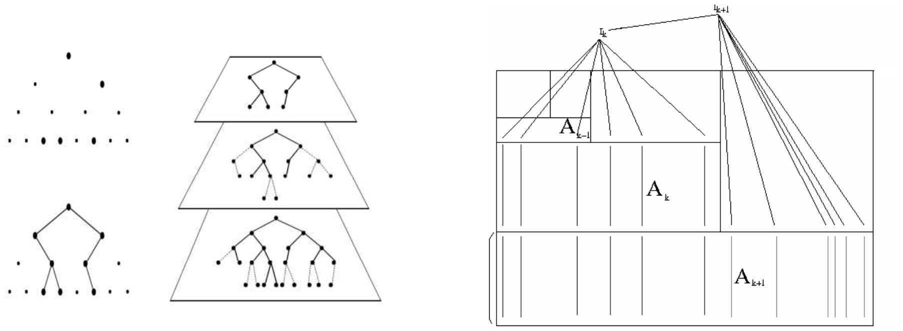

Figure 1: Illustration of the stepwise refinement method. *Left:* Picture of a tree that is iteratively refined. *Right:* A view on the matrix A^∞ and its partition as in (13).

and the approximate solution in this case is

$$u^k \approx \sum_{j \in C^k} x_j^k \, \phi_j.$$

As in classical adaptive methods the hope is to keep the size of the matrix A^k small (i.e., keep R^k and C^k small) and still obtain a good approximation of the solution arising from the full refinement of level k, i.e, a solution obtained from solving (12) for $C^k = R^k = S^k$.

We will now explain how compressed sensing can be used to select small R^k and C^k under the condition that we still obtain a convergent method. For this, assume that we have already selected R^k and C^k. We may start with $R^0 = C^0 = \{1\}$, but in practice one should choose an appropriately fine level. We now refine these sets to $\hat{R}^k := \mathrm{ref}(R^k)$ and $\hat{C}^k := \mathrm{ref}(C^k)$. Then A^k and b^k can be partitioned as follows

$$A^k = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}, \qquad b^k = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \tag{13}$$

where

$$\begin{aligned} A_{11} &= A^\infty[R^k, C^k] & A_{12} &= A^\infty[R^k, \hat{C}^k] & A_{13} &= A^\infty[R^k, \overline{C}^k] \\ A_{21} &= A^\infty[\hat{R}^k, C^k] & A_{22} &= A^\infty[\hat{R}^k, \hat{C}^k] & A_{23} &= A^\infty[\hat{R}^k, \overline{C}^k] \\ A_{31} &= A^\infty[\overline{R}^k, C^k] & A_{32} &= A^\infty[\overline{R}^k, \hat{C}^k] & A_{33} &= A^\infty[\overline{R}^k, \overline{C}^k], \end{aligned}$$

with $\overline{R}^k := \mathbb{N} \setminus (R^k \cup \hat{R}^k)$ and $\overline{C}^k$ defined analogously. Similarly, the right hand side is defined as

$$b_1 = b^\infty[R^k], \quad b_2 = b^\infty[\hat{R}^k], \quad b_3 = b^\infty[\overline{R}^k].$$

The main idea in the construction of R^{k+1} and C^{k+1} is to use the underdetermined subsystem

$$[A_{21} \; A_{22}] \, z = b_2,$$

and to compute a sparse solution, by taking a minimal ℓ_1-solution, i.e.,

$$z^k = \mathrm{argmin}\{\|z\|_1 \, : \, [A_{21} \; A_{22}] \, z = b_2\}.$$

Algorithm 1 Iteratively Refinement Basis Pursuit (IRBP)

Set $R^0 = C^0 = \{1\}$
for $i = 1, \ldots,$ until convergence **do**
  Construct $\hat{R}^k = \mathrm{ref}(R^k)$, $\hat{C}^k = \mathrm{ref}(C^k)$
  Construct $A_{21} = A^\infty[\hat{R}^k, C^k]$, $A_{22} = A^\infty[\hat{R}^k, \hat{C}^k]$, and $b_2 = b^\infty[\hat{R}^k]$.
  Solve the following minimization problem:

$$z^k = \operatorname{argmin}\{\|z\|_1 \,:\, [A_{21}\ A_{22}]z = b_2\}.$$

  Let $C \subseteq C^k \cup \hat{C}^k$ be the index set corresponding to the support of z^k.
  Set $C^{k+1} = C \cup C^k$, $R^{k+1} := C^{k+1}$.
end for

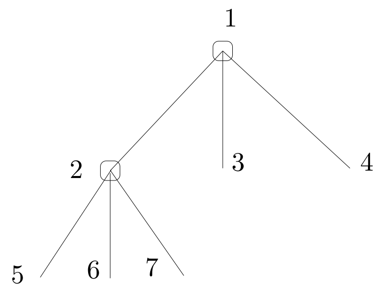

Figure 2: Illustration of Example 3.1.

The submatrix $[A_{21}\ A_{22}]$ is chosen because it combines the refined rows with the full set of columns that are available at the current iteration.

Now assume that $C \subseteq C^k \cup \hat{C}^k$ is the index set corresponding to the support of z^k as defined in Section 2.1. Then the new sets are set to

$$C^{k+1} = C \cup C^k, \qquad R^{k+1} = C^{k+1}.$$

Thus, the support of z^k is only used to select basis functions among $\hat{C}^k$ and the information in $C^k \cap C$ is not used.

Note that by construction

$$C^{k+1} \subseteq \mathrm{tree}(C),$$

where $\mathrm{tree}(C)$ is the set of basis functions on the path of a basis function to the root of the refinement-tree, i.e.,

$$\mathrm{tree}(j) := \{\ell \in \mathbb{N} \,:\, \exists j_1, \ldots, j_s \text{ with } j_1 \in \mathrm{ref}(\ell),\ j_2 \in \mathrm{ref}(j_1),\ \ldots,\ j \in \mathrm{ref}(j_s)\}.$$

The process is terminated if $\|z^{k+1} - z^k\|_2 \le \varepsilon$, where ε is a given tolerance. Since we are using ℓ_1-minimization, which in [10] is called basis pursuit, we call this process *Iteratively Refined Basis Pursuit.* The method is summarized as Algorithm 1. Figure 1 gives an illustration of the process.

Example 3.1. By definition, the first sets are $R^0 = C^0 = \{1\}$. Now suppose that $\mathrm{ref}(1) = \{2, 3, 4\}$, i.e, the initial rows and columns are $\hat{C}^0 = \hat{R}^0 = \{2, 3, 4\}$. We then solve the ℓ_1-minimization problem for the corresponding 3×4 matrix.

Now suppose that $\mathrm{supp}(z^0) = \{1,2\}$, then $C^1 = \{1,2\}$. Assume that $\mathrm{ref}(2) = \{5,6,7\}$, see Figure 2 for an illustration. Then $\hat{C}^1 = \{3,4,5,6,7\}$ and the next matrix is of size 5×7, since $R^1 = \{1,2\}$ and $\hat{R}^1 = \{3,4,5,6,7\}$.

Note that if the support of z^k is full, the above procedure yields a simple refinement process in which $R^k = S^{k-1}$ and $C^k = S^k$.

In general, this process need not converge. In fact, if at some level we see that the error does not decrease, then we back up in the tree of refined basis functions and refine at higher levels until we obtain a decrease in the error.

Remark 3.2. Note that although our description was based on the assumption that test and solution space are the same, the principle of the process does not depend on this assumption. Similar concepts for adaptive refinement methods in the context of wavelets are presented, e.g. in [1, 13, 14].

3.3 Properties of the Proposed Method

In our approach we want to achieve several goals. The solution z^k should be sparse, and z^k should be a good approximation of the solution x^{i+1} of (12). In order to analyze the behavior of the suggested approach, we study the case of two levels and assume that $R^k = C^k$. We will also slightly change notation as follows. For $k \in \mathbb{N}$, we set $n := \# R^k$ and $N := \# (R^k \cup \mathrm{ref}(R^{k+1}))$. We then introduce the following notation for submatrices of A^∞ and subvectors of b^∞ as in (13):

$$A^N = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad b^N = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \tag{14}$$

and $A^n := A_{11}$, $b^n := b_1$. Note that A^N is of size $N \times N$ and A^n of size $n \times n$. The corresponding linear systems are

$$A^\infty x^\infty = b^\infty, \tag{15}$$

$$A^N x^N = b^N, \tag{16}$$

and $A^n x^n = b^n$. In Algorithm 1, we take the last $N - n$ rows of (14) and consider the linear system

$$A^{N-n,N} z^N = b_2 \qquad \text{with} \qquad A^{N-n,N} := [A_{21}\, A_{22}]. \tag{17}$$

We find a solution of this underdetermined system, by taking a minimal ℓ_1-solution, i.e., by solving

$$z^N = \mathrm{argmin}\{\|z\|_1 \,:\, A^{N-n,N} z = b_2\}. \tag{18}$$

As a first step of the analysis we estimate the energy norm error between x^N and z^N in terms of the difference $\|x^N - x^\infty\|$. To derive such a bound, we need to embed z^N and x^n into $\mathbb{R}^\infty$ by appending 0 as follows:

$$\hat{z}^N = \begin{bmatrix} z^N \\ 0 \end{bmatrix}, \ \hat{x}^N = \begin{bmatrix} x^N \\ 0 \end{bmatrix}. \tag{19}$$

We assume that $A^\infty x^\infty = b^\infty$, $A^N x^N = b^N$, and that z^N is determined by (18).

We want to find necessary and sufficient conditions on the matrix A^∞ and thus on the dictionary $\{\phi_i\}$, such that there exists a constant $C_{\frac{n}{N}}$ for which the following inequality holds,

$$(\hat{z}^N - \hat{x}^N)^{\mathrm{T}} A^\infty (\hat{z}^N - \hat{x}^N) \leq C_{\frac{n}{N}} \, (x^\infty - \hat{x}^N)^{\mathrm{T}} A^\infty (x^\infty - \hat{x}^N). \qquad (20)$$

The constant $C_{\frac{n}{N}}$ should only depend on the ratio $\frac{n}{N}$. Considering the results of [9, 15, 22], we may expect that if the matrices A^∞, A^N have a small mutual incoherence or some restricted isometry property, such conditions can be obtained. We will come back to this point in Section 4.

If Inequality (20) holds, then by using the triangle inequality we obtain the error estimate

$$(x^\infty - \hat{z}^N)^{\mathrm{T}} A^\infty (x^\infty - \hat{z}^N) \leq (1 + C_{\frac{n}{N}})(x^\infty - \hat{x}^N)^{\mathrm{T}} A^\infty (x^\infty - \hat{x}_N),$$

which means that the (hopefully sparse) solution z^N obtained by solving (18) is as almost good as the solution of (16).

To estimate the errors between x^∞, x^N, and z^N, we need to consider a refined partition of A^∞ and b^∞ as defined in (14). In order to relate the solutions at different levels of refinement and the solution of the ℓ_1-minimization, we make use of the following Lemmas.

Lemma 3.3. *Let*

$$x^\infty = \begin{bmatrix} x_1^\infty \\ x_2^\infty \\ x_3^\infty \end{bmatrix}, \; x^N = \begin{bmatrix} x_1^N \\ x_2^N \end{bmatrix}, \text{and } z^N = \begin{bmatrix} z_1^N \\ z_2^N \end{bmatrix}$$

be solutions of the problems (15), (16), *and* (18), *respectively. Furthermore, let* $\hat{x}^N$ *and* $\hat{z}^N$ *be as in* (19). *Then inequality* (20) *can be rewritten as*

$$(x_1^N - z_1^N)^{\mathrm{T}} \big(b_1 - [A_{11}\, A_{12}] \begin{bmatrix} z_1^N \\ z_2^N \end{bmatrix} \big) \leq C_{\frac{n}{N}} \, (x_3^\infty)^{\mathrm{T}} \big(b_3 - [A_{31}\, A_{32}] \begin{bmatrix} x_1^N \\ x_2^N \end{bmatrix} \big).$$

Proof. Since

$$A^\infty x^\infty = b^\infty, \; A^\infty \hat{x}^N = \begin{bmatrix} b_1 \\ b_2 \\ [A_{31}\, A_{32}] x^N \end{bmatrix}, \; A^\infty \hat{z}^N = \begin{bmatrix} [A_{11}\, A_{12}] z^N \\ b_2 \\ [A_{31}\, A_{32}] z^N \end{bmatrix},$$

it follows that

$$A^\infty (\hat{x}^N - x^\infty) = \begin{bmatrix} 0 \\ 0 \\ [A_{31}\, A_{32}] z^N - b_3 \end{bmatrix}$$

and

$$A^\infty (\hat{z}^N - \hat{x}^N) = \begin{bmatrix} [A_{11}\, A_{12}] z^N - b_1 \\ 0 \\ [A_{31}\, A_{32}] (z^N - x^N) \end{bmatrix}.$$

Thus, we have

$$(\hat{x}^N - x^\infty)^{\mathrm{T}} A^\infty (\hat{x}^N - x^\infty) = (x_3^\infty)^{\mathrm{T}} \big(b_3 - [A_{31}\, A_{32}] \begin{bmatrix} x_1^N \\ x_2^N \end{bmatrix} \big)$$

and

$$(\hat{z}^N - \hat{x}^N)^{\mathrm{T}} A^\infty (\hat{z}^N - \hat{x}^N) = (x_1^N - z_1^N)^{\mathrm{T}} \big(b_1 - [A_{11}\, A_{12}] \begin{bmatrix} z_1^N \\ z_2^N \end{bmatrix} \big).$$

Plugging these expressions into (20) yields the claim. $\square$

Remark 3.4. A weaker version of (20) and of Lemma 3.3 will be given in Section 4.

The following Lemma gives a condition that has to be satisfied in order to guarantee that the refinement process can be iterated.

Lemma 3.5. *Let*

$$z^N = \begin{bmatrix} z_1^N \\ z_2^N \end{bmatrix}$$

be a solution of (18)*, where* $A^{N-n,N} = [A_{21}\, A_{22}]$ *is as defined in* (17)*, and suppose that* A_{22} *is invertible. If* $z_1^N \neq 0$*, then*

$$\|A_{22}^{-1} A_{21}\|_1 \geq 1. \tag{21}$$

Proof. Since

$$z' = \begin{bmatrix} 0 \\ A_{22}^{-1} b_2 \end{bmatrix}$$

is a feasible solution of (18), it follows that

$$\|z^N\|_1 \leq \|z'\|_1. \tag{22}$$

Moreover, from $A_{21} z_1^N + A_{22} z_2^N = b_2$, we obtain that $z_2^N = A_{22}^{-1} b_2 - A_{22}^{-1} A_{21} z_1^N$. Thus, using (22), we obtain

$$\begin{aligned}
\|A_{22}^{-1} b_2\|_1 &\geq \|z^N\|_1 \\
&= \|z_2^N\|_1 + \|z_1^N\|_1 \\
&= \|A_{22}^{-1} b_2 - A_{22}^{-1} A_{21} z_1^N\|_1 + \|z_1^N\|_1 \\
&\geq \|A_{22}^{-1} b_2\|_1 - \|A_{22}^{-1} A_{21} z_1^N\|_1 + \|z_1^N\|_1.
\end{aligned}$$

Since $\|A_{22}^{-1} A_{21} z_1^N\|_1 \leq \|A_{22}^{-1} A_{21}\|_1 \cdot \|z_1^N\|_1$, we have

$$\|z_1^N\|_1 \leq \|A_{22}^{-1} A_{21}\|_1 \cdot \|z_1^N\|_1,$$

which completes the proof. $\square$

Remark 3.6. Lemma 3.5 implies that a solution of the ℓ_1-minimization problem can only lead to an improvement if the matrix $[A_{21}\, A_{22}]$ satisfies (21). Otherwise, an optimal solution can already be obtained by solving the linear system $A_{22} z_2^N = b_2$. Another observation is that Lemma 3.5 remains true for any nonsingular principal submatrix of A^N.

In order to compare sparse and non-sparse solutions we introduce the short notation $s(x) := \operatorname{supp}(x)$ for a vector $x \in \mathbb{R}^m$ and $\overline{s}(x) = \{1, 2, \ldots, m\} \setminus s(x)$. For $x \in \mathbb{R}^m$ and $S \subset \{1, \ldots, m\}$, we denote

$$(y_S)_i = \begin{cases} y_i & \text{if } i \in S \\ 0 & \text{otherwise.} \end{cases} \qquad i = 1, \ldots, m.$$

We then have the following Lemma.

Lemma 3.7. *Let z^N be a solution of* (18)*, where $A^{N-n,N}$ is as in* (17)*, and let x^N be a solution of $A^N x^N = b^N$, with A^N as in* (14)*. Then for the difference $\delta^N := z^N - x^N$ we have the inequality*

$$\frac{\|\delta^N_{s(x^N)}\|_1}{\|\delta^N\|_1} \geq \frac{1}{2}.$$

Proof. Since x^N is a feasible solution of (18), we have

$$\|x^N + \delta^N\|_1 = \|z^N\|_1 \leq \|x^N\|_1. \tag{23}$$

Furthermore, $x^N + \delta^N = x^N + \delta^N_{s(x^N)} + \delta^N_{\overline{s}(x^N)}$. Therefore, by (23), we have

$$\begin{aligned}\|x^N\|_1 &\geq \|x^N + \delta^N\|_1 \\ &= \|x^N + \delta^N_{s(x_N)}\|_1 + \|\delta^N_{\overline{s}(x^N)}\|_1 \\ &\geq \|x^N\|_1 - \|\delta^N_{s(x^N)}\|_1 + \|\delta^N_{\overline{s}(x^N)}\|_1.\end{aligned}$$

Rewriting yields that

$$\|\delta^N_{s(x^N)}\|_1 \geq \|\delta^N_{\overline{s}(x^N)}\|_1 = \|\delta^N - \delta^N_{s(x^N)}\|_1 \geq \|\delta^N\|_1 - \|\delta^N_{s(x^N)}\|_1,$$

which implies the assertion. □

Remark 3.8. The proof of Lemma 3.7 shows that

$$\|(z^N - x^N)_{\overline{s}(x^N)}\|_1 \leq \|(z^N - x^N)_{s(x^N)}\|_1.$$

In particular, if $(z^N - x^N)_{s(x^N)} = 0$, then we conclude that $z^N = x^N$. If instead of ℓ_1-minimization, we use ℓ_0-minimization and compute

$$w^N = \operatorname{argmin}\{\|w\|_0 \,:\, A^{N-n,N} w = b_2\},$$

then we get the analogous estimate

$$\|(z^N - x^N)_{\overline{s}(x^N)}\|_0 \leq \|(z^N - x^N)_{s(x^N)}\|_0.$$

Remark 3.9. In general, it is not true that z^N and x^N satisfy the inequality $\|z_N\|_0 \leq \|x_N\|_0$. For example if

$$A^4 = \begin{bmatrix} 1 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 1 \end{bmatrix}$$

then

$$A^{3,4} = \begin{bmatrix} 0 & 1 & -\frac{1}{2} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 1 \end{bmatrix}$$

Now consider $b^4 = [\sqrt{2}, 6, -\frac{3}{2}, -1]^{\mathrm{T}}$ and $x^4 = [0, 7, 2, 0]^{\mathrm{T}}$. Then

$$z^4 = [\sqrt{2}, 6, 0, -1]^{\mathrm{T}},$$

and in this case $\|z_N\|_1 = 7 + \sqrt{2} < 7 + 2 = \|x_N\|_1$ and

$$\|z_N\|_0 = 3 > 2 = \|x_N\|_0.$$

In this section we have set the stage for the solution of PDEs via ℓ_1-minimization. In the next section we provide details.

4 The Restricted Isometry Property for Elliptic PDEs

In this section, we again discuss the special case that solution space and test space are the same, i.e., we assume $\mathbb{U} = \mathbb{V} \subset \mathbb{H}$ and we consider the symmetric bilinear form $a(\cdot,\cdot) : \mathbb{U} \times \mathbb{V} \to \mathbb{R}$ associated with the operator L as in (8).

We also assume that there exist constants $\alpha_1 > 0$, $\alpha_2 < \infty$, such that:

$$\alpha_1 \, \|u\|_{\mathbb{H}}^2 \leq a(u,u) \leq \alpha_2 \, \|u\|_{\mathbb{H}}^2, \tag{24}$$

i.e., $a(\cdot,\cdot)$ is *uniformly elliptic* with constant α_1 and *uniformly bounded* with constant α_2.

In order to connect this classical norm equivalence with the k-restricted isometry property, we assume that for the dictionary $\mathcal{D} = \{\phi_k\}_{i=1}^{\infty}$ the following *k-equivalence* between $\|\cdot\|_{\mathbb{H}}$ and the ℓ_2-norm $\|\cdot\|_2$ holds, i.e., we assume that there exist constants $\beta_1 > 0$, $\beta_2 < \infty$ with

$$\beta_1 \, \|u^\infty\|_2 \leq \Big\|\sum_{i=1}^{\infty} u_i \, \phi_i\Big\|_{\mathbb{H}} \leq \beta_2 \, \|u^\infty\|_2, \tag{25}$$

for all infinite vectors u^∞ as in (10) with the property that $\|u^\infty\|_0 \leq k$.

Note that inequality (24) can be written as

$$\alpha_1 \, \Big\|\sum_{i=1}^{\infty} u_i \, \phi_i\Big\|_{\mathbb{H}}^2 \leq \|A^{\infty\,\frac{1}{2}} \, u^\infty\|_2^2 \leq \alpha_2 \, \Big\|\sum_{i=1}^{\infty} u_i \, \phi_i\Big\|_{\mathbb{H}}^2,$$

or equivalently as

$$\alpha_1 \, \Big\|\sum_{i=1}^{\infty} u_i \, \phi_i\Big\|_{\mathbb{H}}^2 \leq (u^\infty)^{\mathrm{T}} A^\infty u^\infty \leq \alpha_2 \, \Big\|\sum_{i=1}^{\infty} u_i \, \phi_i\Big\|_{\mathbb{H}}^2. \tag{26}$$

Combining inequalities (25) and (26), we obtain

$$\alpha_1 \, \beta_1^2 \, \|u^\infty\|_2^2 \leq (u^\infty)^{\mathrm{T}} A^\infty u^\infty \leq \alpha_2 \, \beta_2^2 \, \|u^\infty\|_2^2,$$

for all vectors u^∞ with $\|u^\infty\|_0 \leq k$.

We consider the dictionary $\mathcal{D} = \{\phi_1, \phi_2, \dots\}$, and we choose a set $\mathcal{I}^N = \{q_1, \dots, q_N\} \subset \mathbb{N}$ with associated elements $\phi_{q_1}, \dots, \phi_{q_N} \in \mathcal{D}$. For the theoretical analysis we may assume w.l.o.g. that $\mathcal{I}^N = \{1, 2, \dots, N\}$. This selection can be obtained via an appropriate reordering of the elements ϕ_i of the dictionary.

The corresponding finite stiffness matrix associated with this subset is then $A^N = [a_{ij}] \in \mathbb{R}^{N,N}$ with $a_{ij} = a(\phi_i, \phi_j)$, $i, j = 1, \dots, N$. Since we have assumed uniform ellipticity and since test and solution space are equal, it follows that A^N is symmetric and positive semidefinite with $\operatorname{rank}(A^N) = N$; A^N is positive definite if $\phi_1, \dots, \phi_N$ form a basis.

Since A^N is symmetric positive semidefinite, A^N has a factorization $A^N = (B^N)^{\mathrm{T}} B^N$, where $B^N \in \mathbb{R}^{n,N}$ has full row rank. Hence, there exists a permutation matrix P such that

$$P \, B^N = \begin{bmatrix} B^n \\ B^{N-n} \end{bmatrix},$$

with $B^n \in \mathbb{R}^{n,n}$ invertible. This yields

$$P^{\mathrm{T}} A^N P = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

with stiffness matrix $A^n = A_{11}$. Then we have

$$A^N = (P\,B^N)^{\mathrm{T}} P\,B^N = \begin{bmatrix} B^n \\ B^{N-n} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} B^n \\ B^{N-n} \end{bmatrix} = (B^n)^{\mathrm{T}} B^n + (B^{N-n})^{\mathrm{T}} B^{N-n}.$$

Suppose that it is possible to choose the permutation matrix P in such a way that (measured in spectral norm)

$$\|B^{N-n}\|_2 \leq \epsilon$$

with a small $\epsilon > 0$, i.e.,

$$\|B^{N-n} x\|_2^2 \leq \epsilon\, \|x\|_2^2$$

for all $x \in \mathbb{R}^n$. Suppose further that

$$(1-\delta_k)\, \|x^N\|_2^2 \leq (x^N)^{\mathrm{T}} A^N x^N \leq (1+\delta_k)\, \|x^N\|_2^2$$

or equivalently

$$(1-\delta_k)\, \|x^N\|_2^2 \leq \|B^N x^N\|_2^2 \leq (1+\delta_k)\, \|x^N\|_2^2$$

for all x^N with $\|x^N\|_0 \leq k$.

To get an error estimate between the solution that is based on ℓ_1-minimization and the best k-term approximation, we first prove the following result.

Theorem 4.1. *Let $\mathcal{A} \in \mathbb{R}^{N,N}$ be symmetric positive semi-definite, and consider the solvable linear system $\mathcal{A}x = b$. Let $\mathcal{A} = B^{\mathrm{T}} B$ be a full rank factorization, and let $P \in \mathbb{R}^{N,N}$ be a permutation matrix such that the following properties hold.*

1. $P\,B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$*, and* $\mathcal{A} = B^{\mathrm{T}} B = B_1^{\mathrm{T}} B_1 + B_2^{\mathrm{T}} B_2$*;*

2. *for any solution x of $\mathcal{A}x = b$, $B_2^{\mathrm{T}} B_2\, x$ is k-sparse, i.e., $B_2^{\mathrm{T}} B_2\, x \in \Sigma_k$, where $\Sigma_k = \{z \,:\, \|z\|_0 \leq k\}$;*

3. $\|B_2 x\|_2^2 \leq C_{3,k}\, \sigma_k^2(x)_1$*, where*

$$\sigma_k(x)_1 = \min_{z \in \Sigma_k} \|z - x\|_1$$

   with a constant $C_{3,k}$ that only depends on k;

4. *the $4k$-restricted isometry constant δ_{4k} for B_1 is sufficiently small (for example $\delta_{4k} < \frac{1}{2}$).*

Then

$$(x - \tilde{x})^{\mathrm{T}} \mathcal{A}\, (x - \tilde{x}) \leq C_k\, \sigma_k^2(x)_1, \tag{27}$$

where $\tilde{x}$ is obtained via the solution of the minimization problems

$$\tilde{y} = \operatorname{argmin}_y \|b - B_1^{\mathrm{T}} y\|_1 \quad \text{and} \quad \tilde{x} = \operatorname{argmin}_x \{\|x\|_1 : B_1 x = \tilde{y}\}$$

and the constant C_k only depends on k and the mutual incoherence $\mathcal{M}(B_1)$.

Proof. By Assumption 1, $\mathcal{A}x = b$ implies that $b = B_1^{\mathrm{T}} B_1\, x + B_2^{\mathrm{T}} B_2\, x$. By Assumption 2, it follows that $e = B_2^{\mathrm{T}} B_2\, x$ is k-sparse. Then by Theorem 1.3 of [9] we obtain exact recovery, i.e., if $B_1\, x = \tilde{y}$, then

$$\tilde{y} = \mathrm{argmin}_y \|b - B_1^{\mathrm{T}} y\|_1.$$

The remainder of the proof is then based on Theorems 2.2, 2.4, and Lemma 2.6. Since

$$\mathcal{A} = B^{\mathrm{T}} B = B_1^{\mathrm{T}} B_1 + B_2^{\mathrm{T}} B_2,$$

it follows that

$$(x - \tilde{x})^{\mathrm{T}} \mathcal{A}(x - \tilde{x}) = \|B_1(x - \tilde{x})\|_2^2 + \|B_2(x - \tilde{x})\|_2^2, \tag{28}$$

where $\tilde{x} = \mathrm{argmin}_x \{\|x\|_1 : B_1 x = \tilde{y}\}$. By Assumption 3, we also have

$$\|B_2(x - \tilde{x})\|_2^2 \leq C_{3,k}\, \sigma_k^2(x)_1.$$

W.l.o.g. we can assume that B_1 has unit norm columns. Otherwise instead of the linear equation $B_1 x = \tilde{y}$, we can consider the following linear equation:

$$(B_1 S) S^{-1} x = \tilde{y},$$

where $S = \mathrm{diag}(\frac{1}{\|Be_i\|_2})$ and e_i is the i-th column of the identity matrix. Then the matrix $B_1 S$ has unit norm column and therefore, by Lemma 2.6 we have that

$$\|B_1(x - \tilde{x})\|_2^2 \leq (1 - \mathcal{M}(B_1))\, \|x - \tilde{x}\|_2^2 + \mathcal{M}(B_1)\, \|x - \tilde{x}\|_1^2.$$

By Theorem 2.4, we have $\|x - \tilde{x}\|_2^2 \leq \frac{C_{1,k}^2}{k}\, \sigma_k^2(x)_1$ and by Theorem 2.2 we have $\|x - \tilde{x}\|_1^2 \leq C_{2,k}^2\, \sigma_k^2(x)_1$, where $C_{1,k}$ and $C_{2,k}$ only depend on δ_k. Combining these inequalities with (28), we get

$$(x - \tilde{x})^{\mathrm{T}} \mathcal{A}(x - \tilde{x}) \leq \Big((1 - \mathcal{M}(B_1))\frac{C_{1,k}^2}{k} + \mathcal{M}(B_1)\, C_{2,k}^2 + C_{3,k}^2\Big)\sigma_k^2(x)_1.$$

This concludes the proof. □

Applying Theorem 4.1 to the matrix $\mathcal{A} = (B^{N-n,N})^{\mathrm{T}} B^{N-n,N}$, where matrix $B^{N-n,N} = [A_{21} A_{22}]$ as in (14), we obtain the corresponding estimate for the stiffness matrix.

Corollary 4.2. *Let $\mathcal{A}^N \in \mathbb{R}^{N,N}$ be a symmetric positive semidefinite matrix of rank $N - n$ and $\mathcal{A}^N x^N = b^N$. Let $\mathcal{A}^N = (B^{N-n,N})^{\mathrm{T}} B^{N-n,N}$ be a full rank factorization of $\mathcal{A}^N$, where $B^{N-n,N} = [A_{21} \quad A_{22}]$. If the $4k$ restricted isometry constant δ_{4k} for $B^{N-n,N}$ is sufficiently small (e.g. if $\delta_{4k} < \frac{1}{2}$), then*

$$(x^N - \hat{x})^{\mathrm{T}} \mathcal{A}^N (x^N - \hat{x}) \leq C_k \sigma_k^2(x^N)_1,$$

where $\hat{x}$ is obtained via the solution of the minimization problem

$$\hat{y} = \mathrm{argmin}_y \{\|b^N - {B^{N-n,N}}^{\mathrm{T}} y\|_1\}$$

and

$$\hat{x} = \mathrm{argmin}_z \{\|z\|_1 : B^{N-n,N} z = \hat{y}\},$$

and C_k only depends on k and $\mathcal{M}(B^{N-n,N})$.

Proof. Taking $B_1 = B^{N-n,N}$, $B_2 = 0$, the proof follows from Theorem 4.1. $\square$

Remark 4.3. Equation (27) gives an estimate on the solution of the ℓ_1-minimization problem. If we assume that $\sigma_k(x^N)_1 \leq C_N \sigma_k(x^\infty)_1$, which means that the best approximation of x^N is as good as the best approximation of x^∞, then Theorem 4.1 shows that the solution that we get from ℓ_1-minimization is as good as the best k-term approximation of x^∞, where x^∞ is the solution of original equation $A^\infty x^\infty = b^\infty$.

Remark 4.4. Equation (27) only gives a good bound, if we have

$$C_{2,k}^2 \, \mathcal{M}(B_n) \leq \frac{C_{1,k}^2}{k}(\mu_{\max}^2 - 1),$$

where $\mu_{\max}$ is the largest singular value of B^n. Otherwise we may use the direct estimate $\|B^n(x^N - \tilde{x})\|_2^2 \leq \|B^n\|_2^2 \|x^N - \tilde{x}\|_2^2$ and then apply Theorem 2.2.

5 Numerical Experiments

In this section we present some numerical examples.

5.1 Example: 1D-Poisson Equation

Let us first demonstrate that ℓ_1-minimization can successfully obtain a sparse solution. We consider the Poisson equation

$$-u'' = f \qquad \text{on} \qquad \Omega = (-1, 1),$$

with boundary conditions $u(-1) = 0 = u(1)$ and

$$f(x) = 2 \cdot \alpha^3 \Big(\frac{x + \frac{1}{2}}{1 + \alpha(x + \frac{1}{2})^2} + \frac{x}{1 + (\alpha \cdot x)^2} + \frac{x - \frac{1}{4}}{1 + \alpha(x - \frac{1}{4})^2} + \frac{x - \frac{1}{2}}{1 + \alpha^2(x - \frac{1}{2})^2} \Big),$$

where $\alpha := 100 \cdot \pi$. The exact solution of this problem is

$$\begin{aligned} u(x) = &\arctan(\alpha(x + \tfrac{1}{2})) + \arctan(-\alpha(x + \tfrac{1}{2})) + \arctan(\alpha \cdot x) \\ &+ \arctan(-\alpha \cdot x) + \arctan(\alpha(x - \tfrac{1}{4})) + \arctan(-\alpha(x - \tfrac{1}{4})) \\ &+ \arctan(\alpha(x - \tfrac{1}{2})) + \arctan(-\alpha(x - \tfrac{1}{2})). \end{aligned}$$

We apply a finite element method [3] and use classical *shape functions*

$$\phi(x) := \begin{cases} 1 - |x| & \text{if } -1 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{29}$$

The different refinement levels are given by

$$\phi_{k,\ell}(x) := 2^{-k/2} \, \phi\big(2^{k-1} \, (x + 1) - \ell\big), \; k \in \mathbb{N}, \; \ell = 1, \ldots, 2^{k-1}.$$

Here, the scaling factor $2^{-k/2}$ is used to make the diagonal elements of the stiffness matrix equal to 1. We then have

$$\phi_{k,\ell}(x) = 2^{-k/2} \begin{cases} 1 - |2^{k-1} \, (x + 1) - \ell| & \text{if } -1 + \frac{\ell - 1}{2^{k-1}} \leq x \leq -1 + \frac{\ell + 1}{2^{k-1}} \\ 0 & \text{otherwise,} \end{cases}$$

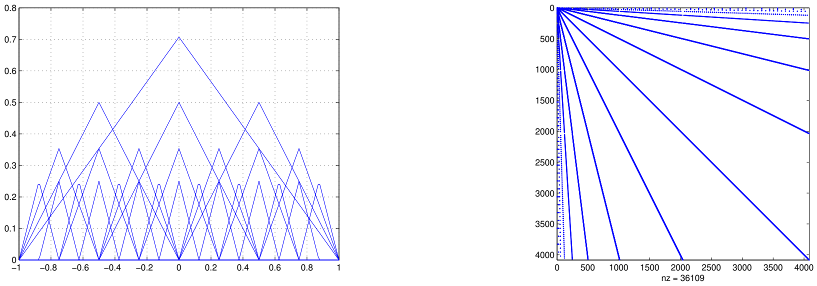

Figure 3: *Left:* four levels of hat functions on the interval $[-1, 1]$. *Right:* nonzeros in the corresponding stiffness matrix with 11 levels.

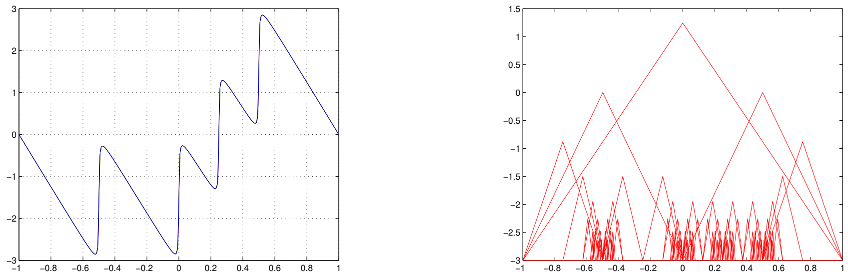

Figure 4: *Left:* exact solution and approximate solution of the example in Section 5.1 at level 8. *Right:* hat functions that are selected by the ℓ_1-minimization problem.

see the left part of Figure 3 for an illustration. On the Nth level, we then have

$$\sum_{k=1}^{N}(2^k - 1) = 2^{N+1} - (N + 2)$$

generating functions and we obtain a stiffness matrix that has a sparsity as depicted on the right part of Figure 3.

For this problem we have numerically computed the mutual incoherence and the restricted isometry constant. The numerical results indicate that for $A^{N-n,N}$ as in (17) we have

$$\mathcal{M}(A^{N-n,N}) = \sqrt{\tfrac{2}{3}},$$

independently of the size of the matrices, and that $\delta_1 = 0$, $\delta_2 = 0.8165$, and for all $k > 1$, we have $\delta_k > 1$.

On level 8 we used the matrix $A^{N-n,N}$ of size 255×502. The least squares solution of $A^N x = b$ has *495* nonzeros, while the minimum ℓ_1-solution only had *57* nonzeros. The left part of Figure 4 depicts the exact solution and the approximate solution at level 8. There is no obvious difference and the relative error in ℓ_2-norm is 0.0627. The right part of Figure 4 shows that our method refines properly at points with large gradients.

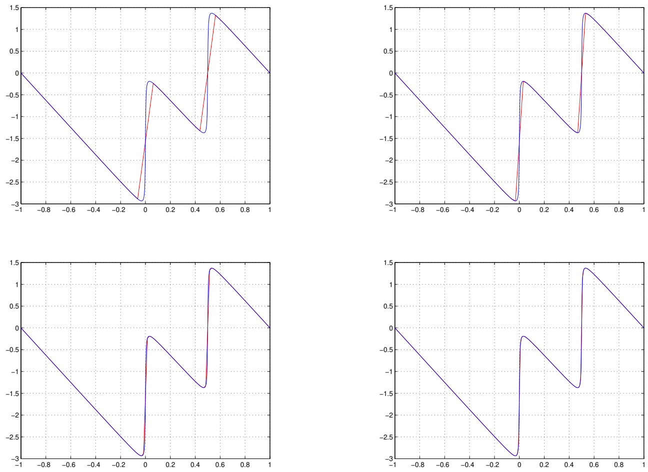

Figure 5: Exact and approximate solution in levels 5 to 8 for Example 5.2

Table 1: Results of Algorithm 1 for Example 5.2. The first column is the refinement step, column 2 gives the size of the matrix $A^{N-n,N}$ for the ℓ_1-minimization problem, column 3 gives the ℓ_0-norm of the least squares solution, column 4 gives the ℓ_0-norm of the ℓ_1-minimal solution, column 5 gives the ℓ_1-norm of the ℓ_1-minimal solution, and column 6 gives the relative error in ℓ_2-norm.

step	size of $A^{N-n,n}$	$\|x\|_0$	$\|z\|_0$	$\|z\|_1$	$\|z-x\|_2/\|x\|_2$
1	21 × 30	26	15	38.91	0.1732
2	39 × 48	57	23	61.74	0.1003
3	63 × 72	120	35	90.94	0.0501
4	99 × 108	247	51	124.21	0.0194

5.2 Application of Algorithm 1 to a 1D-Poisson Equation

To illustrate the behavior of Algorithm 1, we consider the Poisson equation

$$\begin{aligned} &Lu = -u'' = 2\tfrac{(100\pi)^3 x}{1+(100\pi x)^2} + 2\tfrac{(100\pi)^3(x-0.5)}{1+(100\pi(x-0.5))^2}, \quad x \in \Omega = (-1,1), \\ &(u(-1), u(1)) = (0,0). \end{aligned} \tag{30}$$

The exact solution of this problem is

$$\begin{aligned} u = &\arctan(100\,\pi \cdot x) + \arctan(-100\,\pi) \cdot x \\ &+ \arctan(100\,\pi(x - \tfrac{1}{2})) + \arctan(-100\,\pi)(x - \tfrac{1}{2}). \end{aligned}$$

We applied four refinement steps of Algorithm 1 starting as first step from level 5. In turns out that starting from level 4, a straightforward refinement process does not work, since some of the singularities are lost; this is a point where our algorithm would need to backtrack, see Section 3.2.

At each step we determine the new support using ℓ_1-minimization and then refine these nodes and all necessary higher level nodes according to Algorithm 1, see Figure 5. In Table 1 we present the results of four refinement steps of Algorithm 1.

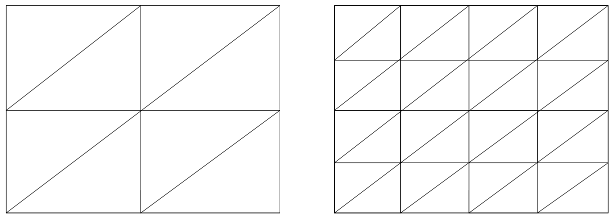

Figure 6: Triangulation on first and second level

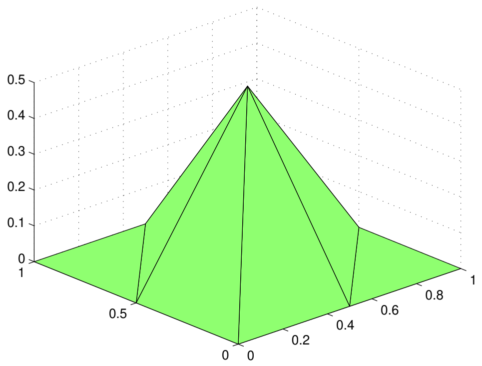

Figure 7: Basis function in the first level

5.3 Application of Algorithm 1 to a 2D-Poisson Equation

As a second example for Algorithm 1, we consider the Poisson equation

$$-\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = f(x,y) \text{ on } [0,1]\times[0,1]$$

with $u(x,y) = 0$ on the boundary of $[0,1]\times[0,1]$.

We use piecewise linear generating functions of the form $u_i(x,y) = a_i x + b_i y + c_i$ on each triangle. Figure 6 depicts the refinement step from the first level (left) to the second level (right). The basis function on level 1 is plotted in Figure 7.

Based on the triangulation on level j and the triangles $\{\Delta_{j,k_i}\}_{i=1}^{6}$, we obtain the generating functions:

$$\phi_{j,(a,b)}(x,y) = \frac{1}{2}\begin{cases} 2^j(x-a)+1 & (x,y)\in\Delta_{j,k_1} \\ 2^j(x-a)-2^j(y-b)+1 & (x,y)\in\Delta_{j,k_2} \\ -2^j(y-b)+1 & (x,y)\in\Delta_{j,k_3} \\ -2^j(x-a)+1 & (x,y)\in\Delta_{j,k_4} \\ -2^j(x-a)+2^j(y-b)+1 & (x,y)\in\Delta_{j,k_5} \\ 2^j(y-b)+1 & (x,y)\in\Delta_{j,k_6} \end{cases}$$

where (a,b) is the center of the basis function.

In general on level n we have $(2^n-1)^2$ basis functions. For this example the exact solution can be sparsely represented by exactly 29 basis functions on the

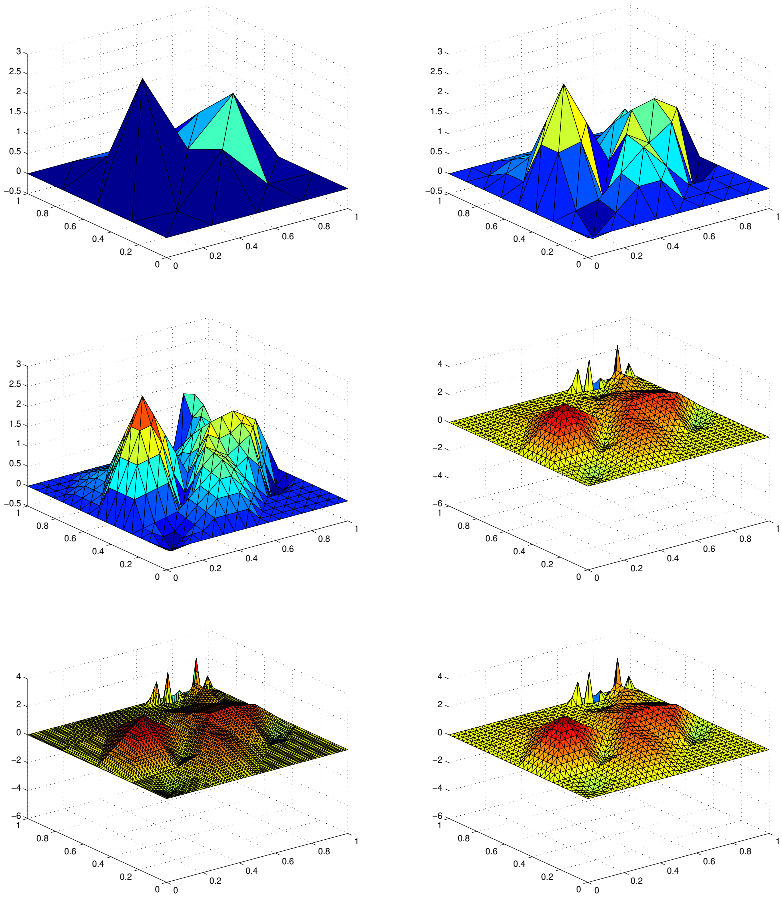

Figure 8: Approximate solutions obtained by Algorithm 1 on levels $2, 3, 4, 5, 6$ and exact solution (last) for Example 5.3.

Table 2: Results of Algorithm 1 for Example 5.3. The first column is the refinement step, column 2 gives the size of the matrix $A^{N-n,N}$, column 3 gives the ℓ_0-norm of the least squares solution, column 4 gives the ℓ_0-norm of the ℓ_1-minimal solution, column 5 gives the ℓ_1-norm of the ℓ_1-minimal solution, and column 6 gives the relative error in ℓ_2-norm.

step	size of $A^{N-n,n}$	$\|x\|_0$	$\|z\|_0$	$\|z\|_1$	$\|z-x\|_2/\|x\|_2$
1	9×10	10	9	14.6719	0.4123
2	49×58	58	25	23.5680	0.1231
3	158×165	161	53	33.8511	0.0146
4	326×333	312	29	57.0000	10^{-8}
5	393×399	345	29	57.0000	10^{-14}

5-th level. We applied four refinement steps of Algorithm 1 starting as first step from level 2. At each step we determine the new support using ℓ_1-minimization and then refine these nodes and all necessary higher level nodes according to Algorithm 1 (see Figure 8). In Table 2 we present the results of four refinement steps of Algorithm 1.

Conclusion

As mentioned in the introduction, many issues of the approach presented in this paper have not yet been resolved and many variations are possible. For instance, it is obvious that a similar approach could be derived using other dictionaries, e.g., wavelets, instead of finite element functions. Furthermore, for practical instances, the solution of the ℓ_1-minimization problem becomes an issue. One approach would be to apply different algorithms, for instance, Orthogonal Matching Pursuit, see [16, 30]. Moreover, the special structure of the stiffness matrices can be exploited and techniques adapted to the iterative procedure could be developed.

Acknowledgment

We thank O. Holtz, and J. Gagelman for many fruitful discussions and A. Jensen for help in the numerical experiments.

References

[1] R. Baraniuk, *Optimal tree approximation using wavelets*, in SPIE Technical Conference on Wavelet Applications in Signal Processing, Denver, CO, July 1999.

[2] R. Becker and R. Rannacher, *An optimal control approach to a posteriori error estimation in finite element methods*, in Acta Numerica, no. 10, Cambridge University Press, 2001, pp. 1–102.

[3] D. Braess, *Finite Elements: Theory, Fast Solvers and Applications in Solid Mechanics*, Cambridge, Cambridge University Press, 3rd ed., 2007.

[4] E. J. Candés, *Compressive sampling*, in Proc. International Congress of Mathematics, Madrid, Spain, 2006, pp. 1433–1452.

[5] E. J. Candès and J. Romberg, *Quantitative robust uncertainty principles and optimally sparse decompositions*, Found. Comput. Math. **6**, no. 2 (2006), pp. 227–254.

[6] E. J. Candès, J. Romberg, and T. Tao, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, IEEE Trans. Inform. Theory **52**, no. 2 (2006), pp. 489–509.

[7] E. J. Candès, J. Romberg, and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure Appl. Math. **59**, no. 8 (2006), pp. 1207–1223.

[8] E. J. Candès and T. Tao, *Decoding by linear programming*, IEEE Trans. Inform. Theory **51**, no. 12 (2005), pp. 4203–4215.

[9] E. J. Candès and T. Tao, *Near-optimal signal recovery from random projections: Universal encoding strategies*, IEEE Trans. Inform. Theory **52**, no. 12 (2006), pp. 5406–5425.

[10] S. S. Chen, *Basis Pursuit*, PhD thesis, Department of Statistics, Department of Statistics, Stanford,CA, 1995.

[11] S. S. Chen, D. L. Donoho, and M. A. Saunders, *Atomic decomposition by basis pursuit*, SIAM J. Sci. Comput. **20**, no. 1 (1999), pp. 33–61.

[12] O. Christensen, *An introduction to frames and Riesz bases*, Birkhäuser, Boston, 2003.

[13] A. Cohen, W. Dahmen, I. Daubechies, and R. DeVore, *Tree approximation and optimal encoding*, Appl. Comput. Harmonic Analysis **11**, no. 2 (2001), pp. 192–226.

[14] A. Cohen, W. Dahmen, and R. DeVore, *Adaptive wavelet methods for elliptic operator equations: convergence rates*, Math. Comput. **70**, no. 233 (2001), pp. 27–75.

[15] A. Cohen, W. Dahmen, and R. DeVore, *Compressed sensing and best k-term approximation*. Preprint, RWTH Aachen, 2006.

[16] G. Davis, S. Mallat, and Z. Zhang, *Adaptive time-frequency decompositions with matching pursuits*, Opt. Eng. **33**, no. 7 (1994), pp. 2183–2191.

[17] D. L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory **52**, no. 4 (2006), pp. 1289–1306.

[18] D. L. Donoho, *For most large underdetermined systems of equations, the minimal ℓ_1-norm near-solution approximates the sparsest near-solution*, Comm. Pure Appl. Math. **59**, no. 7 (2006), pp. 907–934.

[19] D. L. Donoho, *For most large underdetermined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution*, Comm. Pure Appl. Math. **59**, no. 6 (2006), pp. 797–829.

[20] D. L. Donoho and M. Elad, *Optimally sparse representation in general (non-orthogonal) dictionaries via ℓ^1 minimization*, Proc. Natl. Acad. Sci. USA **100**, no. 5 (2003), pp. 2197–2202.

[21] D. L. Donoho, M. Elad, and V. Temlyakov, *Stable recovery of sparse overcomplete representations in the presence of noise*, IEEE Trans. Inform. Theory **52**, no. 1 (2006), pp. 6–18.

[22] D. L. Donoho and X. Huo, *Uncertainty principles and ideal atomic decomposition*, IEEE Trans. Inform. Theory **47**, no. 7 (2001), pp. 2845–2862.

[23] D. L. Donoho, Y. Tsaig, I. Drori, and J.-L. Starck, *Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit*, Tech. Report 2006-02, Standford, Department of Statistics, 2006.

[24] M. R. Garey and D. S. Johnson, *Computers and Intractability. A Guide to the Theory of NP-Completeness*, W. H. Freeman and Company, New York, 1979.

[25] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, New York, NY, USA, 1986.

[26] N. Karmarkar, *A new polynomial-time algorithm for linear programming*, Combinatorica **4**, no. 4 (1984), pp. 373–395.

[27] L. G. Khachiyan, *A polynomial algorithm in linear programming*, Soviet Math. Dokl. **20** (1979), pp. 191–194.

[28] S. Kunis and H. Rauhut, *Random sampling of sparse trigonometric polynomials II – orthogonal matching pursuit versus basis pursuit.* Foundations of Computational Mathematics, to appear, 2007.

[29] B. K. Natarajan, *Sparse approximate solutions to linear systems*, SIAM J. Comput. **24**, no. 2 (1995), pp. 227–234.

[30] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, *Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decompositions*, in Proc. 27th Asilomar Conference on Signals, Systems and Computers, A. Singh, ed., 1993.

[31] H. Rauhut, *Stability results for random sampling of sparse trigonometric polynomials*. Preprint, arXiv:math/0609630v2, 2006.

[32] H. Rauhut, *Random sampling of sparse trigonometric polynomials*, Appl. Comput. Harmonic Analysis **22**, no. 1 (2007), pp. 16–42.

[33] J. A. Tropp, *Greed is good: algorithmic results for sparse approximation*, IEEE Trans. Inform. Theory **50**, no. 10 (2004), pp. 2231–2242.

[34] J. A. Tropp, *Just relax: convex programming methods for identifying sparse signals in noise*, IEEE Trans. Inform. Theory **52**, no. 3 (2006), pp. 1030–1051.