Signal Processing 151 (2018) 99–106

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Short communication

Data-driven parameter selection for activity estimation in nuclear spectroscopy

Tom Trigano[a,*], Yann Sepulcre[b]

[a] Department of Electrical Engineering, SCE, Ashdod Campus, Israel
[b] Unit of Mathematics, SCE, Ashdod Campus, Israel

ARTICLE INFO

Article history:
Received 20 December 2017
Revised 27 April 2018
Accepted 6 May 2018
Available online 7 May 2018

Keywords:
Compressive sensing
Sparse reconstruction
Point processes
Model selection
Adaptive methods
Spectroscopic signal processing

ABSTRACT

One of the main objectives of nuclear spectroscopy is the estimation of the counting rate of unknown radioactive sources. Recently an algorithm based on a sparse reconstruction of the time signal was proposed by the authors to estimate precisely this counting rate, and computable bounds were obtained to quantify the performances. This approach, based on a post-processed approach of a non-negative sparse regression of the time signal, relies on user-defined parameters which are difficult to set up automatically in practice. This paper presents a data-driven strategy to select the underlying parameters. The parameter controlling the sparsity of the regressor is chosen based on cross-validation, while we introduce a new, entropy-based, criterion to select the threshold parameters. Results obtained on simulations illustrate the efficiency of the proposed approach.

1. Introduction

The analysis and processing of signals based on underlying point processes have received a great amount of interest in numerous engineering applications, ranging from biomedical engineering [1] to teletraffic data analysis [2]. Among other aims, practitioners from the field of nuclear spectroscopy often wish to estimate the activity of an unknown radioactive source as well as its contents [3]. Though numerous techniques have been developed in the field of nuclear science for activity estimation [4,5], they are usually limited to low and medium activities.

Due to the recent developments in compressive sensing and to the inherent sparsity of the data at hand in these applications, methods to estimate the activity have been suggested by the authors [6–8], which rely on a preliminary sparse reconstruction of the observed signal, and a thresholding of the obtained regressor. Theoretical and practical results [6,7] obtained on real life signals both show the robustness and usefulness of these approaches, even in the case of very high counting rates. However, one of their shortcomings is their dependency on both a sparsity and a thresholding parameter, which are difficult to tune without preliminary knowledge of the source.

This paper suggests an efficient, fully data-driven, strategy to choose these parameters while guaranteeing a satisfactory estimation of the activity. The main novelty lies in the choice of the threshold parameter, which is based on a local entropy measure. Sparsity of signal representation has indeed been related to several measures of entropy. In [9], the authors introduced some entropy measure associated to a (fixed) vector and several possible orthogonal decompositions, which has to be minimized to obtain some best possible representation. Entropy notions in signal representation can be also combined with an optimal threshold choice, as in [10–12] where iterative schemes are derived for its computation. The main difference between the former approaches and ours lies in applying an entropy-based criterion locally, rather than on the whole signal. The rest of the paper is organized as follows: Section 2 recalls the proposed approach for activity estimation based on spectroscopic data in the homogeneous case. We illustrate empirically that standard cross-validation to choose the thresholding parameter is not a valid strategy. We then present our method for choosing the parameters automatically in Section 3. While the sparsity parameter can be chosen using regular (or modified) cross-validation, the threshold parameter is chosen using a novel criterion. This criterion is based on the entropy related to a non-negative version of the Least Absolute Shrinkage and Selection Operator (LASSO) regressor. Results on simulations presented in Section 4 illustrate the validity of the proposed approach.

* Corresponding author.
E-mail address: thomast@sce.ac.il (T. Trigano).

https://doi.org/10.1016/j.sigpro.2018.05.006
0165-1684/

2. Overview of the activity estimation algorithm

We observe a signal on the time interval [0; T] uniformly sampled on $\mathcal{T} = \{0 = t_0, t_1, t_2, \ldots, t_{N-1} = T\}$ with sampling period Δt,

$$y_i = \sum_{n=1}^{M} E_n \Phi_n(t_i - T_n) + \varepsilon_i, \;\; 0 \le i \le N-1\,, \tag{1}$$

where $\{T_n, n \ge 1\}$ denotes the times of interaction between incoming photons and the detectors, $\{E_n,\ 1 \le n \le M\}$ is a sequence of independent and identically distributed (iid) random variables representing the photonic energies with unknown probability density function, $\{\Phi_n, 1 \le n \le M\}$ is a sequence of functions which characterizes the electric pulse shapes generated by the photons, and $\{\varepsilon_i,\ 0 \le i \le N-1\}$ is a sequence of iid Gaussian random variables with zero mean and variance σ^2 representing the additional noise of the input signal. We assume that $\{T_n, n \ge 1\}$ is a sample path of a Poisson process with unknown, possibly time-varying, activity λ. The estimation methods for either constant or time-varying λ rely on a preliminary thresholded sparse regressor, which is now detailed.

2.1. Activity estimation based on thresholded LASSO

As noted in [6], neither the energies nor the shapes are known. However, due to the physics explaining the recorded shapes in gamma spectroscopy experiments with Germanium detectors, it is relevent to represent the signal defined in (1) with a dictionary of truncated gamma shapes in order to obtain a reliable model. We can assume that an individual pulse contains a rapidly increasing part followed by an exponential decay. We define a set of p pairs of such parameters by $\boldsymbol{\theta} = \{(\theta_1^{(s)}, \theta_2^{(s)});\ s = 1, 2, \ldots, p\}$. For all $s = 1, \ldots, p$, we define the following pulse shapes:

$$\Gamma_s(t) = c_s\, t^{\theta_1^{(s)}} \exp(-\theta_2^{(s)}\, t)\, \mathbf{1}(0 < t \le \tau\, \Delta t), \tag{2}$$

where τ is a positive constant integer defining the common support of the pulse shapes, $\mathbf{1}(0 < t \le \tau\,\Delta t)$ is the indicator function, being equal to 1 when $0 < t \le \tau\,\Delta t$ and to 0 otherwise, and c_s is a normalizing constant chosen so that $\frac{1}{N}\sum_{i=0}^{N-1} \Gamma_s(t_i)^2 = 1$. Obviously, the set $\boldsymbol{\theta}$ should be chosen carefully, in order to be well suited for $\mathbf{y}$. For example, in the applications on real signals performed in [6,8], it was empirically found that choosing $\theta_1^{(s)} = \theta_2^{(s)} \in [0.5, 1.5]$ provided satisfactory results, and we used this choice in this paper as well. Accordingly, we define the following $N \times p$ matrix $\mathbf{A}_k$ whose columns are sampled versions of the previously defined pulse shapes, translated by t_k, $\mathbf{A}_k = [\,\Gamma_s(t_i - t_k)\,]_{0 \le i \le N-1,\, 1 \le s \le p}$. Intuitively, $\mathbf{A}_k$ can be understood as a subdictionary representing the electrical pulses starting from the same sampling time t_k. A global dictionary $\mathbf{A}$ of possible pulses occurring at different times is then defined by regrouping the $\mathbf{A}_k$'s altogether: $\mathbf{A} = [\mathbf{A}_0 \;\; \mathbf{A}_1 \;\; \cdots \;\; \mathbf{A}_{N-1}]$. Given any vector $\boldsymbol{\beta}$ of size pN, it is thus naturally split into blocks of size p: $\boldsymbol{\beta} = [\boldsymbol{\beta}_0^T; \ldots; \boldsymbol{\beta}_{N-1}^T]^T$, where for all i, $\boldsymbol{\beta}_i \in \mathbb{R}^p$. Throughout the paper, we assume that the discretization error can be neglected (a full discussion on the influence of sampling can be found in [13]). Thus, the model investigated in the rest of the paper is

$$\mathbf{y} = \mathbf{A}\boldsymbol{\beta} + \boldsymbol{\varepsilon} = \sum_{k=0}^{N-1} \mathbf{A}_k \boldsymbol{\beta}_k + \boldsymbol{\varepsilon}, \tag{3}$$

where $\mathbf{y} = [y_0, y_1, \ldots, y_{N-1}]^T$ and $\boldsymbol{\varepsilon} = [\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{N-1}]^T$ represent the recorded signal and the additive noise respectively and $\boldsymbol{\beta}$ is a sparse non-negative vector whose non-null entries are related to photon arrival times. Our objective is therefore to estimate λ from (3), and at first to find a sparse estimate of $\boldsymbol{\beta}$, so that its non-zero entries provide relevant information on the active bins present in the signal. This sparse estimate is obtained by solving a non-negative version of the LASSO, that is

$$\widehat{\boldsymbol{\beta}}(r) = \arg\min_{\boldsymbol{\beta} \ge 0} \left\{ \|\mathbf{y} - \mathbf{A}\boldsymbol{\beta}\|_2^2 + r\|\boldsymbol{\beta}\|_1 \right\}. \tag{4}$$

Introduced in [14], (4) provides a regressor $\widehat{\boldsymbol{\beta}}(r)$ whose sparsity is controlled by the parameter r (the larger, the sparser). Note than many alternatives to LASSO exist in the literature, e.g. the sparse Bayesian learning of Wipf and Nagarajan [15] or the reweighted ℓ_1 basis pursuit procedure described in [16]. However, since the dictionary $\mathbf{A}$ has very correlated atoms, consistency in selection cannot be achieved by any of the aforementioned techniques. It can be seen in [6] that a non-negative LASSO behaves as well as the latter techniques for activity estimation, thus justifying our choice.

Given a sparsity parameter r, and assuming that $\widehat{\boldsymbol{\beta}}(r) = [\widehat{\boldsymbol{\beta}}_0^T(r), \widehat{\boldsymbol{\beta}}_1^T(r), \ldots, \widehat{\boldsymbol{\beta}}_{N-1}^T(r)]^T$ has been computed (we refer to [17,18] for fast computation methods), the estimation of λ is performed by replacing a thresholded version of $\widehat{\boldsymbol{\beta}}(r)$ into a known estimate relying on the knowledge of the arrival times. It is shown in [8] that, in the case of a constant activity, the (scalar) λ can be effectively estimated as follows:

$$\hat{\lambda}_{hpp}(r, \eta) = -\frac{1}{\Delta t} \ln\!\left(1 - \frac{1}{N}\sum_{n=0}^{N-1} \mathbf{1}(\|\widehat{\boldsymbol{\beta}}_n(r)\|_1 \ge \eta)\right), \tag{5}$$

with $\eta > 0$ being a user-defined threshold, whereas for time-varying activities, the (functional) λ can be estimated in a non-parametric setting as

$$\hat{\lambda}_{nhpp}(t; r, \eta) = \frac{1}{h}\sum_{i=1}^{\widehat{M}} W\!\left(\frac{t - \widehat{T}_i}{h}\right), \tag{6}$$

where W is a standard kernel function integrating to 1 and $h > 0$ is a user-defined bandwidth parameter, and $\widehat{T}_i = \min\{t_n > \widehat{T}_{i-1} :\ \widehat{\boldsymbol{\beta}}_n(r) > \eta, \widehat{\boldsymbol{\beta}}_{n-1}(r) < \eta\}$, as in [8]. The necessity for a threshold η in both (5) and (6) can be understood from the poor consistency in selection of LASSO[19], which holds under theoretical conditions often unfulfilled in practice. Therefore, only the most representative coefficients of $\widehat{\boldsymbol{\beta}}(r)$ must be selected, this done by adding an additional thresholding step.

Obviously, (5) and (6) are conceptually different, since the first uses properties inherent to the homogeneity of the underlying Poisson process, and the second relies on a kernel function estimate. However, it can be seen that both rely on a thresholded version of a non-negative LASSO regressor, thus, both depend on the parameter (r, η), which must be carefully set up to ensure satisfactory results. In previous contributions, we provided efficient, rule-of-thumb criteria [8] as well as theoretical conditions [6,13] on both r and η which guarantee a satisfactory estimation of the activity, both in the homogeneous and non-homogeneous frameworks. However, the theoretical conditions are difficult to check in practice, and assume prior information on the radioactive source which may be unavailable [13]. Thus, fully automatic strategies for parameter settings remain to be investigated.

2.2. Shortcomings of Cross-Validation

One seemingly relevant strategy would be the use of Cross-Validation (CV) for the choice of η. This being said, a standard cross-validation on the threshold parameter (or on the joint parameter (r, η)) would not provide satisfactory results, since it would systematically yield a larger than expected choice of (r, η). We present in Fig. 1 a comparison between the approximated MSE with full knowledge of λ and the leave-10-out cross validation function, estimated on 1000 simulations in the homogeneous setting (that is, using Eq. (5)).