

Sparse identification of nonlinear functions and parametric Set Membership optimality analysis

Original

Sparse identification of nonlinear functions and parametric Set Membership optimality analysis / Novara, C.. - In: IEEE TRANSACTIONS ON AUTOMATIC CONTROL. - ISSN 0018-9286. - 57:12(2012), pp. 3236-3241.
[10.1109/TAC.2012.2202051]

Availability:

This version is available at: 11583/2497027 since:

Publisher:

IEEE

Published

DOI:10.1109/TAC.2012.2202051

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

Sparse Identification of Nonlinear Functions and Parametric Set Membership Optimality Analysis

Carlo Novara

Abstract—Sparse identification can be relevant in the automatic control field to solve several problems for nonlinear systems such as identification, control, filtering, fault detection. However, identifying a maximally sparse approximation of a nonlinear function is in general an NP-hard problem. The common approach is to use relaxed or greedy algorithms that, under certain conditions, can provide sparsest solutions. In this technical note, a combined ℓ_1 -relaxed-greedy algorithm is proposed and conditions are given, under which the approximation derived by the algorithm is a sparsest one. Differently from other conditions available in the literature, the ones provided here can be actually verified for any choice of the basis functions defining the sparse approximation. A Set Membership analysis is also carried out, assuming that the function to approximate is a linear combination of unknown basis functions belonging to a known set of functions. It is shown that the algorithm is able to exactly select the basis functions which define the unknown function and to provide an optimal estimate of their coefficients.

Index Terms—Nonlinear system identification, set membership optimality, sparse approximation.

I. INTRODUCTION

Sparse approximation consists in approximating a function using “a few” basis functions properly selected within a “large” set. More precisely, a sparse approximation is a linear combination of “many” basis functions, but the vector of linear combination coefficients is sparse, i.e., it has only “a few” nonzero elements. Deriving a sparse approximation of an unknown function from a set of its values (possibly corrupted by noise) is here called sparse identification.

Sparse approximation/identification methods are relevant in many applications, such as compressive sensing [1]–[3], bioinformatics [4], computer vision [5], signal processing [6]–[8], source separation [9], denoising [10], linear regression [11], and regularization [12], and have analogies with Lasso [13] and Support Vector Machines [14]. Sparse approximation/identification methods have also been introduced in the automatic control field with interesting results [15]–[18]. In this field, sparsification methods might be effective to solve several problems such as regularization, basis function selection, regressor selection, nonlinear control design, “fast” online control implementation.

The sparsity of a vector is typically measured by the ℓ_0 quasi-norm, defined as the number of its nonzero elements. Sparse identification can thus be performed by looking for a coefficient vector of the basis function linear combination with a “small” ℓ_0 quasi-norm. However, the ℓ_0 quasi-norm is a nonconvex function and its minimization is in general an NP-hard problem. Two main approaches are commonly adopted to deal with this issue: convex relaxation and greedy algorithms [19]–[22]. In convex relaxation, a suitable convex function, e.g., the ℓ_1 norm, is minimized instead of the ℓ_0 quasi-norm [13], [20]–[22]. In greedy algorithms, the sparse solution is obtained iteratively [19]. The approaches based on ℓ_1 norm minimization give uniform guarantees for sparse recovery and have good stability properties, in the sense that they can effectively deal with compressible and noisy signals [19]. On the other hand, greedy approaches lack these strong properties, but

are very fast and efficient from a computational standpoint [19]. Hybrid algorithms can also be found in the literature, where ℓ_1 norm minimization and iterative procedures are combined to find a sparse solution [23]. A very interesting feature of all these approaches is that, under certain conditions, they provide sparsest solutions, i.e., solutions which also minimize the ℓ_0 quasi-norm [19]–[21], [24]. Although such conditions give an important theoretical motivation for using these relaxed/greedy approaches, their actual verification is often hard from a computational point of view. A remarkable contribution on this topic was provided in [24], [20], [22]. In [24], the conditions for a vector to be the sparsest solution can be actually verified when the basis functions are orthonormal or the union of incoherent orthonormal bases. In [20] and [22], the conditions are of easy verification when the basis functions have “small” mutual coherence.

In the present technical note, a combined ℓ_1 -relaxed-greedy algorithm is proposed for sparse identification and conditions are given, under which the algorithm provides a sparsest solution. Such conditions are actually verifiable for any choice of the basis functions and can be applied to any sparse identification algorithm. Considering a Set Membership framework [25]–[30], an optimality analysis is then carried out in order to show that the approximation obtained by the ℓ_1 -relaxed-greedy algorithm enjoys suitable optimality properties. The noise affecting the data is assumed bounded in norm and the unknown function is assumed to be a linear combination of basis functions whose coefficients have to be estimated. It is supposed that the basis functions are not known but belong to a “large” set of known functions. It is shown that the ℓ_1 -relaxed-greedy algorithm is able to select exactly the basis functions defining the unknown function and to provide an optimal estimate of their coefficients.

II. PROBLEM SETTING

Consider a nonlinear function f^0 defined by

$$y = f^0(x)$$

where $x \in X \subset \mathbb{R}^{n_x}$, $y \in \mathbb{R}$. Suppose that f^0 is not known but a set of noise-corrupted data $D = \{\tilde{x}_k, \tilde{y}_k\}_{k=1}^L$ is available, described by

$$\tilde{y}_k = f^0(\tilde{x}_k) + d_k, \quad k = 1, 2, \dots, L \quad (1)$$

where d_k is noise. Define the following parametrized function:

$$f^a(x) = \sum_{i=1}^N a_i \phi_i(x) = \phi(x) a$$

where $\phi(x) = [\phi_1(x), \phi_2(x), \dots, \phi_N(x)]$, $\phi_i : X \rightarrow \mathbb{R}$ are known basis functions, and $a = (a_1, a_2, \dots, a_N) \in \mathbb{R}^N$ is a coefficient vector. The notations $[\dots, \dots]$ and (\dots, \dots) are used here to indicate a row and a column vector, respectively.

1) *Problem 1:* From the data set D , identify a coefficient vector a such that:

- i) a is sparse;
- ii) the identification error $e(f^a) \doteq \|f^0 - f^a\|_p$, where $\|\cdot\|_p$ is a functional L_p norm, is “small”. ■

In this technical note, following a Set Membership framework [25]–[28], the noise sequence $d = (d_1, d_2, \dots, d_L)$ is assumed to be unknown but bounded, as follows:

$$\|d\|_2 \leq \mu \quad (2)$$

where $\|\cdot\|_2$ is the ℓ_2 vector norm.

Under the assumption (2), a solution to the sparse identification Problem 1 could be found by solving the following optimization problem:

$$a^0 = \arg \min_{a \in \mathbb{R}^N} \|a\|_0 \quad \text{subject to} \quad \|\tilde{y} - \Phi a\|_2 \leq \mu. \quad (3)$$

where $\|\cdot\|_0$ is the ℓ_0 quasi-norm, $\tilde{y} \doteq (\tilde{y}_1, \dots, \tilde{y}_L)$

$$\Phi \doteq \begin{bmatrix} \phi_1(\tilde{x}_1) & \cdots & \phi_N(\tilde{x}_1) \\ \vdots & \ddots & \vdots \\ \phi_1(\tilde{x}_L) & \cdots & \phi_N(\tilde{x}_L) \end{bmatrix} \\ = [\phi_1(\tilde{x}) \quad \cdots \quad \phi_N(\tilde{x})]$$

$\phi_i(\tilde{x}_k) \in \mathbb{R}$ and, with a slight abuse of notation, $\phi_1(\tilde{x}) \doteq (\phi_1(\tilde{x}_1), \dots, \phi_1(\tilde{x}_L)) \in \mathbb{R}^L$. The ℓ_0 quasi-norm is defined as the number of the vector elements which are not null, as follows:

$$\|a\|_0 \doteq \text{card}(\text{supp}(a))$$

where $\text{card}(\cdot)$ is the set cardinality, and $\text{supp}(a)$ is the support of a (the set of indices at which a is not null). The complement of $\text{supp}(a)$ (the set of indices at which a is null) is denoted by $\overline{\text{supp}}(a)$.

Thus, minimizing the ℓ_0 quasi-norm of a vector corresponds to minimizing the number of its nonzero elements, i.e., to maximizing its sparsity. On the other hand, the constraint $\|\tilde{y} - \Phi a\|_2 \leq \mu$ ensures that the identified coefficient vector is consistent with the measured data (1) and the prior assumption on noise (2). According to (3), a^0 is the sparsest coefficient vector consistent with the measured data and the prior assumptions.

Definition 1: For given \tilde{y} , Φ and μ , a coefficient vector is said *maximally sparse* if it is a solution of the optimization problem (3). ■

Unfortunately, the optimization problem (3) cannot be solved in general, since the ℓ_0 quasi-norm is a nonconvex function and its minimization is an NP-hard problem. Convex relaxation, greedy algorithms or hybrid methods can be adopted instead, which, under certain conditions, may provide maximally sparse solutions [19]–[21], [24]. However, the verification of these conditions is in general hard from a computational standpoint, and can actually be performed only for particular types of basis functions.

III. SPARSE IDENTIFICATION OF NONLINEAR FUNCTIONS

In this section, a combined ℓ_1 -relaxed-greedy algorithm, completely based on convex optimization, is proposed for solving the sparse identification Problem 1. Then, a theorem is presented, giving conditions verifiable for any kind of basis functions under which a vector (possibly obtained by the proposed identification algorithm) is maximally sparse.

For a matrix/vector $A \in \mathbb{R}^{K \times N}$, and a set of indices $\lambda = \{i_1, i_2, \dots, i_M\} \subset \{1, 2, \dots, N\}$, let us introduce the notation $A_\lambda \doteq [A_{i_1}, A_{i_2}, \dots, A_{i_M}]$, where A_j are the columns/elements of A . A similar notation is used for a partially ordered set $A = \{A_1, A_2, \dots, A_N\}$: $A_\lambda \doteq \{A_{i_1}, A_{i_2}, \dots, A_{i_M}\}$.

Algorithm 1

1) Solve the optimization problem

$$a^1 = \arg \min_{a \in \mathbb{R}^N} \|a\|_1 \quad \text{subject to} \quad \|\tilde{y} - \Phi a\|_2 \leq \mu. \quad (4)$$

2) Let $r(a^1) \doteq \{i_1, \dots, i_j : 0 < |a_{i_1}^1| \leq \dots \leq |a_{i_j}^1|\}$. Compute the coefficient vector a^* as follows:

```

for  $k = 0 : \text{card}(r(a^1))$ 
  if  $k = 0$ ,  $c^k = a^1$ 
  else  $c^k = \arg \min_{a \in \mathbb{R}^N} \|\tilde{y} - \Phi a\|_2$ 
    subject to  $a_i = 0$ 
     $i \in \overline{\text{supp}}(a^1) \cup r_\lambda(a^1)$ 
     $\lambda = \{1, \dots, k\}$ 
  end
  if  $\|\tilde{y} - \Phi c^k\|_2 \leq \mu$ ,  $a^* = c^k$ 
  else break
end
end. ■

```

The rationale behind Algorithm 1 can be explained as follows: In step 1, an optimization problem similar to (3) is solved, where the ℓ_0 quasi-norm is replaced by the ℓ_1 norm. The ℓ_1 norm is the convex envelope of the ℓ_0 quasi-norm, and its minimization usually yields a sparse vector a^1 [20]–[22]. However, it is not guaranteed that all the nonzero elements of a^1 are necessary to have $\|\tilde{y} - \Phi a^1\|_2 \leq \mu$. Then, in step 2, the nonzero elements of a^1 , ordered by increasing amplitude, are progressively set to zero. The algorithm stops when $\|\tilde{y} - \Phi c^k\|_2 > \mu$. The solution provided by step 2 is thus a vector a^* where the number of nonzero elements is further reduced with respect to the initial sparse solution a^1 .

Algorithm 1 provides an estimate a^* of a^0 , where a^0 is a maximally sparse coefficient vector, solution of the nonconvex optimization problem (3). The following theorem gives conditions ensuring that a^* has the same support as a^0 , and is thus maximally sparse as well. The theorem also provides a bound on the number of nonzero elements of a^* that are in excess with respect to a^0 .

Without loss of generality, assume that the columns of Φ are normalized: $\|\phi_i(\tilde{x})\|_2 = 1$, $i = 1, 2, \dots, N$. Let the vector c^{ver} be defined by the following convex optimization problem:

$$c^{ver} = \arg \min_{a \in \mathbb{R}^N} \|a\|_1 \quad (5)$$

$$\text{subject to} \quad \text{sign}(a_i^*) a_i \geq \eta(a^*), \quad \forall i \in \zeta \quad (6)$$

$$|a_i| < \eta(a^*), \quad \forall i \in \bar{\zeta} \quad (7)$$

$$\|\tilde{y} - \Phi a\|_2 \leq \mu \quad (8)$$

where $\eta(a) \doteq \min_{i \in \text{supp}(a)} |a_i|$, $\zeta \doteq \text{supp}(a^*)$, $\bar{\zeta} \doteq \overline{\text{supp}}(a^*)$. Consider the singular value decomposition

$$USV^T = \Phi \quad (9)$$

where $U \in \mathbb{R}^{L \times L}$ and $V^T \in \mathbb{R}^{N \times N}$ are unitary matrices ($U^T U = U U^T = I$ and $V V^T = V^T V = I$) and $S \in \mathbb{R}^{L \times N}$ is a matrix where the entries outside the main diagonal are all zero, whereas the diagonal entries are the singular values of Φ . Let V_n be the matrix composed of the columns of V corresponding to the zero singular values of Φ . Let $N_n \doteq \text{rank}(V_n) = N - \text{rank}(\Phi)$ and

$$\Gamma_{ij}(V_n) \doteq \max \left(\left| \sum_{k \in L A_{ij}} \nu_{ki} \nu_{kj} \right|, \left| \sum_{k \in S M_{ij}} \nu_{ki} \nu_{kj} \right| \right)$$

where ν_{ki} is the entry (k, i) of V_n , $L A_{ij} \doteq \{l_1, \dots, l_{\bar{q}} : \nu_{l_1 i} \nu_{l_1 j} \geq \dots \geq \nu_{l_{\bar{q}} i} \nu_{l_{\bar{q}} j} > 0\}$, $\bar{q} \leq 2m$ (c^{ver}) is the largest integer for which

$\nu_{\underline{q}i}\nu_{\underline{q}j} > 0$, $SM_{ij} \doteq \{l_1, \dots, l_q : \nu_{l_1i}\nu_{l_1j} \leq \dots \leq \nu_{l_qi}\nu_{l_qj} < 0\}$, $\underline{q} \leq 2m$ (c^{ver}) is the largest integer for which $\nu_{\underline{q}i}\nu_{\underline{q}j} < 0$. Define

$$s \doteq \min_i \left(1 - \sum_{j=1}^{N_n} \Gamma_{ij}(V_n) \right)$$

$$\xi(a) \doteq \beta \frac{|\delta(a)|_1 + |\delta(a)|_{2m}}{\sigma^2(\Phi)}, \quad \beta \doteq \begin{cases} 1, & \text{rank}(\Phi) = N \\ 1 + \frac{1}{s}, & \text{otherwise} \end{cases}$$

where $\sigma(\Phi)$ is the minimum nonzero singular value of Φ , $\delta(a) \doteq \tilde{y} - \Phi a$, $m \equiv m(a) \doteq \|a\|_0$, and

$$|w|_K \doteq \sqrt{\sum_{i \in I_K} (w^T \phi_i(\tilde{x}))^2}$$

where I_K is the set of the K largest inner products $|w^T \phi_i(\tilde{x})|$.

Theorem 1: Let a^0 and c^{ver} be the solutions of the optimization problems (3) and (5), respectively. Let a^* be the parameter vector identified by Algorithm 1. Define $N_e \doteq \|a^*\|_0 - \|a^0\|_0$ and $\lambda^s \doteq \{i : |c_i^{ver}| > \xi(c^{ver})\}$. Assume that:

- i) the constraint (8) is active;
- ii) either $\text{rank}(\Phi) = N$ or $s > 0$.

Then,

$$N_e \leq \overline{N}_e \doteq \|a^*\|_0 - \text{card}(\lambda^s). \quad (10)$$

Moreover, if

$$\xi(c^{ver}) < \eta(a^*) \quad (11)$$

then a^* is maximally sparse ($\overline{N}_e = 0$) and

$$\text{supp}(a^*) = \text{supp}(a^0). \quad (12)$$

Proof: As shown in [24], for any vector $a \in \mathbb{R}^N$ such that $\|a^0\|_0 \leq \|a\|_0$ and $\|\tilde{y} - \Phi a^0\|_2 \leq \|\tilde{y} - \Phi a\|_2$, the following inequality holds:

$$\|a^0 - a\|_\infty \leq \frac{|\delta(a)|_1 + |\delta(a)|_{2m}}{\sigma_{\min, 2m}^2} \quad (13)$$

where $m \doteq \|a\|_0$ and

$$\sigma_{\min, k}^2 \doteq \inf_{\|b\|_0 \leq k} \frac{\|\Phi b\|_2^2}{\|b\|_2^2}.$$

The explicit computation of $\sigma_{\min, k}^2$ is hard in many cases. Babel functions can be used in order to obtain a lower bound on $\sigma_{\min, k}^2$, [24]. However, the bound given by the Babel functions holds only when the basis functions ϕ_i have ‘‘small’’ inner products $\phi_i^T \phi_k$, $i \neq k$, [24]. Moreover, the evaluation of the Babel functions is easy when the basis functions are orthonormal or the union of incoherent orthonormal bases, but may be hard in other cases [24]. Here, a lower bound on $\sigma_{\min, k}^2$ is derived, which is easy to evaluate for any basis functions.

Consider the singular value decomposition (9). Let S_r be the matrix composed only of the rows and columns of S containing the nonzero singular values of Φ . Let V_r be the matrix composed of the columns of V corresponding to the nonzero singular values of Φ . Let U_r be the matrix composed of the columns of U corresponding to the nonzero singular values of Φ . Note that $V_r^T V_r = I$ and $U_r^T U_r = I$. Thus

$$U_r S_r V_r^T = U S V^T = \Phi,$$

$$\Phi^T \Phi = V_r S_r^T U_r^T U_r S_r V_r^T = V_r S_r^T S_r V_r^T$$

and $\|\Phi b\|_2^2 = b^T \Phi^T \Phi b = b^T V_r S_r^T S_r V_r^T b = \|S_r V_r^T b\|_2^2$.

If $\text{rank}(\Phi) = N$, then $\text{rank}(V_r) = \text{rank}(V_r^T) = N$, which implies that, for any $b \in \mathbb{R}^N$, a unique $c \in \mathbb{R}^N$ exists such that $b = V_r c$, $c = V_r^T b$, $\|b\|_2^2 = b^T b = c^T V_r^T V_r c = c^T c = \|c\|_2^2$, and $b = 0$ iff $c = 0$. It follows that

$$\inf_{\|b\|_0 \leq 2m} \frac{\|\Phi b\|_2^2}{\|b\|_2^2} = \inf_{\|b\|_0 \leq 2m} \frac{\|S_r V_r^T b\|_2^2}{\|b\|_2^2}$$

$$= \inf_{\substack{c = V_r^T b \\ \|b\|_0 \leq 2m}} \frac{\|S_r c\|_2^2}{\|c\|_2^2} \geq \inf_c \frac{\|S_r c\|_2^2}{\|c\|_2^2}. \quad (14)$$

In the case where $\text{rank}(\Phi) < N$, we have that $\Phi b = 0$ iff $b = V_n a$ for some $a \in \mathbb{R}^{N_n}$, where V_n is the matrix composed of the columns of V corresponding to the zero singular values of Φ (these columns span the null space of Φ). Any vector a satisfying the equation $b = V_n a$ must also satisfy the equation $\widehat{V}_n a = 0$, where $\widehat{V}_n = (\widehat{V}_{n1}, \dots, \widehat{V}_{nN})$ is the matrix with rows

$$\widehat{V}_{ni} = \begin{cases} 0, & \text{if } b_i \neq 0, \\ V_{ni}, & \text{otherwise,} \end{cases} \quad (15)$$

and V_{ni} are the rows of V_n . Letting w_{ij} be the entry (i, j) of the matrix $\widehat{V}_n^T \widehat{V}_n$, we have that

$$w_{ij} = \sum_{k \in \text{supp}(b)} \nu_{ki} \nu_{kj}$$

where ν_{ki} is the entry (k, i) of V_n . Since $V_n^T V_n = I$ and, consequently, $\sum_{k=1}^N \nu_{ki}^2 = 1$ and $\sum_{k=1}^N \nu_{ki} \nu_{kj} = 0$, $i \neq j$, the following bounds can be derived:

$$|w_{ij}| = \left| \sum_{k \in \text{supp}(b)} \nu_{ki} \nu_{kj} \right| = \left| \sum_{k=1}^N \nu_{ki} \nu_{kj} - \sum_{k \in \text{supp}(b)} \nu_{ki} \nu_{kj} \right|$$

$$\begin{cases} \geq 1 - \left| \sum_{k \in \text{supp}(b)} \nu_{ki}^2 \right|, & i = j \\ \leq \left| \sum_{k \in \text{supp}(b)} \nu_{ki} \nu_{kj} \right|, & i \neq j. \end{cases}$$

Then

$$|w_{ii}| - \sum_{j=1, j \neq i}^{N_n} |w_{ij}| \geq 1 - \left| \sum_{k \in \text{supp}(b)} \nu_{ki}^2 \right|$$

$$- \sum_{j=1, j \neq i}^{N_n} \left| \sum_{k \in \text{supp}(b)} \nu_{ki} \nu_{kj} \right| = 1 - \sum_{j=1}^{N_n} \left| \sum_{k \in \text{supp}(b)} \nu_{ki} \nu_{kj} \right|$$

$$\geq 1 - \sum_{j=1}^{N_n} \Gamma_{ij}(V_n) \geq s, \quad i = 1, \dots, N_n.$$

Since, by assumption, $s > 0$, we have that $|w_{ii}| > \sum_{j=1, j \neq i}^{N_n} |w_{ij}|$ for $i = 1, \dots, N_n$. According to the Gershgorin Circle Theorem, this condition ensures that the matrix $\widehat{V}_n^T \widehat{V}_n$ is nonsingular, implying that \widehat{V}_n is full column rank and that $\widehat{V}_n h = 0$ iff $h = 0$. This means that no vector $h \neq 0$ exists such that $b = V_n h$ and $\|b\|_0 \leq 2m$. In other words, if $\|b\|_0 \leq 2m$, then b cannot lie in the null space of V^T , and thus b must be of the form $b = V_n h + V_r c$, with $c \neq 0$. Clearly, the two vectors h and c satisfy the equality $\widehat{V}_n h + \widehat{V}_r c = 0$, where \widehat{V}_r is defined analogously to \widehat{V}_n in (15). It follows that $h = - \left(\widehat{V}_n^T \widehat{V}_n \right)^{-1} \widehat{V}_n^T \widehat{V}_r c$ and,

consequently, $b = \left(V_r - V_n \left(\widehat{V}_n^T \widehat{V}_n \right)^{-1} \widehat{V}_n^T \widehat{V}_r \right) c$. The following bound thus holds:

$$\begin{aligned} \|b\|_2 &= \left\| V_r - V_n \left(\widehat{V}_n^T \widehat{V}_n \right)^{-1} \widehat{V}_n^T \widehat{V}_r \right\|_2 \|c\|_2 \\ &\leq \left(\|V_r\|_2 + \left\| V_n \left(\widehat{V}_n^T \widehat{V}_n \right)^{-1} \widehat{V}_n^T \widehat{V}_r \right\|_2 \right) \|c\|_2. \end{aligned}$$

Consider now that $\|V_r\|_2, \|V_n\|_2, \|\widehat{V}_n^T\|_2, \|\widehat{V}_r\|_2 \leq 1$ and

$$\left\| \left(\widehat{V}_n^T \widehat{V}_n \right)^{-1} \right\|_2 = \frac{1}{\lambda_{\min} \left(\widehat{V}_n^T \widehat{V}_n \right)} \leq \frac{1}{s}$$

where $\lambda_{\min} \left(\widehat{V}_n^T \widehat{V}_n \right)$ is the smallest eigenvalue of $\widehat{V}_n^T \widehat{V}_n$, and the inequality is due to the Gershgorin Circle Theorem. The bound on $\|b\|_2$ becomes $\|b\|_2 \leq (1 + 1/s) \|c\|_2$. Then

$$\inf_{\|b\|_0 \leq 2m} \frac{\|\Phi b\|_2^2}{\|b\|_2^2} = \inf_{\|b\|_0 \leq 2m} \frac{\|S_r V_r^T b\|_2^2}{\|b\|_2^2} \geq \left(\frac{s}{s+1} \right) \inf_c \frac{\|S_r c\|_2^2}{\|c\|_2^2}. \quad (16)$$

Now, a standard result in matrix theory is that

$$\inf_c \frac{\|S_r c\|_2^2}{\|c\|_2^2} = \underline{\sigma}^2(S_r)$$

where $\underline{\sigma}(S_r)$ is the minimum singular value of S_r . Since $\underline{\sigma}(S_r)$ coincides with $\underline{\sigma}(\Phi)$, the minimum nonzero singular value of Φ , from (14) and (16) we have that

$$\inf_{\|b\|_0 \leq 2m} \frac{\|\Phi b\|_2^2}{\|b\|_2^2} \geq \beta^{-1} \underline{\sigma}^2(\Phi)$$

and, from (13)

$$\|a^0 - a\|_\infty \leq \beta \frac{|\delta(a)|_1 + |\delta(a)|_{2m}}{\underline{\sigma}^2(\Phi)} \doteq \xi(a). \quad (17)$$

Note that this inequality does not necessarily hold for $a = a^*$. Consider then the vector c^{ver} , solution of the optimization problem (5). Since the constraint (8) is active, we have that $\|\tilde{y} - \Phi a^0\|_2 \leq \|\tilde{y} - \Phi c^{ver}\|_2 = \mu$. Moreover, $\|a^0\|_0 \leq \|c^{ver}\|_0$ since a^0 is the sparsest vector satisfying the constraint $\|\tilde{y} - \Phi a^0\|_2 \leq \mu$. Then, inequality (17) holds with $a = c^{ver}$, and this implies that

$$|a_i^0 - c_i^{ver}| \leq \xi(c^{ver}), \quad i = 1, 2, \dots, N.$$

This shows that, if $|c_i^{ver}| > \xi(c^{ver})$, then $a_i^0 \neq 0$ and, consequently

$$\lambda^s \doteq \{i : |c_i^{ver}| > \xi(c^{ver})\} \subseteq \text{supp}(a^0).$$

It follows that $\|a^0\|_0 \geq \text{card}(\lambda^s)$, which yields (10).

Now, the constraints (6) and (7) imply that

$$\text{supp}(a^*) = \{i : |c_i^{ver}| \geq \eta(a^*)\}.$$

Moreover, if condition (11) holds, then

$$\begin{aligned} \{i : |c_i^{ver}| \geq \eta(a^*)\} &\subseteq \{i : |c_i^{ver}| > \xi(c^{ver})\} \\ &\doteq \lambda^s \subseteq \text{supp}(a^0). \end{aligned}$$

It follows that $\text{supp}(a^*) \subseteq \text{supp}(a^0)$. But a^0 is the sparsest vector satisfying the constraint $\|\tilde{y} - \Phi a^0\|_2 \leq \mu$. The relation (12) follows. ■

Remark 1: Verification of the condition (11) in Theorem 1 is computationally simple for any matrix Φ . Indeed, this verification basically requires to evaluate $|\delta(c^{ver})|_1, |\delta(c^{ver})|_2, \underline{\sigma}^2(\Phi)$ and to solve the convex optimization problem (5). All these operations can be easily performed in polynomial time. It must be remarked that the condition (11) has been derived from the condition (12) of Corollary 1 in [24], whose verification is simple when the basis functions are orthonormal or the union of incoherent orthonormal bases. Assumptions (i) and (ii) are technical conditions whose verification is straightforward. When $\text{rank}(\Phi) < N$, assumption (ii) needs that $s > 0$. This in turn requires that the matrix \widehat{V}_n defined in (15) is such that $\widehat{V}_n^T \widehat{V}_n$ is non-singular for all b with $\|b\|_0 \leq 2m$ (c^{ver}), see the proof of Theorem 1. Since \widehat{V}_n is obtained from V_n by setting to zero at most $2m$ (c^{ver}) rows and $V_n^T V_n = I$, assumption (ii) can be met for example when m (c^{ver}) $\ll N$, i.e., when c^{ver} is sufficiently sparse. In this situation, $\widehat{V}_n^T \widehat{V}_n \simeq V_n^T V_n, \Gamma_{ij}(V_n) \simeq 0, s \simeq 1$ and $\beta \simeq 2$, indicating that the maximum sparsity condition (11) is in general more difficult to be met when $\text{rank}(\Phi) < N$.

Remark 2: If the conditions of Theorem 1 are not satisfied and thus the vector a^* derived by Algorithm 1 is not ensured to be maximally sparse, it is anyway of interest to check if a^* is the sparsest solution among all the vectors with $a_i = 0, i \notin \zeta$, satisfying the constraint $\|\tilde{y} - \Phi a\|_2 \leq \mu$. This can be accomplished by applying Theorem 1 to the reduced vector a_ζ^* and matrix Φ_ζ .

Remark 3: The conditions of Theorem 1 are independent on the particular algorithm used to obtain the sparse solution, and therefore can be used to perform the maximal sparsity check for any possible vector.

IV. PARAMETRIC SET MEMBERSHIP OPTIMALITY ANALYSIS

In Section III, an ℓ_1 -relaxed-greedy algorithm has been proposed, able to derive a sparse approximation of the function f^0 , thus allowing the accomplishment of the requirement (i) of the identification Problem 1. In this section, considering a novel optimality notion developed within a Set Membership framework, this approximation is shown to have “small” identification error, thus allowing us to satisfy also the requirement (ii) of Problem 1.

The noise sequence $d = (d_1, d_2, \dots, d_L)$ in (1) is assumed to be bounded according to (2). The function f^0 is assumed to be parametrized as

$$f^0(x) = \sum_{i=1}^N a_i^0 \phi_i(x) \quad (18)$$

where ϕ_i are basis functions and $a^0 = (a_1^0, a_2^0, \dots, a_N^0) \in \mathbb{R}^N$ is a sparse unknown parameter vector, solution of (3). In the present technical note, this assumption is referred to as “parametric f^0 assumption”.

Under these assumptions, the identification Problem 1 reduces to finding an estimate \widehat{a} of a^0 such that:

- (i) $\text{supp}(\widehat{a}) = \text{supp}(a^0)$;
- (ii) the parametric error

$$e^{par}(\widehat{a}) \doteq \|a^0 - \widehat{a}\|_2$$

is “small”.

While Theorem 1 gives a condition under which $\text{supp}(\widehat{a}) = \text{supp}(a^0)$, no exact knowledge of $e^{par}(\widehat{a})$ is available, being a^0 unknown. However, from (18) and (2), we have that $a^0 \in FPS$, where FPS is the Feasible Parameter Set.

Definition 2: The Feasible Parameter Set is

$$FPS \doteq \{a \in \mathbb{R}^N : \text{supp}(a) = \text{supp}(a^0), \|\tilde{y} - \Phi a\|_2 \leq \mu\}$$

where μ is the noise bound assumed in (2) and a^0 is the ‘‘true’’ parameter vector in (18). ■

According to this definition, FPS is the set of all parameter vectors consistent with prior assumptions and data, and is the smallest set guaranteed to contain a^0 . Thus, a tight bound on $e^{par}(\hat{a})$ is given by the following worst-case error.

Definition 3: Worst-case parametric error of an estimate \hat{a} :

$$EP(\hat{a}) \doteq \sup_{a \in FPS} \|a - \hat{a}\|_2. \quad \blacksquare$$

This leads to the notion of optimal estimate, defined as an estimate which minimizes the worst-case parametric error.

Definition 4: An estimate a^{op} is optimal if

$$EP(a^{op}) = \inf_{\hat{a} \in \mathbb{R}^N} EP(\hat{a}). \quad \blacksquare \quad (19)$$

The following result shows that, under the assumptions of Theorem 1, the estimate a^* identified by Algorithm 1 from the data set D described in (1) is optimal. This result also gives an explicit expression of the worst-case parametric error.

Theorem 2: Assume that:

- i) The noise affecting the measurements $\{\tilde{x}_k, \tilde{y}_k\}_{k=1}^L$ is bounded according to (2).
- ii) The function f^0 to identify is parametrized according to (18), where a^0 is a sparse unknown parameter vector, solution of (3).
- iii) $\text{rank}(\Phi_\zeta) = \text{card}(\zeta)$, where $\zeta = \text{supp}(a^*)$.
- iv) Assumptions (i) and (ii) of Theorem 1 hold and $\xi(c^{ver}) < \eta(a^*)$, where c^{ver} is a solution of the optimization problem (5) and a^* is the parameter vector identified by Algorithm 1.

Then:

- i) a^* is an optimal estimate of a^0 .
- ii) The worst-case parametric error of a^* is given by

$$EP(a^*) = \bar{\sigma}(\Phi_\zeta^\dagger) \sqrt{\mu^2 - \|\delta(a^*)\|_2^2}$$

where $\Phi_\zeta^\dagger \doteq (\Phi_\zeta^T \Phi_\zeta)^{-1} \Phi_\zeta^T$ and $\bar{\sigma}(\Phi_\zeta^\dagger)$ is the maximum singular value of Φ_ζ^\dagger .

Proof: From Theorem 1, if $\xi(c^{ver}) < \eta(a^*)$, then $\text{supp}(a^*) = \text{supp}(a^0)$, where a^0 is a solution of the optimization problem (3). Moreover, the condition $\|\delta(c^k)\|_2 \leq \mu$ in step 2 of Algorithm 1 implies that $\|\tilde{y} - \Phi a^*\|_2 \leq \mu$. It follows that $a^* \in FPS$.

Now, consider an optimal estimate a^{op} . Then

$$\begin{aligned} EP(a^{op}) &= \inf_{\hat{a} \in \mathbb{R}^N} \sup_{a \in FPS} \|a - \hat{a}\|_2 \\ &= \inf_{\hat{a} \in \mathbb{R}^N} \sup_{\substack{a \in \mathbb{R}^N \\ \text{supp}(a) = \text{supp}(a^0) \\ \|\tilde{y} - \Phi a\|_2 \leq \mu}} \|a - \hat{a}\|_2 \\ &= \inf_{\hat{a} \in \mathbb{R}^m(a^*)} \sup_{a \in \mathbb{R}^m(a^*): \|\tilde{y} - \Phi_\zeta a\|_2 \leq \mu} \|a - \hat{a}\|_2. \end{aligned}$$

where $\zeta = \text{supp}(a^*) = \text{supp}(a^0)$. In [26], it is shown that the least-squares estimate

$$a_\zeta^{ls} = \arg \min_{a \in \mathbb{R}^m(a^*)} \|\tilde{y} - \Phi_\zeta a\|_2$$

where $m \equiv m(a^*)$, is an optimal estimate

$$a_\zeta^{ls} = \arg \inf_{\hat{a} \in \mathbb{R}^m} \sup_{a \in \mathbb{R}^m(a^*): \|\tilde{y} - \Phi_\zeta a\|_2 \leq \mu} \|a - \hat{a}\|_2.$$

The corresponding worst-case estimation error is given by

$$\sup_{a \in \mathbb{R}^m: \|\tilde{y} - \Phi_\zeta a\|_2 \leq \mu} \|a - a_\zeta^{ls}\|_2 = \bar{\sigma}(\Phi_\zeta^\dagger) \sqrt{\mu^2 - \|\delta(a_\zeta^{ls})\|_2^2}$$

where $\delta(a_\zeta^{ls}) = \tilde{y} - \Phi_\zeta a_\zeta^{ls}$, see [26].

On the other hand, a^* is the solution of the optimization problem in step 2 of Algorithm 1, that is a least-squares problem restricted to the elements of a^* whose indexes belong to $\text{supp}(a^0)$ (the other elements are null). It follows that $a_\zeta^* = a_\zeta^{ls}$ and, consequently

$$\begin{aligned} EP(a^*) &= \inf_{\hat{a} \in \mathbb{R}^N} \sup_{a \in FPS} \|a - \hat{a}\|_2 \\ &= \bar{\sigma}(\Phi_\zeta^\dagger) \sqrt{\mu^2 - \|\delta(a^*)\|_2^2} \end{aligned}$$

where $\delta(a^*) = \delta(a_\zeta^*) = \tilde{y} - \Phi_\zeta a_\zeta^* = \delta(a_\zeta^{ls})$. The two claims of the theorem are thus proven. ■

Remark 4: The optimality notion (19) is stronger than the ‘‘standard’’ worst-case optimality notion. Indeed, the ‘‘standard’’ feasible parameter set is defined as

$$\overline{FPS} \doteq \{a \in \mathbb{R}^N : \|\tilde{y} - \Phi a\|_2 \leq \mu\}$$

see [26]. The ‘‘standard’’ optimal estimate is consequently defined as an estimate a^{sc} such that

$$\overline{EP}(a^{sc}) = \inf_{\hat{a} \in \mathbb{R}^N} \sup_{a \in \overline{FPS}} \|a - \hat{a}\|_2.$$

Since $a^0 \in FPS \subseteq \overline{FPS}$, it follows that $EP(a^{op}) \leq \overline{EP}(a^{sc})$, showing that an optimal estimate a^{op} has better estimation accuracy (in a worst-case sense) with respect to a ‘‘standard’’ optimal estimate. Note also that the classical least squares estimate $a^{ls} = \arg \min_{a \in \mathbb{R}^N} \|\tilde{y} - \Phi a\|_2$ is a ‘‘standard’’ optimal estimate of a^0 [26], and thus $EP(a^{op}) \leq \overline{EP}(a^{ls})$.

Remark 5: Theorem 2 shows that, besides giving optimal estimates, the presented ℓ_1 -relaxed-greedy algorithm is able to perform *exact basis function selection*, i.e., to select, within a ‘‘large’’ set of basis functions, the ones defining the unknown function f^0 . It must be remarked that exact selection is here performed for a finite number of data. (See [31] for an asymptotic analysis, regarding the case where the number of data tends to infinity).

V. CONCLUSIONS

A combined ℓ_1 -relaxed-greedy algorithm for sparse identification of nonlinear functions has been proposed. Conditions for maximum sparsity have been provided, which can be verified for any choice of the basis functions. It has been shown that the algorithm is able to perform exact basis function selection and has improved optimality properties over standard least-squares and other sparse techniques. An application of the proposed algorithm to a problem of practical interest regarding the identification of vehicle vertical dynamics can be found in [30].

REFERENCES

- [1] D. Donoho, ‘‘Compressed sensing,’’ *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [2] E. Candes, J. Romberg, and T. Tao, ‘‘Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,’’ *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [3] E. Candes and T. Tao, ‘‘Near-optimal signal recovery from random projections: Universal encoding strategies?,’’ *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

- [4] Z. Liu, S. Lin, and M. Tan, "Sparse support vector machines with l_p penalty for biomarker identification," *IEEE/ACM Trans. Comput. Biol. and Bioinform.*, vol. 7, no. 1, pp. 100–107, Jan. 2010.
- [5] N. Ozay, M. Sznaier, and O. Camps, "Sequential sparsification for change detection," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2008, pp. 1–6.
- [6] R. Gribonval and E. Bacry, "Harmonic decomposition of audio signals with matching pursuit," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 101–111, Jan. 2003.
- [7] P. Frossard, P. Vandergheynst, R. Figueras i Ventura, and M. Kunt, "A posteriori quantization of progressive matching pursuit streams," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 525–535, Feb. 2004.
- [8] J. Tropp, J. Laska, M. Duarte, J. Romberg, and R. Baraniuk, "Beyond Nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Trans. Inform. Theory*, vol. 56, no. 1, pp. 520–544, Jan. 2010.
- [9] M. Zibulevsky and B. Pearlmutter, "Blind source separation by sparse decomposition in a signal dictionary," *Neural Comput.*, vol. 13, no. 4, pp. 863–882, 2001.
- [10] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1999.
- [11] A. Miller, *Subset Selection in Regression*. London, U.K.: Chapman & Hall, 2002.
- [12] I. Daubechies, M. Debrise, and C. D. Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, pp. 1413–1457, 2004.
- [13] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *Royal. Statist. Soc B.*, vol. 58, no. 1, pp. 267–288, 1996.
- [14] F. Girosi, "An equivalence between sparse approximation and support vector machines," *Neur. Comput.*, vol. 10, no. 6, pp. 1455–1480, 1998.
- [15] Y. Chen, Y. Gu, and A. Hero, "Sparse LMS for system identification," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2009, pp. 3125–3128.
- [16] C. Feng, C. M. Lagoa, and M. Sznaier, "Hybrid system identification via sparse polynomial optimization," in *Proc. American Control Conf. (ACC)*, Jun. 2010, pp. 160–165.
- [17] B. M. Sanandaji, T. L. Vincent, and M. B. Wakin, "Concentration of measure inequalities for compressive toeplitz matrices with applications to detection and system identification," in *Proc. 49th IEEE Conf. Decision and Control (CDC)*, Dec. 2010, pp. 2922–2929.
- [18] N. Ozay, M. Sznaier, C. Lagoa, and O. Camps, "A sparsification approach to set membership identification of switched affine systems," *IEEE Trans. Autom. Control*, vol. 57, no. 3, pp. 634–648, Mar. 2012.
- [19] J. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.
- [20] J. Fuchs, "Recovery of exact sparse representations in the presence of bounded noise," *IEEE Trans. Inform. Theory*, vol. 51, no. 10, pp. 3601–3608, Oct. 2005.
- [21] J. Tropp, "Just relax: Convex programming methods for identifying sparse signals in noise," *IEEE Trans. Inform. Theory*, vol. 52, no. 3, pp. 1030–1051, Mar. 2006.
- [22] D. Donoho, M. Elad, and V. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 6–18, Jan. 2006.
- [23] E. J. Candes, M. B. Wakin, and S. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. and Applic.*, vol. 14, no. 5, pp. 877–905, 2008.
- [24] R. Gribonval, R. Figueras i Ventura, and P. Vandergheynst, "A simple test to check the optimality of a sparse signal approximation," *Signal Process.*, vol. 86, no. 3, pp. 496–510, 2006.
- [25] M. Milanese and A. Vicino, "Optimal algorithms estimation theory for dynamic systems with set membership uncertainty: An overview," *Automatica*, vol. 27, pp. 997–1009, 1991.
- [26] M. Milanese, "Properties of least squares in set membership identification," *Automatica*, vol. 31, no. 2, pp. 327–332, 1995.
- [27] A. Kurzhanski and V. Veliov, *Modeling Techniques for Uncertain Systems*. Basel, Switzerland: Birkhäuser, 1994.
- [28] J. Chen and G. Gu, *Control-Oriented System Identification: An H_∞ Approach*. New York: Wiley, 2000.
- [29] M. Milanese and C. Novara, "Model quality in identification of nonlinear systems," *IEEE Trans. Autom. Control*, vol. 50, no. 10, pp. 1606–1611, Oct. 2005.
- [30] C. Novara, "Sparse identification of nonlinear functions and parametric set membership optimality analysis," in *Proc. American Control Conf.*, San Francisco, CA, 2011.
- [31] C. Novara, K. Hsu, M. Milanese, and K. Poolla, "Parametric identification of structured nonlinear systems," *Automatica*, vol. 7, no. 4, pp. 711–721, 2011.

Convex Dwell-Time Characterizations for Uncertain Linear Impulsive Systems

Corentin Briat and Alexandre Seuret

Abstract—New sufficient conditions for the characterization of dwell-times for linear impulsive systems are proposed and shown to coincide with continuous decrease conditions of a certain class of looped-functionals, a recently introduced type of functionals suitable for the analysis of hybrid systems. This approach allows to consider Lyapunov functions that evolve nonmonotonically along the flow of the system in a new way, broadening then the admissible class of systems which may be analyzed. As a byproduct, the particular structure of the obtained conditions makes the method is easily extendable to uncertain systems by exploiting some convexity properties. Several examples illustrate the approach.

Index Terms—Dwell-time, impulsive systems, looped-functionals, robustness, stability.

I. INTRODUCTION

Impulsive systems [1]–[6] are an important class of hybrid systems admitting discontinuities in their trajectories at certain time-instants. They arise in several fields like epidemiology [7], [8], sampled-data and networked control systems [9], etc. Among the wide class of impulsive dynamical systems, we may distinguish systems whose impulse-times depend on the system state and those for which they are external and only time-dependent. Linear systems of the latter class may be represented in the following form:

$$\begin{aligned} \dot{x}(t) &= Ax(t), \quad t \neq t_k, \quad k \in \{1, 2, \dots\} \\ x^+(t) &= Jx(t), \quad t = t_k, \quad k \in \{1, 2, \dots\} \\ x(t_0) &= x_0 \end{aligned} \quad (1)$$

where $x, x_0 \in \mathbb{R}^n$ are the system state and the initial condition respectively. The state $x(t)$ is assumed to be left-continuous with $x^+(t) := \lim_{s \uparrow t} x(s)$ and the matrices A and J may be uncertain. The sequence $\{t_1, t_2, t_3, \dots\}$ is a strictly increasing sequence of impulse times in (t_0, ∞) for some initial time t_0 . The sequence is assumed to be finite, or infinite and unbounded. The distance between two consecutive impulse times is denoted by $T_k := t_{k+1} - t_k$ with the additional assumption that $\varepsilon < T_k$, for some $\varepsilon > 0$, which excludes the possible existence of a finite accumulation point. Given a set $\mathcal{S} \subseteq (0, \infty)$, we define $\mathbb{1}_{\mathcal{S}} := \{\{t_1, t_2, \dots\} : t_{k+1} - t_k \in \mathcal{S}, \quad k \in \mathbb{N}\}$ to characterize the sequence of impulse times.

C. Briat was previously with the ACCESS Linnaeus Centre, KTH, Stockholm, Sweden. He is now with the Swiss Federal Institute of Technology–Zürich (ETH-Z), Department of Biosystems Science and Engineering (D-BSSE), Mattenstrasse 26, 4058 Basel, Switzerland (e-mail: corentin@briat.info; corentin.briat@bsse.ethz.ch).

A. Seuret is with the CNRS, Control Systems Department GIPSA Lab INP BP 46, ENSIEG Domaine Universitaire, 38400 Saint Martin d’Hères, France (e-mail: aseuret@gmail.com, alexandre.seuret@gipsa-lab.genoble-inp.fr).