On the Uniqueness of Nonnegative Sparse Solutions to Underdetermined Systems of Equations *

Alfred M. Bruckstein, Michael Elad and Michael Zibulevsky†

July 20th, 2008

Abstract

An underdetermined linear system of equations $\mathbf{Ax} = \mathbf{b}$ with non-negativity constraint $\mathbf{x} \geq 0$ is considered. It is shown that for matrices $\mathbf{A}$ with a row-span intersecting the positive orthant, if this problem admits a sufficiently sparse solution, it is necessarily unique. The bound on the required sparsity depends on a coherence property of the matrix $\mathbf{A}$. This coherence measure can be improved by applying a conditioning stage on $\mathbf{A}$, thereby strengthening the claimed result. The obtained uniqueness theorem relies on an extended theoretical analysis of the $\ell_0 - \ell_1$ equivalence developed here as well, considering a matrix $\mathbf{A}$ with arbitrary column norms, and an arbitrary monotone element-wise concave penalty replacing the ℓ_1-norm objective function. Finally, from a numerical point of view, a greedy algorithm – a variant of the matching pursuit – is presented, such that it is guaranteed to find this sparse solution. It is further shown how this algorithm can benefit from well-designed conditioning of $\mathbf{A}$.

Keywords: Basis pursuit, greedy algorithm, linear system, positive orthant, sparse solution, uniqueness, ℓ_1 .

1 Introduction

This paper is devoted to the theoretical analysis of underdetermined linear system of equations of the form $\mathbf{Ax} = \mathbf{b}$ ($\mathbf{A} \in \mathbb{R}^{n \times k}$ with $k > n$ and $\mathbf{b} \in \mathbb{R}^n$) with non-negativity constraint $\mathbf{x} \geq 0$. Such problems are frequently encountered in signal and image processing, in handling of multi-spectral data, considering non-negative factorization for recognition, and more

*This research was supported in part by Israel Science Foundation Grant No. 796/05, by E. and J. Bishop Research Fund and by New York Metropolitan Research Fund.

†Department of Computer Science, Technion–Israel Institute of Technology, Technion City, Haifa 32000, Israel, Email: [elad,freddy,mzib]@cs.technion.ac.il

(see [2, 24, 27, 18, 28, 31] for representative work). In this paper we do not dwell on the applicative side of this problem and instead concentrate on the theoretical behavior of such systems.

When considering an under-determined linear system (i.e. $k > n$), with a full rank matrix $\mathbf{A}$, the removal of the non-negativity requirement $\mathbf{x} \geq 0$ leads to an infinite set of feasible solutions. How is this set reduced when we further require a non-negative solution? How can solutions be effectively found in practice? Assuming there could be several possible solutions, the common practice is the definition of an optimization problem of the form

$$(P_f): \qquad \min_{\mathbf{x}} \; f(\mathbf{x}) \;\; \text{subject to} \;\; \mathbf{A}\mathbf{x} = \mathbf{b} \text{ and } \mathbf{x} \geq 0, \tag{1-1}$$

where $f(\cdot)$ measures the quality of the candidate solutions. Possible choices for this penalty could be various entropy measures, or general ℓ_p-norms for various p in the range $[0, \infty)$. Popular choices are $p = 2$ (minimum ℓ_2-norm), $p = 1$ (minimum ℓ_1-norm), and $p = 0$ (enforcing sparsity). For example, recent work reported in [11, 12, 13] proves that the $p = 0$ and $p = 1$ choices lead to the same result, provided that this result is sparse enough. This work also provides bounds on the required sparsity that guarantee such equivalence.

Clearly, if the set of feasible solutions $\{\mathbf{x} | \; \mathbf{A}\mathbf{x} = \mathbf{b} \text{ and } \mathbf{x} \geq 0\}$ contains only one element, then all the above choices of $f(\cdot)$ will lead to the same solution. In such a case, the above-discussed ℓ_0-ℓ_1 equivalence becomes an example of a much wider phenomenon. Surprisingly, this is exactly what happens when a sufficiently sparse solution exists, and when one considers matrices $\mathbf{A}$ with a row-span intersecting the positive orthant. The main result shown of this paper proves the uniqueness of such a sparse solution, and provides a bound on $\|\mathbf{x}\|_0$ below which such a solution is guaranteed to be unique.

There are several known results reporting an interesting behavior of sparse solutions of a general under-determined linear system of equations, when minimum of ℓ_1-norm is imposed on the solution (this is the Basis Pursuit algorithm) [6, 17]. These results assume that the columns of the coefficient matrix have a unit ℓ_2-norm, stating that the minimal ℓ_1-norm solution coincides with the sparsest one for sparse enough solutions. As mentioned above, a similar claim is made in [11, 12, 13] for non-negative solutions, leading to stronger bounds.

In this work we extend the basis pursuit analysis, presented in [6, 17], to the case of a matrix with arbitrary column norms and an arbitrary monotone element-wise concave penalty replacing the ℓ_1-norm objective function. A generalized theorem of the same flavor is obtained. Using this result, we get conditions of uniqueness of sparse solution of non-negative system of equations, as mentioned above. Interestingly, there is no need in a sparsifying measure such as the ℓ_1 penalty – a non-negativity constraint is sufficient to lead to the unique (and sparsest) solution in these cases.

The bound on the required sparsity for guaranteed uniqueness depends on a coherence property of the matrix $\mathbf{A}$ that undergoes a conditioning stage. This conditioning allows for a left multiplication by any invertible matrix, and a right multiplication by a diagonal positive matrix of specific nature. These give some degrees of freedom in improving the coherence measure and strengthening the uniqueness claim. We demonstrate this property, and present some preliminary ways to exploit it.

Returning to the practical side of things, and assuming that we are interested in the sparsest (and possibly the only) feasible solution,

$$(P_0^+): \qquad \min_{\mathbf{x}} \ \|\mathbf{x}\|_0 \ \text{ subject to } \ \mathbf{A}\mathbf{x} = \mathbf{b} \text{ and } \mathbf{x} \geq 0, \tag{1-2}$$

there are several possible numerical methods for solving this problem. In this paper we present a variant of the orthogonal matching pursuit (OMP) for this task. We provide a theoretical analysis of this algorithm that follows the one shown in [29, 9] and shows that it is guaranteed to lead to the desired solution, if it is indeed sparse enough.

The structure of this paper is as follows: In Section 2 we extend the basis pursuit analysis to the case of arbitrary monotone element-wise concave penalty and matrix $\mathbf{A}$ with arbitrary column norms. This analysis relies on a special definition of coherence measure of the matrix $\mathbf{A}$. In Section 3 we develop the main theoretical result in this paper, claiming that a sufficiently sparse solution of $\{\mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$ is unique. We also introduce a conditioning stage that improves the coherence of the involved matrix, and thus strengthen the uniqueness claim. Section 4 presents the OMP variant for the non-negative problem, along with empirical and theoretical analysis of its performance.

2 Basis Pursuit: An Extended Result

2.1 General

In this section we develop a theorem claiming that a sufficiently sparse solution of a general under-determined linear system[1] $\mathbf{Dz} = \mathbf{b}$ is necessarily a minimizer of a separable concave function. It extends previous results from [6, 19, 16] in the following ways:

- It does not assume normalization of the columns in $\mathbf{D}$. Note that the work reported in [16] also allows for varying norms by pre-weighting the ℓ_1 measure – here our approach is different because of the following differences;
- It relies on a different feature of the matrix $\mathbf{D}$ – a one-sided coherence measure. As we shall see in Section 4, this generally implies a weaker bound;
- The objective function is more general than the ℓ_1-norm used in [6]. In fact, it is similar to the one proposed by Gribonval and Nielsen in [19], but due to the above changes, the analysis is rather different.

The result presented in this section constitutes a moderate contribution over the above-mentioned literature, and its importance is mainly in serving as grounds for the analysis of the non-negativity constraint that follows in Section 3.

2.2 The One-Sided Coherence and Its Use

For an arbitrary $n \times k$ matrix $\mathbf{D}$ with columns $\mathbf{d}_i$ we define its one-sided coherence as

$$\rho(\mathbf{D}) = \max_{i,j;j\neq i} \frac{|\mathbf{d}_i^T \mathbf{d}_j|}{\|\mathbf{d}_i\|_2^2}. \tag{2-1}$$

Defining the Gram matrix $\mathbf{G} = \mathbf{D}^T\mathbf{D}$, its elements satisfy

$$\frac{|G_{ij}|}{G_{ii}} \leq \rho(\mathbf{D}) \quad \forall i, j \neq i. \tag{2-2}$$

[1]In this Section we introduce a different notation for the linear system: $\mathbf{Dz} = \mathbf{b}$, instead of $\mathbf{Ax} = \mathbf{b}$. The reason for this change will be clarified in Section 3.

Note that the coherence measure used extensively in past work for the analysis of the matching and the the basis pursuit algorithms [8, 15, 6, 17, 29, 9, 30] is different, defined as a two-sided expression of the form[2]

$$\mu(\mathbf{D}) = \max_{i,j;j\neq i} \frac{|\mathbf{d}_i^T \mathbf{d}_j|}{\|\mathbf{d}_i\|_2 \|\mathbf{d}_j\|_2}. \tag{2-3}$$

The relation between the two is given in the following Lemma.

Lemma 1 *For any $\mathbf{D} \in \mathbb{R}^{n\times k}$ the inequality $\rho(\mathbf{D}) \geq \mu(\mathbf{D})$ holds true.*

Proof: Assume that the pair of columns i_0, j_0 are those leading to the two-sided coherence value μ. Thus we have

$$\mu(\mathbf{D}) = \frac{\left|\mathbf{d}_{i_0}^T \mathbf{d}_{j_0}\right|}{\|\mathbf{d}_{i_0}\|_2 \cdot \|\mathbf{d}_{j_0}\|_2}. \tag{2-4}$$

Assuming with no loss of generality that $\|\mathbf{d}_{i_0}\|_2 \geq \|\mathbf{d}_{j_0}\|_2$, we have

$$\mu(\mathbf{D}) \quad = \quad \frac{\left|\mathbf{d}_{i_0}^T \mathbf{d}_{j_0}\right|}{\|\mathbf{d}_{i_0}\|_2 \cdot \|\mathbf{d}_{j_0}\|_2} \leq \frac{\left|\mathbf{d}_{i_0}^T \mathbf{d}_{j_0}\right|}{\|\mathbf{d}_{i_0}\|_2 \cdot \|\mathbf{d}_{j_0}\|_2} \cdot \frac{\|\mathbf{d}_{i_0}\|_2}{\|\mathbf{d}_{j_0}\|_2} = \frac{\left|\mathbf{d}_{i_0}^T \mathbf{d}_{j_0}\right|}{\|\mathbf{d}_{j_0}\|_2^2} \leq \rho_c(\mathbf{D}), \tag{2-5}$$

as claimed. □

Despite the fact that the two-sided coherence is smaller (and thus better, as we shall see shortly), the analysis in this paper calls for the one-sided version. We start our analysis with characterizing the null-space of $\mathbf{D}$ using this measure.

Lemma 2 *Any vector δ from the null-space of $\mathbf{D}$ satisfies*

$$\|\delta\|_\infty \leq \frac{\rho(\mathbf{D})}{1+\rho(\mathbf{D})}\|\delta\|_1 = t_{\mathbf{D}}\|\delta\|_1, \tag{2-6}$$

where $t_{\mathbf{D}}$ is defined as $t_{\mathbf{D}} = \frac{\rho(\mathbf{D})}{1+\rho(\mathbf{D})}$.

Proof: Multiplying the null-space condition $\mathbf{D}\delta = 0$ by $\mathbf{D}^T$, and using $\mathbf{G} = \mathbf{D}^T\mathbf{D}$, we get $\mathbf{G}\delta = 0$. The i-th row of this equation,

$$G_{ii}\delta_i + \sum_{j\neq i} G_{ij}\delta_j = 0, \quad i = 1, 2, \ \ldots \ , k, \tag{2-7}$$

[2]This definition is for matrices with arbitrary column norms. In the case of ℓ_2 normalized columns, as assumed in the above-mentioned citation, the denominator simply vanishes.

gives us

$$\delta_i = -\sum_{j\neq i} \frac{G_{ij}}{G_{ii}}\delta_j, \quad i = 1, 2, \ \dots \ , k. \tag{2-8}$$

Taking absolute value of both sides, we obtain

$$|\delta_i| = \left|\sum_{j\neq i} \frac{G_{ij}}{G_{ii}}\delta_j\right| \leq \sum_{j\neq i} \left|\frac{G_{ij}}{G_{ii}}\right| |\delta_j| \leq \rho(\mathbf{D}) \sum_{j\neq i} |\delta_j|, \tag{2-9}$$

where the last inequality is due to (2-2). Adding a term $\rho(\mathbf{D})|\delta_i|$ to both sides, we obtain

$$(1 + \rho(\mathbf{D}))|\delta_i| \leq \rho(\mathbf{D}) \sum_{j=1}^{k} |\delta_j| = \rho(\mathbf{D})\|\delta\|_1, \quad i = 1, 2, \ \dots \ , k, \tag{2-10}$$

implying

$$|\delta_i| \leq \frac{\rho(\mathbf{D})}{1 + \rho(\mathbf{D})}\|\delta\|_1 = t_{\mathbf{D}}\|\delta\|_1, \quad i = 1, 2, \ \dots \ , k. \tag{2-11}$$

Thus, $\|\delta\|_\infty \leq t_{\mathbf{D}}\|\delta\|_1$, as claimed. $\square$

2.3 Sufficient Sparsity Guarantees Unique Global Optimality

Theorem 1 *Consider the following optimization problem:*

$$\min_{\mathbf{z}} \sum_i \varphi(|z_i|), \quad s.t. \quad \mathbf{Dz} = \mathbf{b}, \tag{2-12}$$

with a concave semi-monotone increasing function[3] $\varphi(\cdot)$. *A sparse (with* T *non-zeros) feasible solution* $\bar{\mathbf{z}}$ *(i.e.* $\mathbf{D}\bar{\mathbf{z}} = \mathbf{b}$*) is a unique global optimum of the above optimization problem, if*

$$T \equiv \|\bar{\mathbf{z}}\|_0 < \frac{1}{2t_D}, \tag{2-13}$$

where t_D *is given by* (2-6)*.*

[3]We assume w.l.o.g. that $\varphi(0) = 0$. Also, $\forall t > 0$ we assume the following properties: (i) Non-triviality - $\varphi(t) > 0$; (ii) Monotonicity - $\varphi'(t) \geq 0$; and (iii) Concavity - $\varphi''(t) \leq 0$. While these conditions are stated in terms of derivatives, they can be easily replaced with more general statements that avoid continuity assumption.

We shall provide a brief and shortened proof of this theorem below, as it is quite similar to proofs found in [6, 19].

Proof: We are going to show that under the conditions of the theorem, any feasible non-zero perturbation $\delta \in \mathbb{R}^k$ (such that $\mathbf{D}(\bar{\mathbf{z}} + \delta) = \mathbf{b}$) increases the objective function,

$$\sum_i \varphi(|\bar{z}_i + \delta_i|) > \sum_i \varphi(|z_i|). \tag{2-14}$$

We start from the observation that $\mathbf{D}\delta = \mathbf{0}$, and therefore by Lemma 2 we have

$$|\delta_i| \leq t_{\mathbf{D}} \|\delta\|_1 \equiv \delta_{\text{tol}}, \quad i = 1, ..., k. \tag{2-15}$$

Taking into account monotonicity of $\varphi(\cdot)$, decrease of the objective function is possible when the non-zero elements $|\bar{z}_i|$ are reduced. Maximal reduction takes place when $|\bar{z}_i|$ is decreased by the maximal possible value: δ_{tol} or to zero if $|\bar{z}_i| < \delta_{\text{tol}}$. Due to concavity of φ, the function decrease will be larger when $|\bar{z}_i| = \delta_{\text{tol}}$, comparing to $|\bar{z}_i| > \delta_{\text{tol}}$, because in the latter case reduction of the argument will fall into the area of lower slope of φ. With T nonzeros in $\bar{\mathbf{z}}$, the overall possible decrease of the objective function is therefore

$$\text{Amount of Maximal Decrease} = T\varphi(\delta_{\text{tol}}). \tag{2-16}$$

The remaining elements of δ, which correspond to the positions of zeros in $\bar{\mathbf{z}}$, will cause increase in the objective function. The minimal total increase happens if the remaining total amplitude

$$\|\delta\|_1 - T\delta_{\text{tol}} = \|\delta\|_1 - Tt_{\mathbf{D}}\|\delta\|_1 \tag{2-17}$$

falls to the area of possibly lowest slope of φ. Due to the concavity of φ, this will happen if we assign the maximum possible amplitude $\delta_i = \delta_{\text{tol}} = t_{\mathbf{D}}\|\delta\|_1$ to each element until the remaining total amplitude is gone, avoiding creation of small elements with high slope of φ. Thus, the number of these elements is

$$\frac{\|\delta\|_1 - Tt_{\mathbf{D}}\|\delta\|_1}{t_{\mathbf{D}}\|\delta\|_1} = \frac{1 - Tt_{\mathbf{D}}}{t_{\mathbf{D}}} \tag{2-18}$$

leading to the increase of the objective function by

$$\text{Amount of Minimal Increase} = \frac{1 - Tt_{\mathbf{D}}}{t_{\mathbf{D}}}\varphi(\delta_{\text{tol}}). \tag{2-19}$$

In order for $\bar{\mathbf{z}}$ to be a unique global minimizer, the change of the objective function should be positive, i.e. by (2-16) and (2-19) we get the condition

$$T\varphi(\delta_{\text{tol}}) < \frac{1 - Tt_{\mathbf{D}}}{t_{\mathbf{D}}}\varphi(\delta_{\text{tol}}), \tag{2-20}$$

which is satisfied for $T < \frac{1}{2t_D}$ as claimed. □

3 Introducing the Non-Negativity Constraint

We now turn to the main result of this paper, showing that if a system of equations $\mathbf{Ax} = \mathbf{b}$ with non-negativity constraint has a sufficiently sparse solution, then it is a unique one. This claim is shown to be true for a specific class of matrices $\mathbf{A}$: such that their row-span intersects the positive orthant. In this section we also show how to better condition the linear system in order to strengthen this uniqueness theorem.

3.1 Canonization of the System

We start with a system of linear of equations with non-negativity constraint

$$\mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq 0. \tag{3-1}$$

Suppose that the span of the rows of $\mathbf{A}$ intersects the positive orthant, i.e. a *strictly* positive vector $\mathbf{w}$ can be obtained as a linear combination of the rows of $\mathbf{A}$:

$$\exists\ \mathbf{h} \quad \text{s.t.} \quad \mathbf{h}^T\mathbf{A} = \mathbf{w}^T > 0. \tag{3-2}$$

We refer to this class of matrices hereafter as $\mathcal{M}^+$. This class includes all the purely positive or negative matrices, any matrix with at least one strictly positive or negative row, and more.

Assuming that $\mathbf{A} \in \mathcal{M}^+$, we can find a suitable $\mathbf{h}$ (there could be infinitely many in general) and a corresponding $\mathbf{w}$. Define the diagonal and strictly positive definite matrix $\mathbf{W} = \text{diag}(\mathbf{w})$. By changing the variables $\mathbf{z} = \mathbf{Wx}$ we obtain an equivalent system of the form

$$\mathbf{DW}^{-1}\mathbf{Wx} = \mathbf{Dz} = \mathbf{b}, \quad \mathbf{z} \geq 0, \tag{3-3}$$

where $\mathbf{D} = \mathbf{A}\mathbf{W}^{-1}$. There is a one-to-one correspondence between the solution sets

$$\{\mathbf{x} \mid \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq 0\} \quad \text{and} \quad \{\mathbf{z} \mid \mathbf{D}\mathbf{z} = \mathbf{b},\ \mathbf{z} \geq 0\}, \tag{3-4}$$

and the cardinalities of the corresponding solutions are equal. An interesting and important property of the new system $\mathbf{D}\mathbf{z} = \mathbf{b}$ is the following: Multiplying Equation (3-1) by $\mathbf{h}^T$, we have

$$\mathbf{h}^T\mathbf{A}\mathbf{x} = \mathbf{w}^T\mathbf{x} = c, \tag{3-5}$$

where $c = \mathbf{h}^T\mathbf{b}$. This implies

$$\mathbf{1}^T\mathbf{W}\mathbf{x} = \mathbf{1}^T\mathbf{z} = c, \tag{3-6}$$

i.e. the sum of the entries of *any* solution $\mathbf{z}$ in $\{\mathbf{z} \mid \mathbf{D}\mathbf{z} = \mathbf{b},\ \mathbf{z} \geq 0\}$ is the constant c. The non-negativity constraint implies further that these set of solutions satisfy $\|\mathbf{z}\|_1 = c$. This property will be found useful in developing the next result.

3.2 Main Result

Theorem 2 *Given the system* $\mathbf{D}\mathbf{z} = \mathbf{b},\ \mathbf{z} \geq 0$*, such that all its solutions satisfy* $\mathbf{1}^T\mathbf{z} = c$*, if* $\bar{\mathbf{z}}$ *is a solution to this problem with*[4] $\|\bar{\mathbf{z}}\|_0 < \frac{1}{2t_{\mathbf{D}}}$*, then* $\bar{\mathbf{z}}$ *is the unique solution, i.e.* $\{\mathbf{z} \mid \mathbf{D}\mathbf{z} = \mathbf{b},\ \mathbf{z} \geq 0\}$ *is a singleton.*

Proof: Taking into account non-negativity of $\mathbf{z}$, we rewrite the condition $\mathbf{1}^T\mathbf{z} = c$ differently, as $\|\mathbf{z}\|_1 = c$. The vector $\bar{\mathbf{z}}$ is a sparse (with less than $1/2t_{\mathbf{D}}$ non-zeros) feasible solution of the linear programming problem

$$\min_{\mathbf{z}} \ \|\mathbf{z}\|_1 \quad \text{subject to} \quad \mathbf{D}\mathbf{z} = \mathbf{b}. \tag{3-7}$$

Notice that we do not specify the constraint $\mathbf{1}^T\mathbf{z} = c$ because any feasible solution of this problem must satisfy this condition anyhow. Also, we do not add a non-negativity constraint – the problem is defined as described above, and we simply observe that $\bar{\mathbf{z}}$ is a feasible solution.

[4]The definition of t_D is given in Section 2.

By Theorem 1, $\bar{\mathbf{z}}$ is necessarily the unique global minimizer of (3-7), i.e. any other feasible vector $\mathbf{z}$ satisfying $\mathbf{Dz} = \mathbf{b}$ has larger ℓ_1-norm, $\|\mathbf{z}\|_1 > c$. Hence, being non-negative, it cannot satisfy $\mathbf{1}^T\mathbf{z} = c$, and therefore it can not be a solution of $\mathbf{Dz} = \mathbf{b},\ \mathbf{z} \geq 0$. $\square$

We add the following brief discussion to get more intuition on the above theorem. Assume that a very sparse vector $\bar{\mathbf{z}}$ has been found to be a feasible solution of $\mathbf{Dz} = \mathbf{b},\ \mathbf{z} \geq 0$. At least locally, if we aim to find other feasible solutions, we must use a deviation vector that lies in the null-space of $\mathbf{D}$, i.e., $\mathbf{D}\delta = \mathbf{0}$. Positivity of the alternative solution $\bar{\mathbf{z}} + \delta$ forces us to require that at the off-support of $\bar{\mathbf{z}}$, all entries of δ are non-negative. Thus, the above theorem is parallel to the claim that such constrained vector is necessarily the trivial zero.

We now turn to present the consequence of the above Theorem to the original system we start from.

Corollary 1 *We are given the system* $\mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq 0$*, and* $\mathbf{A} \in \mathcal{M}^+$*. We canonize this system to obtain the matrix* $\mathbf{D}$ *as shown in Section 3.1. Then, if* $\bar{\mathbf{x}}$ *is a solution to this problem with* $\|\bar{\mathbf{x}}\|_0 < \frac{1}{2t_{\mathbf{D}}}$*, then* $\bar{\mathbf{x}}$ *is the unique solution, i.e.* $\{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq 0\}$ *is a singleton.*

Proof: Due to Theorem 2 we know that there is only one solution to the canonical system $\mathbf{Dz} = \mathbf{b},\ \mathbf{z} \geq 0$. Thus, due to the bijective relation to the original system $\mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq 0$ this uniqueness follows. $\square$

Note that the condition for this uniqueness is posed in terms of the coherence property of the canonic matrix $\mathbf{D}$. Furthermore, the fact that $\mathbf{A} \in \mathcal{M}^+$ implies that there exists $\mathbf{h}$ to drive this canonization process. However, there could be many possible such vectors, and each may lead to a different coherence. This brings us to the next discussion on better conditioning of the system.

3.3 Reducing The Coherence of the System Matrix

When converting (3-1) into (3-3), we use $\mathbf{D} = \mathbf{AW}^{-1}$ with $\mathbf{W}$ being a diagonal matrix based on an arbitrary positive vector $\mathbf{w}$ from the linear span of rows of $\mathbf{A}$. Thus, there is a flexibility in forming the canonic system due to the choice of this linear combination, governed by $\mathbf{h}$. Naturally, we would like to choose $\mathbf{h}$ so as to minimize $\rho(\mathbf{D})$.

There is yet another way to manipulate the coherence measure and sharpen the claim of Theorem 2 and Corollary 1. After canonization, one can multiply the linear system $\mathbf{Dz} = \mathbf{b}$ by *any* invertible matrix $\mathbf{P}$, and get a new and equivalent system $\mathbf{PDz} = \mathbf{Pb}$ which has the same set of solutions. Furthermore, the property $\mathbf{1}^T\mathbf{z} = c$ remains true for this system as well, and thus Theorem 2 is still valid, although with a bound using $\rho(\mathbf{PD})$.

In general, one could pose an optimization problem of minimizing the coherence using the best possible $\mathbf{h}$ and $\mathbf{P}$:

$$\min_{\mathbf{P},\mathbf{h}} \ \rho(\mathbf{PA}\text{diag}^{-1}(\mathbf{A}^T\mathbf{h})) \ \text{ s.t. } \ \mathbf{A}^T\mathbf{h} > 0. \tag{3-8}$$

Multiplying the vector $\mathbf{h}$ by a constant does not change the result, and therefore the ill-posed constraint $\mathbf{A}^T\mathbf{h} > 0$ can be changed to $\mathbf{A}^T\mathbf{h} \geq 1$. This optimization problem is nonconvex and requires a numerical technique for its solution. Having chosen $\mathbf{h}$, the optimization over the choice of $\mathbf{P}$ could be done using the algorithm presented in [14]. Fixing $\mathbf{P}$, one could devise a way to optimize the objective with respect to $\mathbf{h}$. This way, a block-coordinate-descent method could be envisioned.

Assuming the the optimal values $\hat{\mathbf{P}}$ and $\hat{\mathbf{h}}$ could be found somehow, this implies that there is a fundamental bound on the required sparsity of the solution $\bar{\mathbf{x}}$ that depends only on $\mathbf{A}$, and which guarantees uniqueness. Could this fundamental bound be found (or approximated) theoretically? We leave this and related problems for a future work, and refer hereafter to a specific case of positive matrices.

3.4 Positive Matrices

For a positive matrix $\mathbf{A}$, the simple choice $\mathbf{h} = \mathbf{1}$ leads to a $\mathbf{D} = \mathbf{AW}^{-1}$ that normalizes (in ℓ_1-norm) the columns of $\mathbf{A}$. An efficient coherence reduction can be obtained in this case by subtracting the mean of each column from its entries. This is achieved by

$$\mathbf{PD} = \left(\mathbf{I} - \frac{1}{n}\mathbf{1}\mathbf{1}^T\right)\mathbf{D}. \tag{3-9}$$

Note that the matrix $\mathbf{P}$ is non-invertible, and thus we should use $(\mathbf{I} - \frac{1-\epsilon}{n}\mathbf{1}\mathbf{1}^T)$ instead, with ϵ being a small positive constant $0 < \epsilon << 1$.

In order to illustrate the effect of this conditioning by $\mathbf{P}$, we present the following test. We construct random matrices $\mathbf{A}$ of size $n \times 2n$, where n is in the range $[50, 1000]$, and the entries are drawn independently from a uniform distribution in the range $[0, 1]$ – thus obtaining positive matrices. By normalizing the columns to a unit ℓ^1-norm we obtain the canonical matrices $\mathbf{D}$. Naturally, the measures $\rho(\mathbf{D})$ and $\rho(\mathbf{PD})$ are random variables, with mean and standard deviation being a function of n.

Note that past study of matrices with zero-mean Gaussian entries leads to the conclusion that $\mu(\mathbf{D})$ is asymptotically proportional to $1/\sqrt{n}$ up to a log factor [8, 25]. Can the same be said about $\rho(\mathbf{D})$ and $\rho(\mathbf{PD})$? Figure 1 shows the obtained coherence measures averaged over 100 experiments for each value of n. Without the multiplication by $\mathbf{P}$, the coherence tends towards a higher (and fixed!) value, being 0.75. The multiplication by $\mathbf{P}$ causes a reduction of the coherence, behaving like $1/\sqrt{n}$. Note that we do not provide proofs for these properties, as these are deviations from the main point of this paper.

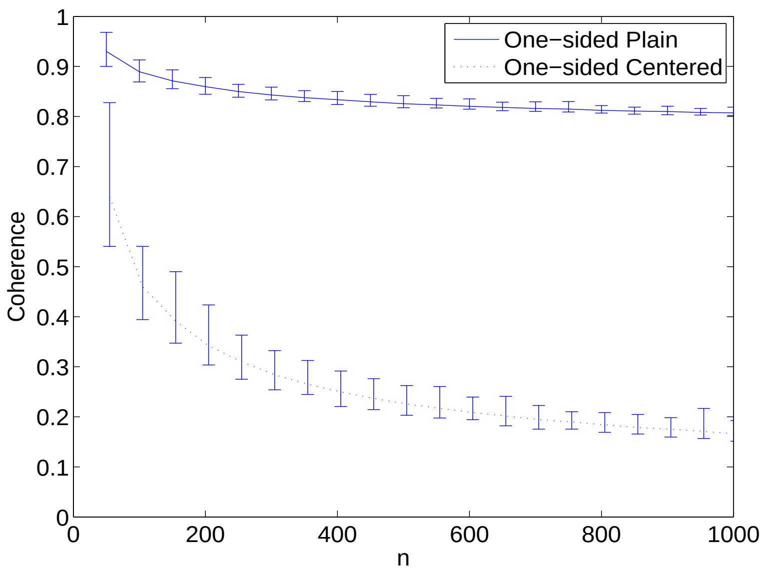

Figure 1: One sided centered versus plain coherence for random matrices of size $n \times 2n$ with varying n in the range $[50, 1000]$. The vertical bars show the estimated standard-deviation.

4 Orthogonal Matching Pursuit Performance

4.1 Approximation Algorithm

We have defined an optimization task of interest, (P_0^+) in (1-2) but this problem is very hard to solve in general. We could replace the ℓ_0-norm by an ℓ_1, and solve a linear programming problem. This is commonly done in a quest for sparse solutions of general linear systems, with good theoretical foundations. In fact, based on Theorem 2, one could replace the ℓ_0 with any other norm, or just solve a non-negative feasibility problem, and still get the same result, if it is indeed sparse enough. However, when this is not the case, we may deviate strongly from the desired solution of (P_0^+).

An alternative to the ℓ_1 measure is a greedy algorithm. We present this option in this section and study its performance, both empirically and theoretically. Specifically, we consider the use of the orthogonal matching pursuit (OMP) algorithm [21, 23], finding the sparsest and non-negative solution to $\mathbf{Ax} = \mathbf{b}$ one atom at a time. The algorithm is described in Figure 2, operating on the canonic pair $(\mathbf{D}, \mathbf{b})$. This is a modified version of the regular OMP that takes into account the non-negativity of the sought solution.

When using the OMP, one can either operate on the original system $\mathbf{Ax} = \mathbf{b}$, or its canonic version $\mathbf{Dz} = \mathbf{b}$ (and with better conditioning by $\mathbf{P}$). While the solutions of all these are equivalent, how well will the OMP perform in the different settings? In the following subsections we offer two kinds of answers - an empirical one and a theoretical one. We start with an empirical evidence.

4.2 Experimental Results

We experiment with random non-negative and canonic matrices $\mathbf{D}$, as in Section 3.4. We consider a fixed size 100×200, and generate 1000 random sparse representations with varying cardinalities in the range $1 - 40$, checking the performance of the OMP in recovering them. We test both the OMP on the plain system (note that the first canonization step that leads to ℓ_1-normalized columns has no impact on the OMP behavior) and the conditioned one with the centering by $\mathbf{P}$. The implemented OMP is as described in Figure 2. The *Update*

Task: Solve approximately (P_0^+) : $\min_{\mathbf{z}\geq\mathbf{0}} \;\|\mathbf{z}\|_0$ subject to $\mathbf{Dz}=\mathbf{b}$

Initialization: Fix $i=0$ and set

- The temporary solution $\mathbf{z}^i=0$.
- The temporary residual $\mathbf{r}^i=\mathbf{b}-\mathbf{Dz}^i=\mathbf{b}$.
- The temporary solution support $\mathcal{S}^i=Support\{\mathbf{z}^i\}=\phi$.

Main Iteration: Increment i, and apply

- **Sweep:** Compute the following errors for all $1\leq j\leq k$:

$$\epsilon(j)=\min_{z_j\geq 0}\|\mathbf{d}_j z_j-\mathbf{r}^{i-1}\|_2^2=\|\mathbf{r}^{i-1}\|_2^2-\frac{\max\{\mathbf{d}_j^T\mathbf{r}^{i-1},0\}^2}{\|\mathbf{d}_j\|_2^2}.$$

- **Update Support:** Find j_0 such that $\forall\ j\in\mathcal{S}^c,\ \epsilon(j_0)\leq\epsilon(j)$, and update $\mathcal{S}^i=\mathcal{S}^{i-1}\cup\{j_0\}$ accordingly.
- **Update Solution:** Compute $\mathbf{z}^i$

$$\mathbf{z}^i=\min_{\mathbf{z}}\ \|\mathbf{Dz}-\mathbf{b}\|_2^2 \text{ subject to } Support\{\mathbf{z}^i\}=\mathcal{S}^i \text{ and } \mathbf{z}\geq 0.$$

- **Update Residual:** Compute $\mathbf{r}^i=\mathbf{b}-\mathbf{Dz}^i$.
- **Stopping Rule:** If $\|\mathbf{r}^i\|_2^2<T$, stop. Otherwise, apply another iteration.

Output: The desired solution is $\mathbf{z}^i$.

Figure 2: The orthogonal matching pursuit (OMP) algorithm for solving (P_0^+).

Solution step is implemented using Matlab's `lsqnonneg` instruction.

For comparison, we also test the Basis Pursuit (BP), solving the problem

$$\min_{\mathbf{z}} \; \mathbf{1}^T\mathbf{z} \quad \text{s.t.} \quad \mathbf{D}\mathbf{z} = \mathbf{b}, \; \mathbf{z} \geq 0. \tag{4-1}$$

When there exist only one solution, this method necessarily finds it exactly. On the other hand, when there are several possible solutions, it does not necessarily find the sparsest one, thus leading errors. Note that this alternative requires many more computations, as its complexity is much[5] higher. Also, conditioning of the form discussed above does not affect its solutions.

Figure 3 shows the relative average number of wrong atoms detected in the tested algorithms. Figure 4 shows the average number of badly represented signals (i.e., those not satisfying $\mathbf{D}\mathbf{z} = \mathbf{b}$). As can be seen in both graphs, the conditioned OMP performs much better. We also see, as expected, that BP outperforms both greedy options and yielding very low error rate, with the obvious added cost in complexity. Notice that the BP's representation error is zero simply because BP always finds a solution to satisfy $\mathbf{D}\mathbf{z} = \mathbf{b}$, whereas the OMP is operated with a fixed (assumed to be known) number of atoms.

4.3 Theoretical Study

Turning to a theoretical assessment of the OMP, we follow previous results developed in [29, 9], we state the following Theorem without proof[6].

Theorem 3 *For the linear system of equations* $\mathbf{D}\mathbf{z} = \mathbf{b}$ *(where* $\mathbf{D} \in \mathbb{R}^{n\times k}$*), if a non-negative solution exists such that*

$$\|\mathbf{z}\|_0^0 < \frac{1}{2}\left(1 + \frac{1}{\mu\{\mathbf{D}\}}\right), \tag{4-2}$$

then OMP (as described in Figure 2) is guaranteed to find it exactly.

[5] As an example, the Matlab run-time ratio BP-versus-OMP for 1000 examples was found to be roughly $500/\|\mathbf{z}\|_0$.

[6] The proof is similar to the one given in [29, 9].

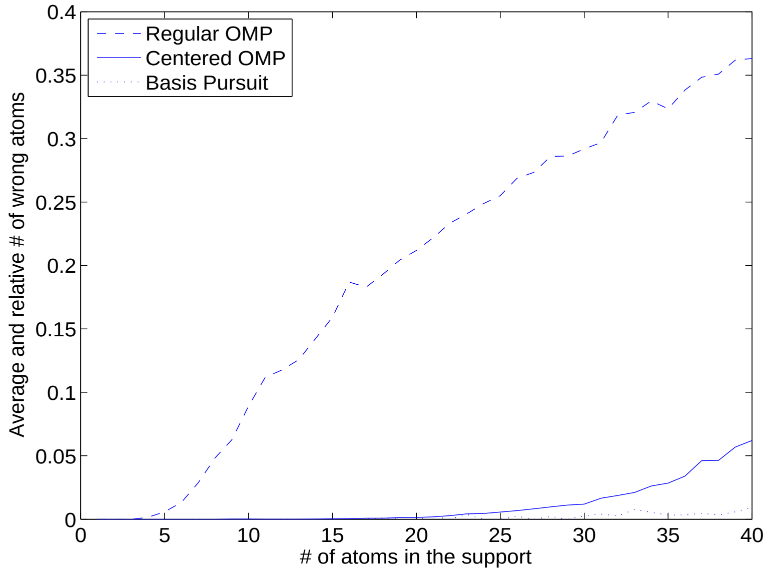

Figure 3: Performance comparison between regular and centered OMP. This graph shows the relative and average number of wrongly detected atoms as a function of the original cardinality.

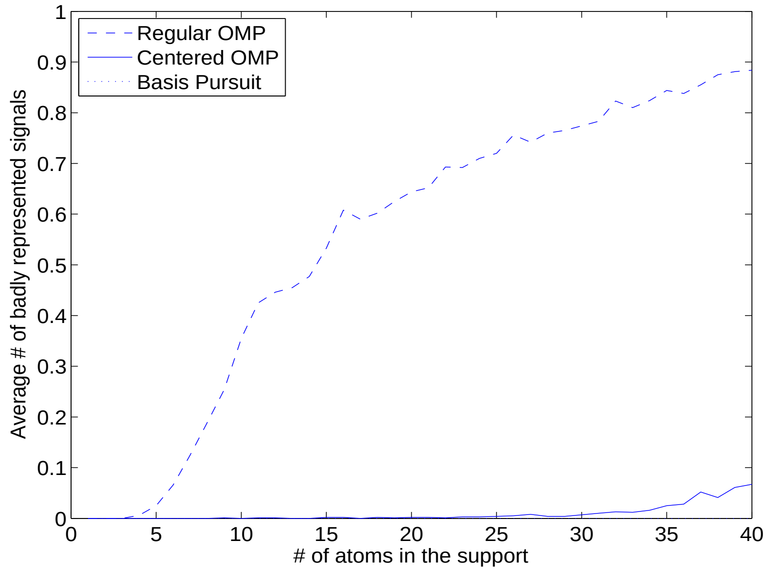

Figure 4: Performance comparison between regular and centered OMP. This graph shows the average number of wrongly represented signals from the test set.

The above theorem is quite weak in general since $\mu\{\mathbf{D}\}$ may be too large. For example, in the non-negative case, $\mu\{\mathbf{D}\}$ tends to 0.75, implying that we can handle only empty supports. one may apply the very same analysis to the conditioned problem, and obtain a better result.

Theorem 4 *For the linear system of equations* $\mathbf{PDz} = \mathbf{Pb}$ *(where* $\mathbf{D} \in \mathbb{R}^{n\times k}$*,* $\mathbf{P} \in \mathbb{R}^{n\times n}$ *is invertible,) if a non-negative solution exists such that*

$$\|\mathbf{z}\|_0^0 < \frac{1}{2}\left(1 + \frac{1}{\mu\{\mathbf{PD}\}}\right), \tag{4-3}$$

then OMP (as described in Figure 2) is guaranteed to find it exactly.

Considering the non-negative case, as in the experiments above, the centered OMP is indeed performing much better, and theoretically we see that this is an expected phenomenon. However, as mentioned in past work on the analysis of pursuit algorithms, we should note that the bounds we provide here are far worse compared to the actual (empirical) performance, as they tend to be over-pessimistic. In the experiments reported in the previous section we have $\mu\{\mathbf{D}\} = 0.858$ and $\mu\{\mathbf{PD}\} = 0.413$, implying that at best one can recover supports of cardinality $T = 1$. Clearly, the OMP succeeds far beyond this point.

5 Conclusions

Linear systems of equations with a positivity constraint come up often in applications in signal and image processing. Solving such systems is usually done by adding conditions such as minimal ℓ_2 length, maximal entropy, maximal sparsity, and so on. In this work we have shown that if a sparse enough solution exists, then it is the only one, implying that all the mentioned measures lead to the same solution. We also have proposed an effective conditioning for improving the chances of such linear system to be handled well by greedy algorithms.

There are several directions in which this work should/could be extended, and several intriguing questions that form the grounds for such extended work. What is the optimal choice for the canonization parameters? Answering this may lead to stronger bounds for the discovered uniqueness. Also, as there is a clear gap between the proved bound and the empirical behavior, can we strengthen the bounds by relying on a probabilistic analysis? These questions and more promise a fruitful path for more work on this topic.

Acknowledgements

The authors would like to thank Prof. Amnon Shashua from the Hebrew University in Jerusalem, for a lively discussion about the entropy measure and non-negative matrix factorizations, which helped in focusing this work. Special thanks are also in order to the two reviewers who's comments helped improving this paper.

References

[1] S.S. Chen, D.L. Donoho, and M.A. Saunders, "Atomic decomposition by basis pursuit", *SIAM Review*, Vol. 43, pp. 129–159, Jan. 2001.

[2] M. Chu and R. Plemmons, "Nonnegative matrix factorization and applications", *IMAGE – Bulletin of the International Linear Algebra Society*, Vol. 34, pp. 2–7, July 2005.

[3] R.R. Coifman and M.V.Wickerhauser,"Entropy-based algorithms for best basis selection", *IEEE Trans. on Information Theory*, Vol. 38, pp. 713–718, Mar. 1992.

[4] D. De-Vries and A.J. Berkhout, "Velocity analysis based on minimum entropy Geophysics", *Geophysics*, Vol 49, pp. 2132–2142, Dec. 1984.

[5] D.L. Donoho, "On minimum entropy segmentation", In *Wavelets: Theory, Algorithms, and Applications*, edited by C.K.Chui et. al., pp. 233–269, Academic Press, 1994.

[6] D.L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization", *Proceedings Of The National Academy Of Sciences Of The United States Of America*, Vol. 100, pp. 2197–2202, Mar. 2003.

[7] D.L. Donoho and M. Elad, "On the stability of the basis pursuit in the presence of noise", *Signal Processing*, Vol. 86, pp. 511–532, Mar. 2006.

[8] D.L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition", *IEEE Trans. on Information Theory*, Vol. 47, pp. 2845–2862, Nov. 2001.

[9] D.L. Donoho, M. Elad, and V.N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise", *IEEE Trans. on Information Theory*, Vol. 52 pp. 6–18, Jan. 2006.

[10] D.L. Donoho, and P.B. Stark, "Uncertainty principles and signal recovery", *SIAM J. Appl. Math.*, Vol. 49, pp. 906–931, June 1989.

[11] D.L. Donoho and J. Tanner, "Sparse nonnegative solutions of underdetermined linear equations by linear programming", *Proceedings of the National Academy of Sciences*, Vol. 102, No. 27, pp. 9446–9451, 2005.

[12] D.L. Donoho and J. Tanner, "Thresholds for the recovery of sparse solutions via L1 minimization", *Proceedings of the Conference on Information Sciences and Systems*, March 2006.

[13] D.L. Donoho and J. Tanner, "Neighborliness of randomly-projected simplices in high dimensions", *Proceedings of the National Academy of Sciences*, Vol. 102, No. 27, pp. 9452–9457, 2005.

[14] M. Elad, "Optimized projections for compressed-sensing", *IEEE Trans. on Signal Processing*, Vol. 55, No. 12, pp. 5695–5702, Dec. 2007.

[15] M. Elad and A.M. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases",*IEEE Trans. on Information Theory*, Vol. 48, pp. 2558–2567, Sep. 2002.

[16] O.D. Escoda, L. Granai and P. Vandergheynst, "On the use of a priori information for sparse signal approximations", *IEEE Trans. Signal Processing*, Vol 54, No 9, pp 3468–3482, 2006.

[17] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases", *IEEE Trans. on Information Theory*, Vol. 49, pp. 3320–3325, Dec. 2003.

[18] T. Hazan, S. Polak and A. Shashua, Sparse image coding using a 3d non-negative tensor factorization, *International Conference on Computer Vision (ICCV)*, Beijing, China, October, 2005.

[19] R. Gribonval and M. Nielsen, "On the strong uniqueness of highly sparse representations from redundant dictionaries", *ICA 2004: Lecture Notes In Computer Science*, Vol. 3195, pp. 201–208, 2004.

[20] K. Kreutz-Delgado, B. D. Rao, K. Engan, T.W. Lee and T. Sejnowski, Convex/Schur-convex (csc) log-priors and sparse coding, Proceedings of the *Sixth Joint Symposium on Neural Computation*, Pasadena, CA, My 1999.

[21] S. Mallat, and Z. Zhang, "Matching pursuit with time-frequency dictionaries", *IEEE Trans. Signal Processing*, Vol. 41, pp. 3397–3415, Dec. 1993.

[22] M. Nafie, M. Ali, and A.H. Tewfik, "Optimal subset selection for adaptive signal r.epresentation", Proceedings of the *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 1996.

[23] Y. C. Pati, R. Rezaiifar, and P. S. Krishnaprasad, Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition, in Conf. Rec. 27th Asilomar Conf. Signals, Syst. Comput., Vol. 1, 1993.

[24] P. Pauca, J. Piper, and R. Plemmons, Nonnegative matrix factorization for spectral data analysis, *Linear Algebra and Applications*, 2006.

[25] S. Raboi, *On the recovery of missing samples via sparsity conditions*, M.Sc. Thesis, The Electrical Engineering department, the Technion - Israel Institute of Technology, February 2007.

[26] B.D. Rao and K. Kreutz-Delgado, "Deriving algorithms for computing sparse solutions to linear inverse problems", *Asilomar Conference on Signals, Systems and Computers*, Vol. 1, pp. 955–959, Monterey, CA, November 1997.

[27] A. Shashua and T. Hazan, “Non-negative tensor factorization with applications to statistics and computer vision”, *International Conference on Machine Learning (ICML)*, August 2005.

[28] A. Shashua, R. Zass and T. Hazan, “Multi-way clustering using super-symmetric non-negative tensor factorization”, Proc. of the *European Conference on Computer Vision (ECCV)*, May 2006, Graz, Austria.

[29] J.A. Tropp, “Greed is good: Algorithmic results for sparse approximation”, *IEEE Trans. on Information Theory*, Vol. 50, pp. 2231–2242, Oct. 2004.

[30] J.A. Tropp, “Just relax: Convex programming methods for identifying sparse signals in noise”, *IEEE Trans. on Information Theory*, Vol. 52, pp. 1030–1051, Mar. 2006.

[31] R. Zass and A. Shashua, “Nonnegative sparse PCA”, *Neural Information and Processing Systems (NIPS)*, December, 2006.

Biographies

Alfred Bruckstein

Alfred Bruckstein received the B.Sc. and M.Sc. degrees in Electrical Engineering, from the Technion, Israel Institute of Technology, in 1977 and 1980, respectively. His M.Sc. thesis dealt with models of coding in the nervous system. In 1980 he joined the Information Systems Laboratory in the Electrical Engineering Department at Stanford University, in California, as a Research Assistant and worked on several topics related to direct and inverse scattering and signal processing. In October 1984, he completed the requirements for the Ph.D. degree in Electrical Engineering with a thesis on Scattering Models in Signal Processing.

In the autumn of 1981 he spent two month at the Physics Department of Groningen University, with the Biophysics Group, analyzing the responses of movement-sensing giant neurons in the visual system of the fly. In the summers of 1980 and 1982 he visited the Massachusetts Institute of Technology, at the Man-Vehicle Laboratory and the Information and Decision Systems Laboratory, respectively.

From October 1984 to June 1989, he has been with the Faculty of Electrical Engineering at the Technion, Haifa. In December 1987 Dr. Bruckstein was tenured at the Technion, in January 1989 he became an Associate Professor there, and in 1995 he was promoted to Full Professorship in the Computer Science Department. Since 1999, he is holder of the Technion Ollendorff Chair in Science. At the Technion he served from 2002 to 2005 as the Dean of the Technion Graduate School, and is presently the Head of Technion's Excellence Program.

His present research interests are in distributed (swarm or multi-a(ge)nt) robotics, image and signal processing, computer vision, pattern recognition, and various aspects of applied geometry. He has done work in the past, and is still interested in algorithmic aspects of inverse scattering problems, and in point processes as mathematical models in neurophysiology.

Alfred Bruckstein authored and co-authored over one hundred journal papers in the fields of interest mentioned.

He is a member of IEEE, SIAM, the American Mathematical Society and the Mathematical Association of America.

Michael Elad

Michael Elad received his B.Sc (1986), M.Sc.(supervision by Prof. David Malah - 1988) and D.Sc. (supervision by Prof. Arie Feuer 1997) from the department of Electrical engineering at the Technion, Israel. From 1988 to 1993 he served in the Israeli Air Force. From 1997 to 2000 he worked at Hewlett-Packard laboratories as an R&D engineer. From 2000 to 2001 he headed the research division at Jigami corporation, Israel.

During the years 2001 to 2003 Michael spent a post-doc period as a research-associate with the computer science department at Stanford university (SCCM program). On September 2003 Michael returned to the Technion, assuming a tenure-track assistant professorship position in the department of Computer science. On May 2007 Michael was tenured to an associate professorship.

Michael Elad works in the field of signal and image processing, specializing in particular on inverse problems, sparse representations and overcomplete transforms. Michael received the Technion's best lecturer award five times (1999, 2000, 2004, 2005, and on 2006), he is the recipient of the Solomon Simon Mani award for excellence in teaching in 2007, and he is also the recipient of the Henri Taub Prize for academic excellence (2008). Michael is a senior IEEE member, and he is currently serving as an associate editor for IEEE Trans. on image processing.

Michael Zibulevsky

Michael Zibulevsky received his M.Sc. degree in Electrical Engineering from MIIT in Moscow, and the D.Sc. degree in Operations Research (Nonlinear Optimization) from the Technion, Israel Institute of Technology, in 1996. In D.Sc. Thesis he introduced Augmented Lagrangian method for Semidefinite Programming and corresponding duality analysis. Dr. Zibulevsky spent 2 years in University of New-Mexico, working on blind source separation via sparse signal representations, forming the foundation to this field. Currently he is with the Department of Computer Science at the Technion. His area of interests includes nonlinear optimization, signal/image processing, with applications to blind source separation, computer tomography, ultrasound imaging, EEG/MEG. Dr. Zibulevsky is also very active in the field of sparse approximation and its application to image processing.