

Autoregressive Process Parameter Estimation from Compressed Sensing Measurements

Original

Autoregressive Process Parameter Estimation from Compressed Sensing Measurements / Testa, M., Magli, E.. - STAMPA. - (2015), pp. 488-492. (2015 49th Asilomar Conference on Signals, Systems and Computers 2015) [10.1109/ACSSC.2015.7421176].

Availability:

This version is available at: 11583/2638900 since: 2016-04-01T17:29:02Z

Publisher:

IEEE

Published

DOI:10.1109/ACSSC.2015.7421176

Terms of use:

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

Publisher copyright

IEEE postprint/Author's Accepted Manuscript

©2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

Autoregressive Process Parameter Estimation from Compressed Sensing Measurements

Matteo Testa and Enrico Magli

Department of Electronics and Telecommunications – Politecnico di Torino (Italy)

Abstract—In this paper we introduce a least squares estimator of the regression coefficients of an autoregressive process acquired by means of Compressed Sensing (CS). Unlike common CS problems in which we only know that the signal is sparse, using the proposed autoregressive model we can gain knowledge about the structure of the original signal without recovering it. This problem is addressed by introducing an ad-hoc sensing matrix able to preserve the structure of the regression. We numerically validate the performance of this matrix. Moreover, we present applications that naturally exploit this additional information we can directly obtain from the compressed data, and particularly power spectral density estimation from CS measurements.

Keywords—Compressed Sensing, Autoregressive process, compressibility, parameter estimation.

I. INTRODUCTION

Compressed Sensing (CS) [1] [2] is a new paradigm in which the signal acquisition and compression collapse into a single operation. The biggest advantage of CS is the ability to sense a signal at a lower rate than the classical Nyquist rate. CS recovery relies on the fact that the acquired signal must be sparse in some domain. This assumption plays then a central role in the recovery process since it is the foundation which allows to guarantee the exact recover of the original signal with overwhelming probability. However, this operation is generally expensive in terms of required computational power and, if only few parameters are needed, is preferred to infer them from the compressed measurements.

In many CS applications, once the compressed measurements have been obtained, no knowledge about the nature of the compressed signal is available, except for the assumption that it is sparse in some domain. However, natural signals are typically not exactly sparse, but rather approximately sparse or “compressible”. In many cases, such compressibility implies that the signal spectrum is decreasing [3]. This kind of information on the signal structure has indeed been used to improve the CS reconstruction [4].

In this paper we are interested in a more general model to understand the underlying structure of the compressed signal. The goal is hence not only to improve the recovery process, but also to gain knowledge on the signal structure by means of few inferred parameters. Such structure can then be used *e.g.* to perform signal processing operations directly in the compressed domain [5] [6], such as detection, parameter estimation, filtering, and so on.

In this paper we propose to model the structure of the sensed signal using an autoregressive (AR) process. The reasons are twofold: the first is related to its good approximation of real processes since this model is suited to efficiently represent signals with a natural spectrum. The other is that, since it is a parametric model, few coefficients are expected to suffice to characterize the signal nature. This latter reason has great importance for all the settings in which the communication cost must be kept low and hence it is worth to transmit only few parameters instead of the signal, even though it is compressed.

Although to the best of our knowledge there are no papers in literature dealing with the estimation of AR parameters from compressed sensing measurements, there are a few papers addressing the recovery of “compressed” AR(p) processes. The common basic assumption is that the driving process of the AR(p) model must be sparse. In [7] the authors estimate the regression coefficients before performing any compressed sensing, and then use them later as a way to construct a sparsifying basis in order to recover the signal. Another technique is proposed in [8] where, using a circulant sensing matrix and exploiting its convolution property, the authors first recover the AR process, and then estimate its parameters. The same authors, in a later paper [9], introduce a functional in which the regression coefficients are estimated along with the driving process, resulting in a modified LASSO for signal recovery.

Unlike previous works, this paper addresses the problem of the standalone estimation of the underlying signal structure (in terms of AR coefficients) *directly* from its CS measurements. In this paper, to reach the goal we introduce a new design for sensing matrices which, along with the least-squares (LS) estimator we propose, allows to estimate the coefficients of the uncompressed AR signal starting from its CS measurements. Moreover, we present two applications exploiting the estimated regression coefficients: the first one is an improved recovery which takes advantage of the estimated model parameters, the latter one is the spectral estimation of a signal from its compressed measurements. We also show that, given only the CS measurements with no additional information, the improved recovery we propose outperforms other existing techniques lowering the recovery MSE up to -7 dB.

II. BACKGROUND

A. Compressed Sensing

CS allows to exactly recover a sparse signal given a small number of its random projections. Let $v \in \mathbb{R}^N$ be a sparse signal (*i.e.*, it has a number of non-zero entries $s \ll N$), and $\Phi \in \mathbb{R}^{M \times N}$ with $M \ll N$ be the sensing matrix. Then, given $y = \Phi v \in \mathbb{R}^M$, v can be exactly recovered

This work has received funding from the European Research Council under the European Communitys Seventh Framework Programme (FP7/2007-2013) ERC Grant agreement no. 279848.

with overwhelming probability. However, most signals are not sparse in the domain of acquisition and hence it is necessary to introduce a sparsifying basis $\Psi \in \mathbb{R}^{N \times N}$. The compressed sensing acquisition model can be then written as $y = \Phi \Psi v$. A possible way to recover the signal is through the *LASSO*, defined as

$$\arg \min_u \lambda \|v\|_1 + \|y - \Phi \Psi v\|_2.$$

B. Autoregressive processes

An AR process of order p is a parametric model able to describe the time-varying nature of a process in which the output values linearly depend on their previous values. More formally

$$x_t = \sum_{i=1}^p x_{t-i} a_i + u_t, \quad (1)$$

where u is called driving process and a_i are the coefficients of the regression. In other words, it can be seen as a filtering operation over a process u with an all-pole filter with coefficients given by $a = [a_1 \dots a_i \dots a_p]^\top$.

III. COMPRESSED LEAST SQUARES ESTIMATION OF AR COEFFICIENTS

The main goal of this paper is to estimate the coefficients of an AR(p) process directly from the compressed measurements. In order to reach the goal, among the different classes of the estimators of the AR coefficients available in literature, we focus on the LS estimator. In fact, this choice, along with a new sensing matrix design we introduce (with excellent recovery performances as shown in Sec. V), allow us to explicitly infer the regression coefficients.

Let us start by introducing some notation and an LS estimator for the uncompressed domain. Given an AR(p) process $x \in \mathbb{R}^N$, we define x^+ as a subset of x composed by its samples with index from $(p+1)$ to N . Let us also define the matrix $X \in \mathbb{R}^{(N-p) \times p}$ constructed in the following way

$$X = \begin{bmatrix} x_p & x_{p-1} & \dots & x_1 \\ x_{p+1} & x_p & \dots & x_2 \\ \vdots & \dots & \dots & \vdots \\ x_{N-1} & x_{N-2} & \dots & x_{N-p} \end{bmatrix}. \quad (2)$$

Since by (1) we obtain

$$x^+ = Xa, \quad (3)$$

we can hence write an LS estimator for the parameters of a process as the minimizer of the following functional

$$\arg \min_{\hat{a}} \|x^+ - X\hat{a}\|_2 \quad (4)$$

or, more concisely, as $\hat{a} = X^\dagger x^+$ where “ \dagger ” denotes the pseudo-inverse.

In order to have an analogous LS estimator for the compressed domain we need to introduce a sensing matrix able to preserve the structure of the regression. As we can see from (2), the LS estimator for a process of order p , needs $p+1$ shifted versions of the input signal. Hence, the idea is to build

a sensing matrix from which, given the output measurements, is possible to extract the compressed $p+1$ shifted versions of x . This means that the sensing matrix should be made of $p+1$ sub-blocks Φ' where each of them senses a shifted version of the given signal x .

Then, if we use (3), multiplying both sides by the sensing block Φ' we get

$$y^+ = Y a, \quad (5)$$

where $y^+ = \Phi' x^+$ and $Y = \Phi' X$. Hence, if the sensing matrix is made up of shifted sensing blocks it is possible to extract the quantities y^+ and Y from the measurement vector y .

More formally, let us assume that the main block $\Phi' \in \mathbb{R}^{\mu \times (N-p)}$ with $\mu = M/(p+1)$, has entries distributed according to $\phi'_{ij} \sim \mathcal{N}(0, \frac{1}{M})$. Then, the proposed sensing matrix $\Phi \in \mathbb{R}^{M \times N}$ is made of $p+1$ circulant blocks of Φ' as depicted in Fig. 1. In Sec. V we will show that this matrix has very good properties for CS reconstruction.

$$\begin{bmatrix} \phi'_{1,1} & \phi'_{1,2} & \phi'_{1,3} & \phi'_{1,4} & \phi'_{1,5} & \phi'_{1,6} & 0 & 0 \\ \phi'_{2,1} & \phi'_{2,2} & \phi'_{2,3} & \phi'_{2,4} & \phi'_{2,5} & \phi'_{2,6} & 0 & 0 \\ 0 & \phi'_{1,1} & \phi'_{1,2} & \phi'_{1,3} & \phi'_{1,4} & \phi'_{1,5} & \phi'_{1,6} & 0 \\ 0 & \phi'_{2,1} & \phi'_{2,2} & \phi'_{2,3} & \phi'_{2,4} & \phi'_{2,5} & \phi'_{2,6} & 0 \\ 0 & 0 & \phi'_{1,1} & \phi'_{1,2} & \phi'_{1,3} & \phi'_{1,4} & \phi'_{1,5} & \phi'_{1,6} \\ 0 & 0 & \phi'_{2,1} & \phi'_{2,2} & \phi'_{2,3} & \phi'_{2,4} & \phi'_{2,5} & \phi'_{2,6} \end{bmatrix}$$

Fig. 1. Circulant blocks structure of Φ .

In order to obtain the measurements of shifted versions of the original signal (which are needed to obtain y^+ and Y), we exploit the structure of the sensing matrix in Fig. 1. As can be seen, each sub-block acquires a shifted version of the input signal through the sub-sensing matrix Φ' . Hence we can write $y_{1+\mu(k-1) \rightarrow k\mu} = \Phi' x_{k \rightarrow (N-p+k-1)}$, which means that the vector y is made of $p+1$ blocks of length μ which are the measurements corresponding to different shifts of x .

In particular, using (5) we define the compressed LS estimator for AR(p) coefficients, as:

$$\arg \min_{\hat{a}} \|y^+ - Y\hat{a}\|, \quad (6)$$

where the chosen M must be an integer multiple of $p+1$. The performance of this estimator is validated in Sec. V.

It is worth noting that the proposed estimator, working directly in the reduced space of the measurements domain, is computationally less demanding with respect to the corresponding LS estimator (4) in the uncompressed domain. The complexity strictly depends on the value of p and the length of signal N according to $O(p^2(N-p))$, where the most influential term is N because the order of the process is typically small. Therefore, the required computational power for the proposed estimator drastically reduces to $O(p^2\mu)$ with $\mu \ll (N-p)$.

IV. APPLICATIONS

A. Sparsely-driven AR process recovery

As previously discussed, gaining knowledge of the underlying signal structure can improve the CS recovery process: in fact the proposed estimator allows to recover a compressively

sensed AR(p) process under some sparsity assumptions. In this section we consider a process obtained by filtering a sparse driving process. This model is of interest in speech processing, since the speech signal can be modeled as an AR process where the residual (driving process of the model) is sparse as in Multi-Pulse excitation coding [10].

To tackle this problem we need to introduce a basis able to represent our process sparsely. Since, by assumption, the driving noise of AR(p) process is sparse, our sparsifying basis will be that filtering matrix which reverses the AR(p) filtering effect.

Given the AR(p) process x and the sparse driving process u , we can write $u = Ax$, where A is the matrix performing the inverse filtering operation on x , i.e, the lower-triangular Toeplitz matrix created from the length- N vector $[1a^T 0 \dots 0]$. Then, the basis we are looking for is $H \triangleq A^{-1}$.

When considering the compressed measurements we can equivalently write $y = \Phi x = \Phi H u$, which leads to the following LASSO problem for the AR(p) recovery:

$$\arg \min_u \|y - \Phi \tilde{H} u\|_2 + \lambda \|u\|_1, \quad (7)$$

where \tilde{H} is the sparsifying basis constructed from the estimate of a in the CS domain. It is worth noting that using the proposed estimator along with the LASSO in (7) to improve the signal recovery only requires the CS measurements and hence no side informations or training data must be known.

B. Compressive spectral estimation

Signal recovery in CS is in general a computationally expensive task. With this in mind, there are some cases in which, exploiting the signal structure, the recovery could be avoided. Examples include filtering unwanted spectral components directly in the compressed domain or discarding the signal because the spectrum does not contains certain required components.

It is known [11] that the Power Spectral Density (PSD) estimate of a signal R_y is related to the coefficients of the regression according to

$$R_y = \left| \frac{1}{A(\omega)} \right|^2, \quad (8)$$

where $A(\omega) = 1 + a_1 e^{-i\omega} + \dots + a_p e^{-ip\omega}$. Hence, the proposed estimator, along with the specifically designed sensing matrix allows to use the AR parameters estimated from the CS measurements to build an estimate of the PSD of the signal. As an example, let us assume $y = \Phi x$ is a signal composed by three distinct spectral lines; if additional spectral components are present the signal is assumed to be corrupted and its recovery is not needed. This kind of problems is effectively solved by the proposed framework; indeed, having an estimate of the AR coefficients \hat{a} allows to build a PSD estimate of the signal and perform decisions prior to execute the computationally demanding recovery step.

V. NUMERICAL EXPERIMENTS

This section is devoted to the numerical validation of the properties of the proposed estimator and of the performances of such estimates when used in the aforementioned applications.

A. Sensing matrix validation

In this subsection, we validate the recovery ability of the proposed sensing matrix compared to some of the most used ones in literature for which theoretical results on the recovery performance exists [12] [13]. More in detail, we show that these matrices can indeed be used as standard sensing matrices within CS framework for generic sparse signals. For this experiment we fix a sparsity level $s = 100$ and randomly pick the support of the non-zero components. Hence, we compared the performances of these matrices by running 1000 different Monte Carlo runs over different M values by compressing sparse signals and then recovering them using LASSO. The error metric is the relative recovery error defined as $\|x - \hat{x}\|_2 / \|x\|_2$ where \hat{x} is the recovered signal. The results (Fig. 2) show that recovery error of the proposed matrix is lower bounded by the Gaussian sensing matrix and upper bounded by the Bernoulli one. The recovery performance of the proposed matrix is hence comparable to that of the circulant one showing that the proposed sensing matrix is comparable to other popular ones, and has a negligible performance loss with respect to a Gaussian matrix.

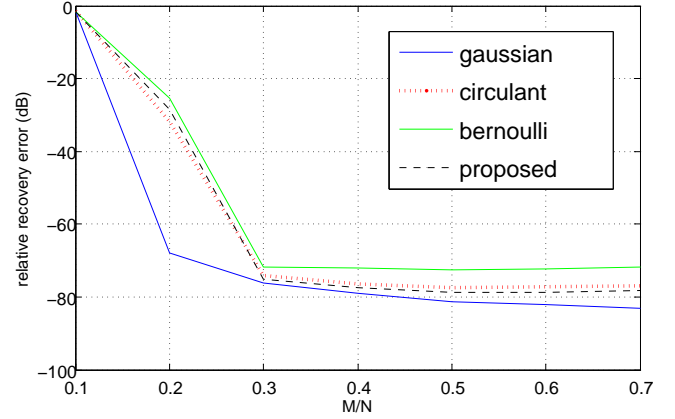


Fig. 2. Comparison of recovery ability of different sensing matrices using a signal of length $N = 1000$, sparsity $s = 100$ and $p = 9$.

B. Sparsely-driven AR recovery

Here we simulate the recovery of a sparsely driven AR(p) process. First, we generate a sparse synthetic signal and use LASSO as in Sec. IV-A to recover the signal (Fig. 5). The error metric we use for this experiment is the same as for the sensing matrices comparison. The results are averaged over 100 different trials. In particular, we compare the proposed method with the recovery using the Discrete Fourier Transform (DFT) sparsifying basis, and the one proposed in [7]. It is important to highlight that, while both the proposed and the DFT-based recovery do not require any additional information apart for the CS measurements, the recovery in [7] requires the a-priori knowledge of the AR coefficients of the process. The DFT-based recovery shows a very large error, showing that the DFT basis is not good for sparsifying this class of signals. Conversely, the other two methods show lower error, which decreases as M increases. Moreover, the errors converge as $M \rightarrow N$ because the two methods use the same sparsifying basis, but constructed with different AR coefficients estimates (from CS measurements and from original signal). Finally, in Fig. 3 we show the recovery of a *natural* signal: a vocalized

tract of a speech signal. For this experiment, in order to run a fair comparison, we compare the techniques which require only the CS measurements: the DFT-based recovery and the proposed technique. As we can see, with a small number of measurements, the signal recovered using the proposed method approximates the original one very well. In contrast, the recovery assuming DFT-based sparsity is not able to approximate the original signal accurately. In fact, the MSE of the proposed recovery (shown in Fig. 3) is -30.46 dB conversely to the DFT-based which is -23.4 dB.

C. Compressive spectral estimation

To evaluate the performance of the compressive spectral estimation, we start by focusing on the performance of the estimated AR coefficients. Indeed we study the estimation error of the AR(p) coefficients comparing it with the LS estimator in the uncompressed domain, namely we analyze the performance loss deriving from working in the compressed domain. As shown in Figure 4(a), the error, defined as $\|a - \hat{a}\|_2 / \|a\|_2$, is small and it decreases as M/N approaches 1. However, the two estimators do not reach the same value for $M = N$. This is due to the fact that since $Y = \Phi'X$ and Φ' has $\mu = M/(p+1)$ rows, the maximum value assumed by M is $M = N$ then $\mu = N/(p+1)$ will always be smaller than N . This means that in the case of $M = N$, the block Φ' is still a rectangular matrix and the problem is under-determined.

Then, since there is a direct relationship between the AR coefficients and the PSD of the signal as in (8), the goodness of the estimated parameters can also be seen from the PSD perspective. In Fig. 4(b) we show the comparison between the spectrum of a signal made of three sinusoidal components estimated using the FFT for the original signal and the compressive spectral estimation for CS measurements. As can be seen, the spectrum generated with the estimated AR coefficients accurately approximates the spectrum spikes of the original signal.

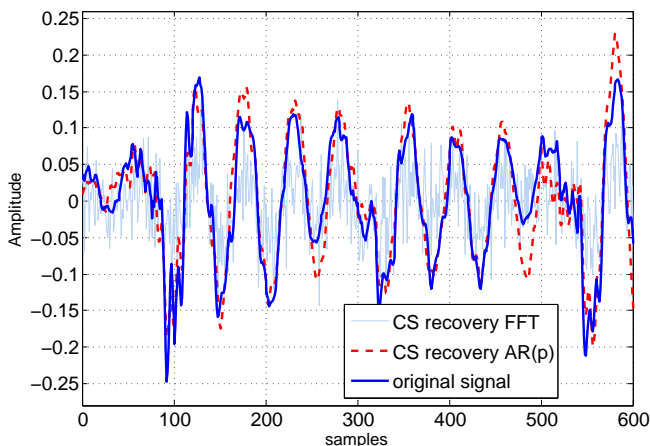
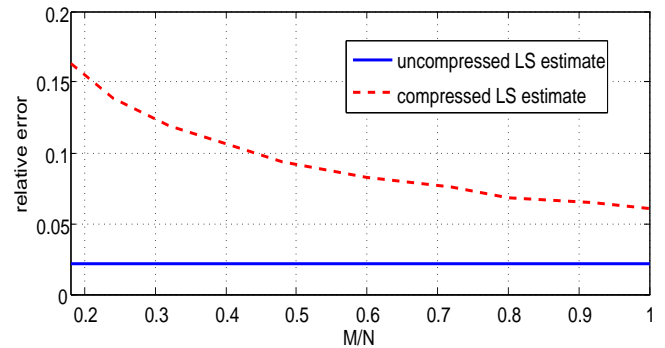
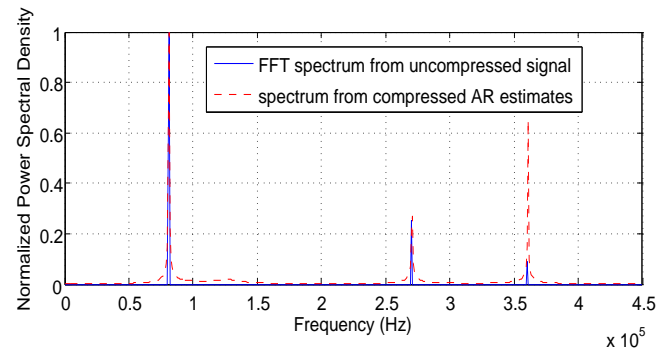


Fig. 3. Recovery comparison of a vocalized tract of speech signal of length $N = 600$ with $M = 200$ and $p = 10$. We compare the original signal with the recovered versions made by using the DFT as a sparsifying basis, and the LASSO in Sec. IV-A with the estimated regression coefficients.



(a) Relative error comparison between uncompressed and compressed LS AR(p) coefficients estimators. $N = 1000$ and $p = 4$, since in cases of interest p is small, usually $p \leq 10$.



(b) FFT and compressed AR spectrum comparison. For the experiment $N = 1000$, $M = 100$ and $p = 9$.

Fig. 4. Performance evaluation of the proposed estimator. In (a) we show the performance loss deriving from using compressed measurements instead of uncompressed data, in (b) we show an example of compressive spectral estimation via estimated AR coefficients.

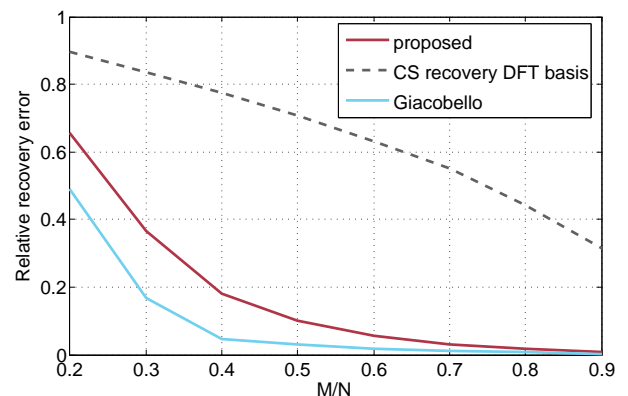


Fig. 5. Relative recovery error comparison of a synthetic sparse signal with $N = 1000$, and sparsity $s = 100$. We compare the recovery proposed by Giacobello *et al.* in [7], a CS recovery assuming the sparsity to be in the frequency domain and the proposed method with regression coefficients estimated from the measurements.

VI. CONCLUSIONS

In this paper we proposed to model the underlying structure of a compressively sensed signal with as an AR(p) process. We hence derived a new sensing matrix design which along with the proposed LS estimator allows to estimate the process parameters directly from CS measurements.

Starting from the intuition of modeling the signal struc-

ture with an AR process, we showed that according to the proposed setup, it is possible to extract information from the compressed signal and in turn improve the recovery. In fact, our experiments on synthetic and real signals showed that the proposed technique outperforms other recovery methods which only rely on the CS measurements without the ability of gaining information on the signal structure.

Moreover, we also showed that the AR coefficients are not only valuable in improving the recovery but that also very effective *per se*. We indeed showed that, when considering frequency sparse signals, the PSD of the uncompressed signal is effectively estimated from its random projections.

Future work will move in the direction of providing theoretical analysis of the proposed estimator and sensing matrix.

REFERENCES

- [1] Emmanuel J Candès, Justin Romberg, and Terence Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on*, vol. 52, no. 2, pp. 489–509, 2006.
- [2] David L Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [3] Gregory W Wornell, "Wavelet-based representations for the 1/f family of fractal processes," *Proceedings of the IEEE*, vol. 81, no. 10, pp. 1428–1450, 1993.
- [4] Richard G Baraniuk, Volkan Cevher, Marco F Duarte, and Chinmay Hegde, "Model-based compressive sensing," *Information Theory, IEEE Transactions on*, vol. 56, no. 4, pp. 1982–2001, 2010.
- [5] Mark A Davenport, Michael B Wakin, and Richard G Baraniuk, "Detection and estimation with compressive measurements," *Dept. of ECE, Rice University, Tech. Rep*, 2006.
- [6] Diego Valsesia and Enrico Magli, "Compressive signal processing with circulant sensing matrices," *arXiv preprint arXiv:1403.2835*, 2014.
- [7] Daniele Giacobello, "Retrieving Sparse Patterns Using a Compressed Sensing Framework: Applications to Speech Coding Based on Sparse Linear Prediction," 2010.
- [8] Venkatesh Saligrama and Manqi Zhao, "Compressed sensing of autoregressive processes," *SPIE Optical ...*, vol. 7446, pp. 1–8, 2009.
- [9] Manqi Zhao and Venkatesh Saligrama, "Noisy filtered sparse processes: Reconstruction and compression," *49th IEEE Conference on Decision and Control (CDC)*, pp. 2930–2935, Dec. 2010.
- [10] Bishnu S Atal and J Remde, "A new model of lpc excitation for producing natural-sounding speech at low bit rates," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'82*. IEEE, 1982, vol. 7, pp. 614–617.
- [11] Steven M Kay, *Modern spectral estimation*, Pearson Education India, 1988.
- [12] Holger Rauhut, "Circulant and toeplitz matrices in compressed sensing," *arXiv preprint arXiv:0902.4394*, 2009.
- [13] Emmanuel J Candès and Terence Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *Information Theory, IEEE Transactions on*, vol. 52, no. 12, pp. 5406–5425, 2006.