IMI

Preprint Series

Industrial Mathematics Institute

2000:12

A counter-example to the convergence of greedy algorithms under very weak conditions

R. Gribonval

Department of Mathematics
University of South Carolina

A counter-example to the convergence of greedy algorithms under very weak conditions*

R. Gribonval

IMI, Dept of Mathematics, Univ. of South Carolina, Columbia, SC 29208, USA
E-mail: remi@math.sc.edu

In a separable Hilbert space $\boldsymbol{H}$, greedy algorithms iteratively define m-term approximants to a given vector from a complete redundant dictionary $\mathcal{D}$. With very large dictionaries, the pure greedy algorithm cannot be implemented and must be replaced with a weak greedy algorithm. A conjecture about the convergence of ***very weak greedy algorithms*** arises naturally from the observation of numerical experiments. We introduce, study and disprove this conjecture.

Key Words: nonlinear approximation, greedy algorithms, redundant dictionary, convergence.

1. INTRODUCTION

Given a complete dictionary $\mathcal{D}$ of unit vectors (or *atoms*) in a separable Hilbert space $\boldsymbol{H}$, one can consider the approximations of a vector R_1 by linear combinations of atoms taken from $\mathcal{D}$. When the dictionary is indeed an orthonormal basis, the best m-term approximation to R_1 can actually be constructed. But whenever the dictionary is redundant, there is no unique linear decomposition of R_1, and the best m-term approximation may be difficult to build. A greedy algorithm (known as Matching Pursuit in signal processing [8], or Projection Pursuit in statistics [6]) provides such an m-term approximation by constructing a sequence $R_m \in \boldsymbol{H}, m \geq 1$ such that at each step

$$R_m = \langle R_m, g_m \rangle \, g_m + R_{m+1}, \; g_m \in \mathcal{D} \tag{1}$$

* This research was supported by the National Science Foundation (NSF) Grant No. DMS-9872890.

with

$$|\langle R_m, g_m\rangle| = \sup_{g\in\mathcal{D}} |\langle R_m, g\rangle| . \tag{2}$$

It was proved that such a sequence converges weakly [6] and strongly [7] to zero. When the dictionary $\mathcal{D}$ is very large, the choice (2) of the best atom from $\mathcal{D}$ maybe so computationnaly costly that a sub-optimal choice has to be considered. More recently, the convergence of greedy algorithms was proved under the weaker sufficient condition

$$|\langle R_m, g_m\rangle| \geq t_m \sup_{g\in\mathcal{D}} |\langle R_m, g\rangle| , \tag{3}$$

provided that (t_m) complies with some additional condition. In [9] the convergence is proved if $\exists t > 0, \forall m, t_m \geq t$. In [10] the stronger result of convergence whenever $\sum_m t_m/m = \infty$ is proved, and the remark was made that divergence may occur when $\sum_m t_m^2 < \infty$. An open question consists in filling the gap between these two conditions on (t_m).

In [1, 2], a pursuit with sub-dictionaries of local maxima of $\mathcal{D}$ was defined for the approximation of images. The same technique was developped in [4, 5] in order to accelerate the analysis of sound signals. Between iterations m_p and $m_{p+1} - 1$, a sub-dictionary $\mathcal{D}_p$ of local maxima is used. It is computed at the iteration m_p and contains the best atom for this iteration. The convergence of such an algorithm was proved [1] under the restrictive assumption that $m_{p+1} - m_p$ is bounded. The proof used the convergence of weak greedy algorithms [9]. The results from [10] show that a weaker sufficient condition for the convergence of this algorithm is $\sum_p \frac{1}{m_p} = \infty$. For example if $m_p = p \log p$, there is convergence but $m_{p+1} - m_p$ is not bounded.

With a multiscale time-frequency dictionary of chirps [5], we suggested a much weaker two-step choice of the "best" atom of $\mathcal{D}$. At first we only consider a complete sub-dictionary $\mathcal{D}^\star$ of reasonable size, and select the best atom $g_m^\star \in \mathcal{D}^\star$. Then we improve this choice by choosing a "locally optimal" atom $g_m \in \mathcal{D}$ in a "neighborhood" of $g_m^\star$. The final choice thus complies with

$$|\langle R_m, g_m\rangle| \geq |\langle R_m, g_m^\star\rangle| = \sup_{g\in\mathcal{D}^\star} |\langle R_m, g\rangle| . \tag{4}$$

Such a stepwise choice $g_m \in \mathcal{D}$ is generally much weaker than (3), if no additional assumption is made about $\mathcal{D}$ and $\mathcal{D}^\star$. We call the corresponding class of algorithms "very weak greedy algorithms" (VWGA).

A VWGA is "stepwise better" than a pure greedy algorithm in $\mathcal{D}^\star$. A natural question is whether such a choice of $g_m \in \mathcal{D}$ will improve the speed of convergence, compared to a pursuit in $\mathcal{D}^\star$. That is to say, if one is doing a pursuit in $\mathcal{D}^\star$ (which is known to converge), is it a good idea to get at each step a "better atom" in $\mathcal{D}$? Does it "take advantage" of the extra redundancy given by $\mathcal{D}$? The intuition tells us that it should converge, may be with an improvement of the speed of convergence. In finite dimension, the convergence of VWGA is actually trivial.

The goal of this report is to show that the intuition is false in infinite dimension. Not only the VWGA can converge more slowly than a pursuit in $\mathcal{D}^\star$ (this is already known from a counter-example of DeVore and Temlyakov [3], where $\mathcal{D}^\star$ an orthonormal basis and $\mathcal{D} = \mathcal{D}^\star \cup \{g_0\}$ with g_0 a "bad" vector), but it may not converge at all. We show this by building a counter-example.

In the first section, we give a precise statement of the natural conjecture, and define the notion of (pure, weak and very weak) greedy sequence. In the second section we build a very weak greedy sequence that is our counter-example. In the third section we make some comments about the implications of our result.

2. GREEDY ALGORITHMS AND SEQUENCES

The natural conjecture about VWGA is the following

Conjecture 2.1. *Let $\mathcal{D}^\star \subset \mathcal{D}$ two complete dictionaries. Let $\{R_m\}_{m\geq 1}$ such that, for all m, (1) holds with some $g_m \in \mathcal{D}$ chosen such that (4) is true. Then $\|R_m\|_2 \to 0$.*

If this conjecture were true, then it should be true when $\mathcal{D}^\star = \mathcal{B}$ is an orthonormal basis of $\boldsymbol{H}$ and $\mathcal{D} \supset \mathcal{B}$ is any dictionary containing this basis. It would imply the following result.

Conjecture 2.2. *Let $\mathcal{B} = \{e_n, n \in \mathbb{N}\}$ an orthonormal basis of $\boldsymbol{H}$, and define for any $x \in \boldsymbol{H}$: $\|x\|_\infty := \sup_{n\in\mathbb{N}} |\langle x, e_n\rangle|^2$. Let $\{R_m\}_{m\geq 1}$ a sequence such that for all $m \geq m_0$,*

$$\langle R_m, R_{m+1}\rangle = \|R_{m+1}\|_2^2 \tag{5}$$

and

$$\|R_m\|_2^2 - \|R_{m+1}\|_2^2 \geq \sup_{n\in\mathbb{N}} |\langle R_m, e_n\rangle|^2 = \|R_m\|_\infty^2 \tag{6}$$

Then $\|R_m\|_2 \to 0$.

Proof. Let R_m comply with (5) and (6). We shall prove that (1) and (4) hold for $m \geq m_0$, for some dictionary $\mathcal{D}$ that we will specify later on. As the convergence is an asymptotic property, it is clear that we can replace, in conjecture 2.1, "(1) and (4) hold for all m" by "(1) and (4) hold for all $m \geq m_0$". Thus, if conjecture 2.1 is true, we will get $\|R_m\|_2 \to 0$.

Let us define

$$g_m := (R_m - R_{m+1}) / \|R_m - R_{m+1}\|_2 \tag{7}$$

and

$$\mathcal{D} := \mathcal{B} \cup \{g_m, m \in \mathbb{N}\}\,. \tag{8}$$

From the definition of g_m, $R_m = \delta_m g_m + R_{m+1}$ for some δ_m. From (5) we get $\langle R_{m+1}, g_m\rangle = 0$, thus $\delta_m = \langle R_m, g_m\rangle$, which shows (1).

Now from (6) we have $|\langle R_m, g_m\rangle|^2 = \|R_m\|_2^2 - \|R_{m+1}\|_2^2 \geq \sup_{e_k \in \mathcal{B}} |\langle R_m, e_k\rangle|^2$ which shows (4). ∎

This enables us to only deal with properties of *sequences* in $\boldsymbol{H}$: we can forget about the algorithmic nature of the iterative decomposition. We will call *greedy sequence* any sequence R_m in $\boldsymbol{H}$ complying with (5). A greedy sequence that additionnaly complies with (6) will be called a *very weak* greedy sequence, by opposition to a *pure* greedy sequence which is supposed to satisfy the stronger condition (which is equivalent to (2))

$$\|R_m\|_2^2 - \|R_{m+1}\|_2^2 \geq \sup_{g \in \mathcal{D}} |\langle R_m, g\rangle|^2\,.$$

We shall prove that conjecture 2.1 is false by building a counter-example to conjecture 2.2, that is to say a very weak greedy sequence R_m such that $\|R_m\|_2$ is bounded below by some $c > 0$.

3. A COUNTER-EXAMPLE

For convenience, our counter-example $\{R_m\}_{m \geq 1}$ will be defined through its normalization $S_m := R_m / \|R_m\|_\infty$ in $\|.\|_\infty$. Let us first show that, for any $\lambda > 0$ and any sequence $\{S_m\}_{m \geq 1}$ such that $\|S_m\|_\infty = 1$ there is a unique sequence

$$R_m = \|R_m\|_\infty S_m.$$

such that (5) holds and $\|R_1\|_\infty = \lambda$. It is done by proving that the sequence of norms $\{\|R_m\|_\infty\}_{m \geq 1}$ is defined by λ and the sequence $\{S_m\}_{m \geq 1}$, which comes from the fact that (5) is equivalent to

$$\|R_{m+1}\|_\infty / \|R_m\|_\infty = \langle S_{m+1}, S_m\rangle \,/\, \langle S_{m+1}, S_{m+1}\rangle\,. \tag{9}$$

Condition (6) on $\{R_m\}_{m\geq 1}$ then becomes

$$\langle S_m, S_m\rangle - \langle S_{m+1}, S_m\rangle^2 / \langle S_{m+1}, S_{m+1}\rangle \geq 1. \tag{10}$$

Let us remark that for any greedy sequence, $\{\|R_m\|_2\}_{m\geq 1}$ is a decreasing positive sequence, hence it has a limit. For a very weak greedy sequence, the fact that (6)/(10) holds for $m \geq m_0$ thus implies $\|R_m\|_\infty \to 0$. Suppose now that $\{R_m\}_{m\geq 1}$ is a counter-example to conjecture 2.2, which is equivalent to $\lim \|R_m\|_2 > 0$: $\{R_m\}_{m\geq 1}$ corresponds to some "bad" sequence $\{S_m\}_{m\geq 1}$ complying with (10) such that

$$\|R_m\|_2 = \|R_m\|_\infty \|S_m\|_2 \geq c, \tag{11}$$

where $\|R_m\|_\infty$ is defined through (9) and $c > 0$ is some constant. Hence we must have $\|S_m\|_2 \to \infty$.

We are going to specify some badly behaved sequence $\{S_m\}_{m\geq 1}$ of vectors, such that $\|S_m\|_\infty = 1$, $\|S_m\|_2 \to \infty$, and (10)-(11) hold. Our setting is now $\boldsymbol{H} = l^2(\mathbb{N})$ and we define, for $0 < \epsilon < 1$

$$S(\epsilon) := \{(1-\epsilon)^n\}_{n\geq 0}. \tag{12}$$

One can easily check that $\|S(\epsilon)\|_\infty = 1$ and for all $0 < \epsilon, \eta < 1$,

$$\langle S(\epsilon), S(\eta)\rangle = \sum_{n=0}^{\infty}((1-\epsilon)(1-\eta))^2 = (\epsilon+\eta-\epsilon\eta)^{-1} = \frac{\frac{1}{\epsilon}\frac{1}{\eta}}{\frac{1}{\epsilon}+\frac{1}{\eta}-1}. \tag{13}$$

In particular

$$\|S(\epsilon)\|_2^2 = \left(2\epsilon - \epsilon^2\right)^{-1}. \tag{14}$$

Using this family of "bad vectors" we can now build the counter-example we have announced.

PROPOSITION 3.1. *Let $\boldsymbol{H} = l^2(\mathbb{N})$ and $\mathcal{B}$ its canonical basis. Let $\alpha > 2$ and $\epsilon_m = m^{-\alpha}$. Let $S_m = S(\epsilon_m)$. Let $R_m = \|R_m\|_\infty S_m$ where $\{\|R_m\|_\infty\}_{m\geq 1}$ is inductively defined by (9), with an arbitrary initial value $\|R_1\|_\infty > 0$. Then $\{R_m\}_{m\geq 1}$ is a counter example, that is to say : there exists $m_0 \in \mathbb{N}$ such that for all $m \geq m_0$, relations (5) and (6) hold, and*

$$\exists\ c > 0, \forall m, \|R_m\|_2 \geq c \tag{15}$$

Notations. We use the symbol $a_m \asymp b_m$ to denote the existence of constants c and C such that $ca_m \leq b_m \leq Ca_m$ for m big enough. The notation $a_m \sim b_m$ means that $a_m/b_m \to 1$. Finally $a_m = \mathcal{O}(b_m)$ is written when a_m/b_m is bounded.

Proof. Using (13) we get

$$\langle S_m, S_m\rangle - \frac{\langle S_{m+1}, S_m\rangle^2}{\langle S_{m+1}, S_{m+1}\rangle} = \frac{m^{2\alpha}\left((m+1)^\alpha - m^\alpha\right)^2}{(2m^\alpha - 1)\left((m+1)^\alpha + m^\alpha - 1\right)^2} \sim \frac{m^\alpha}{8}\left((1+1/m)^\alpha - 1\right)^2 \asymp m^{\alpha-2} \qquad (16)$$

which proves that (10) is true, because its left hand side goes to infinity.

From (14) we know that $\|S_m\|_2^2 \sim 1/(2\epsilon_m) \asymp m^\alpha$. It is thus sufficient to prove that $\|R_m\|_\infty \asymp m^{-\alpha/2}$ to obtain (11) and reach our conclusion. To get this result we use the sequence $v_m := \log \|R_{m+1}\|_\infty / \|R_m\|_\infty, m \geq 1$ and show that

$$v_m = -\frac{\alpha}{2}\log\frac{m+1}{m} + \mathcal{O}(m^{-2}). \qquad (17)$$

Indeed, from (9) we have $v_m = \log\langle S_{m+1}, S_m\rangle - \log\langle S_{m+1}, S_{m+1}\rangle$, thus using (13) we get

$$v_m + \log\frac{\epsilon_m + \epsilon_{m+1}}{2\epsilon_{m+1}} = \log(1-\epsilon_m/2) - \log(1-\epsilon_m\epsilon_{m+1}/(\epsilon_m+\epsilon_{m+1})) = \mathcal{O}(m^{-\alpha}). \qquad (18)$$

Moreover, one easily gets

$$\log\frac{\epsilon_m + \epsilon_{m+1}}{2\epsilon_{m+1}} = \log\left[1 + \frac{1}{2}\left((1+1/m)^\alpha - 1\right)\right] = \frac{\alpha}{2m} + \mathcal{O}(m^{-2}). \qquad (19)$$

As $1/m = \log(m+1) - \log m + \mathcal{O}(m^{-2})$, and $\alpha > 2$, (18) and (19) lead to (17). To finish with,

$$\log\|R_m\|_\infty - \log\|R_1\|_\infty + \frac{\alpha}{2}\log m = \sum_{k=1}^{m-1}\left(v_k + \frac{\alpha}{2}\log\frac{k+1}{k}\right) = \sum_{k=1}^{m-1}\mathcal{O}(1/m^2) \qquad (20)$$

has a limit $K \in \mathbb{R}$, which proves that

$$\|R_m\|_\infty m^{\alpha/2} \to C > 0. \qquad (21)$$

∎

4. COMMENTS AND CONSEQUENCES

This counter-example gives some additional information on the properties of weak greedy sequences. Let us state some (known) results about greedy sequences

LEMMA 4.1. *Let $\{R_m\}_{m\geq 1}$ be any* very weak *greedy sequence. If* $\sum \|R_m - R_{m+1}\|_2 < \infty$ *then* $\|R_m\|_2 \to 0$.

Proof. As $\boldsymbol{H}$ is complete, $\sum \|R_m - R_{m+1}\|_2 < \infty$ implies the (strong) convergence of $\{R_m\}_{m\geq 1}$ to some $R_\infty \in \boldsymbol{H}$. But we have seen that $\|R_m\|_\infty \to 0$ because of (6). Fatou's lemma thus gives $R_\infty = 0$. ∎

The counter-example we have built must thus comply with $\sum \|R_m - R_{m+1}\|_2 = \infty$. On the other hand, it is easy to see that

LEMMA 4.2. *For every greedy sequence $\{R_m\}_{m\geq 1}$,*

$$\sum \|R_m - R_{m+1}\|_2^2 < \infty.$$

Proof. The stepwise orthogonal decomposition (1) implies a stepwise energy conservation $\|R_m\|_2^2 = \|R_m - R_{m+1}\|_2^2 + \|R_{m+1}\|_2^2$, which gives $\sum \|R_m - R_{m+1}\|_2^2 \leq \|R_1\|^2$ by a telescoping sequence argument. ∎

We currently know that, for any counter-example to conjecture 2.2 :

$$\sum \|R_m - R_{m+1}\|_2 = \infty \tag{22}$$

$$\sum \|R_m - R_{m+1}\|_2^2 < \infty. \tag{23}$$

What about the convergence of $\sum \|R_m - R_{m+1}\|_2^p, 1 < p < 2$? The particular counter-example we built in proposition 3.1 does show that such a convergence is not sufficient to ensure the strong convergence of R_m to zero.

LEMMA 4.3. *The counter-example $\{R_m\}_{m\geq 1}$ to conjecture 2.2 built in proposition 3.1 complies, for all $p > 1$, with*

$$\sum \|R_m - R_{m+1}\|_2^p < \infty$$

Proof. We show that $\|R_m - R_{m+1}\|_2 \asymp m^{-1}$, which gives the result. Using the asymptotic rates (16) and (21) we get

$$\begin{aligned} \|R_m - R_{m+1}\|_2^2 &= \|R_m\|_2^2 - \|R_{m+1}\|_2^2 \\ &= \|R_m\|_\infty^2 \left(\langle S_m, S_m \rangle - \langle S_{m+1}, S_m \rangle^2 / \langle S_{m+1}, S_{m+1} \rangle \right) \\ &\asymp m^{-\alpha} m^{\alpha-2} = m^{-2}. \end{aligned}$$

∎

We know from [10] that, for any counter-example to conjecture 2.2, $\sum t_m/m < \infty$, where we use (7) and (8) to define

$$t_m := \frac{|\langle R_m, g_m \rangle|}{\sup_{g \in \mathcal{D}} |\langle R_m, g \rangle|} = \frac{\|R_m - R_{m+1}\|_2}{\sup_{g \in \mathcal{D}} |\langle R_m, g \rangle|}. \tag{24}$$

We would like to know whether we can find any counter-example to conjecture 2.2 such that $\sum t_m^2 = \infty$. It would show that $\sum t_m^2 = \infty$ is not a sufficient condition to ensure the convergence of weak greedy algorithms.

Let us start by studying the asymptotic behavior of t_m for any very weak greedy sequence $\{R_m\}_{m \geq 1}$. It is actually easy to see that $t_m = \|R_m - R_{m+1}\|_2 / \sup_p |\langle R_m, g_p \rangle|$ where $g_p = (R_p - R_{p+1}) / \|R_p - R_{p+1}\|_2$. For each p, one can write

$$\begin{aligned} |\langle R_m, g_p \rangle|^2 &= \frac{\left[\|R_m\|_\infty \|R_p\|_\infty \left(\langle S_m, S_p \rangle - \frac{\langle S_{p+1}, S_p \rangle}{\langle S_{p+1}, S_{p+1} \rangle} \langle S_m, S_{p+1} \rangle \right) \right]^2}{\|R_p - R_{p+1}\|_2^2} \\ &= \|R_m\|_\infty^2 \frac{\left(\langle S_m, S_p \rangle - \frac{\langle S_{p+1}, S_p \rangle}{\langle S_{p+1}, S_{p+1} \rangle} \langle S_m, S_{p+1} \rangle \right)^2}{\langle S_p, S_p \rangle - \frac{\langle S_{p+1}, S_p \rangle}{\langle S_{p+1}, S_{p+1} \rangle} \langle S_p, S_{p+1} \rangle} \end{aligned}$$

so that

$$t_m^2 = \frac{\|R_m - R_{m+1}\|_2^2}{\|R_m\|_\infty^2 K_m} \tag{25}$$

with $K_m = \sup_p K_{m,p}$ and

$$K_{m,p} = \frac{\left(\langle S_m, S_p \rangle - \frac{\langle S_{p+1}, S_p \rangle}{\langle S_{p+1}, S_{p+1} \rangle} \langle S_m, S_{p+1} \rangle \right)^2}{\langle S_p, S_p \rangle - \frac{\langle S_{p+1}, S_p \rangle}{\langle S_{p+1}, S_{p+1} \rangle} \langle S_p, S_{p+1} \rangle} \tag{26}$$

Let us now restrict the study to the very specific case of sequences $\{R_m\}_{m\geq 1}$ which associated sequence $\{S_m\}_{m\geq 1}$ can be written as $\{S(1/u_m)\}_{m\geq 1}$ (using definition (12)). We know that if $\{R_m\}_{m\geq 1}$ is a counter-example to conjecture 2.2, then $\|S(1/u_m)\|_2 \to \infty$, which shows that $u_m \to \infty$ thanks to (14). Let us show that it implies $\sum t_m^2 < \infty$.

LEMMA 4.4. *For any sequence $u_m \to \infty$, there exists $0 < \beta_1 < \beta_2 < \infty$, $\eta > 0$, and $m_0 \in \mathbb{N}$, such that for all $m \geq m_0$ there exists $p_m \in \mathbb{N}$ complying with*

$$u_{p_m} \in [\beta_1 u_m, \beta_2 u_m] \tag{27}$$

$$u_{p_m+1} \notin [(1-\eta)/(1+\eta)u_m, (1+\eta)/(1-\eta)u_m] \tag{28}$$

Proof. For every $\eta > 0$ and $x > 0$, denote $I_\eta(x)$ the interval $[(1-\eta)/(1+\eta)x, (1+\eta)/(1-\eta)x]$. Suppose the conclusion is false. The we know that for every $0 < \beta_1 < \beta_2 < \infty$, every $\eta > 0$ and every M, there exists $m \geq M$ such that for all p, $u_p \in [\beta_1 u_m, \beta_2 u_m] \Rightarrow u_{p+1} \in I_\eta(u_m)$. Let us take $\beta_1 = (1-\eta)/(1+\eta)$ and $\beta_2 = (1+\eta)/(1-\eta)$, for some arbitrary η. In this case, we know that, for some m, for all p, $u_p \in I_\eta(u_m) \Rightarrow u_{p+1} \in I_\eta(u_m)$. But we also know that $u_m \in I_\eta(u_m)$, so it becomes clear that by induction, $f(n) \in I_\eta(u_m)$ for all $n \geq m$, which is in contradiction with $u_m \to \infty$. ▮

PROPOSITION 4.1. *For every very weak greedy sequence which can be written as $R_m = \|R_m\|_\infty S(1/u_m)$ and is a counter-example to conjecture 2.2, we have*

$$t_m \asymp \|R_m - R_{m+1}\|_2 \,,$$

hence

$$\sum t_m^2 < \infty.$$

Proof. One can check that

$$K_{m,p} = u_m^2 \frac{2u_p - 1}{(u_p + u_m - 1)^2} \left(\frac{u_{p+1} - u_m}{u_{p+1} + u_m - 1} \right)^2 . \tag{29}$$

It is then easy to show that

$$K_m = \sup_{p,m} K_{m,p} \leq \frac{u_m^2}{2u_m - 1} \sim u_m/2. \tag{30}$$

Moreover, using lemma 4.4, one gets a sequence p_m such that

$$K_{m,p_m} \geq u_m^2 \frac{2\beta_1 u_m - 1}{((\beta_2+1)u_m - 1)^2}\eta^2 \sim \frac{2\beta_1\eta^2}{\beta_2^2} u_m \tag{31}$$

which shows that $K_m \asymp u_m$. From (14) this becomes $K_m \asymp \|S(1/u_m)\|_2^2$ thus, using (25), we get

$$t_m^2 \asymp \|R_m - R_{m+1}\|_2^2 \,/\, \left(\|R_m\|_\infty^2 \|S(1/u_m)\|_2^2\right) = \|R_m - R_{m+1}\|_2^2 \,/\, \|R_m\|_2^2$$

which finally gives $t_m^2 \asymp \|R_m - R_{m+1}\|_2^2$ using (11). We get the square summability of t_m from lemma 4.2. ▮

5. CONCLUSION

The family of potential counter-examples $R_m = \|R_m\|_\infty S(\epsilon_m)$ that we have built does not discard the possibility that $\sum t_m^2 = \infty$ might be a sufficient condition to ensure the convergence of a weak greedy algorithm. These counter-examples show that too weak a choice of g_m in a greedy algorithm can prevent the algorithm from converging. However, some of our numerical experiments [5] do show convergence of a very weak greedy algorithm in the multiscale time-frequency dictionary of Gaussian chirps $\mathcal{D}$, with an improvement of the speed of convergence compared to a pure greedy algorithm in the multiscale Gabor dictionary $\mathcal{D}^\star$. One simple reason for the convergence is that numerical experiments use finite dimensional data. But this does not explain the improvement in the rate of convergence. The reason for the good behaviour of this algorithm still has to be investigated. It may be due to the particular structure of $\mathcal{D}^\star$ and $\mathcal{D}$ and/or to the properties of the choice functional $R \mapsto g(R)$ which defines a particular set of very weak greedy sequences $R_{m+1} = R_m - \langle R_m, g(R_m)\rangle\, g(R_m)$.

ACKNOWLEDGMENT

The author wants to thank V. Temlyakov for his interest in this research and his very valuable remarks.

REFERENCES

1. F. Bergeaud. *Représentations adaptatives d'images numériques,* Matching Pursuit. PhD thesis, Ecole Centrale Paris, 1995.
2. F. Bergeaud and S. Mallat. Matching pursuit : adaptive representations of images and sounds. *Journal of Computational and Applied Mathematics*, 1996.

3. R.A. DeVore and V.N. Temlyakov. Some remarks on greedy algorithms. *Advances in Computational Mathematics*, 5:173–187, 1996.
4. R. Gribonval. *Approximations non-linéaires pour l'analyse de signaux sonores*. PhD thesis, Université Paris IX Dauphine, September 1999.
5. R. Gribonval. Fast ridge pursuit with a multiscale dictionary of Gaussian chirps. *IEEE Trans. Signal Process.(submitted)*, 2000. preprint at http://www.math.sc.edu/~imip/00papers/0002.ps.
6. P. J. Huber. Projection pursuit. *The annals of statistics*, 13(2):435–475, 1985.
7. L.K. Jones. On a conjecture of Huber concerning the convergence of PP-regression. *The Annals of Statistics*, 15:880–882, 1987.
8. S. Mallat and Z. Zhang. Matching pursuit with time-frequency dictionaries. *IEEE Trans. Signal Process.*, 41(12):3397–3415, December 1993.
9. L. Rejtö and G.G. Walter. Remarks on projection pursuit regression and density estimation. *Stochastic Analysis and Applications*, 10(2):213–222, 1992.
10. V.N. Temlyakov. Weak greedy algorithms. *Advances in Computational Mathematics*, 12(2,3):213–227, 2000.