Signal Processing 98 (2014) 275–283

Contents lists available at ScienceDirect

Signal Processing

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

Compressed sensing with linear correlation between signal and measurement noise

Thomas Arildsen [*,♣], Torben Larsen

Aalborg University, Department of Electronic Systems, Signal and Information Processing, Niels Jernes Vej 12, DK-9220 Aalborg, Denmark

ARTICLE INFO

Article history:
Received 13 January 2013
Received in revised form
16 October 2013
Accepted 22 October 2013
Available online 21 November 2013

Keywords:
Compressed sensing
Convex optimization
Correlated noise
Quantization

ABSTRACT

Existing convex relaxation-based approaches to reconstruction in compressed sensing assume that noise in the measurements is independent of the signal of interest. We consider the case of noise being linearly correlated with the signal and introduce a simple technique for improving compressed sensing reconstruction from such measurements. The technique is based on a linear model of the correlation of additive noise with the signal. The modification of the reconstruction algorithm based on this model is very simple and has negligible additional computational cost compared to standard reconstruction algorithms, but is not known in existing literature. The proposed technique reduces reconstruction error considerably in the case of linearly correlated measurements and noise. Numerical experiments confirm the efficacy of the technique. The technique is demonstrated with application to low-rate quantization of compressed measurements, which is known to introduce correlated noise, and improvements in reconstruction error compared to ordinary Basis Pursuit De-Noising of up to approximately 7 dB are observed for 1 bit/sample quantization. Furthermore, the proposed method is compared to Binary Iterative Hard Thresholding which it is demonstrated to outperform in terms of reconstruction error for sparse signals with a number of non-zero coefficients greater than approximately 1/10th of the number of compressed measurements.

1. Introduction

In the recently emerged field of compressed sensing, one considers linear measurements $\mathbf{y}$ of a sparse vector $\mathbf{x}$, possibly affected by noise as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{1}$$

where the measurements $\mathbf{y} \in \mathbb{R}^{M \times 1}$, the sparse vector $\mathbf{x} \in \mathbb{R}^{N \times 1}$, the additive noise $\mathbf{n} \in \mathbb{R}^{M \times 1}$, the system matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$, and $M \ll N$ [1–3]. $\mathbf{A}$ is generally the product of a measurement matrix and a dictionary matrix: $\mathbf{A} = \mathbf{\Phi\Psi}$, where $\mathbf{\Phi} \in \mathbb{C}^{M \times N}$, $\mathbf{\Psi} \in \mathbb{C}^{N \times N}$. For simplicity, we assume that $\mathbf{\Psi}$ is an orthonormal basis although more general dictionaries are indeed possible [4].

The essence of compressed sensing, as Donoho et al. have shown in [1,2], is that the under-determined equation system (1) can be solved provided that:

1. The vector $\mathbf{x}$ is sparse; i.e., only few (K) elements in $\mathbf{x}$ are non-zero:

   $$K = |\{x_i | x_i \neq 0, \; i = 1, \ldots, N\}| \tag{2}$$

   $\mathbf{x}$ can also be approximated sparsely if it is compressible [3, Section 3.3], meaning that its coefficients sorted by magnitude decay rapidly to zero.
2. The system matrix $\mathbf{A}$ obeys the Restricted Isometry Property (RIP) with isometry constant $\delta_K > 0$, defined as follows:

   $$(1-\delta_K)\|\mathbf{x}\|_{\ell_2}^2 \le \|\mathbf{A}\mathbf{x}\|_{\ell_2}^2 \le (1+\delta_K)\|\mathbf{x}\|_{\ell_2}^2, \tag{3}$$

* Corresponding author. Tel.: +45 99409844.
E-mail addresses: tha@es.aau.dk (T. Arildsen), tl@es.aau.dk (T. Larsen).
♣ EURASIP member.

0165-1684/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sigpro.2013.10.021

for any at most K-sparse vector $\mathbf{x}$ such that [5]

$$\delta_K + \delta_{2K} + \delta_{3K} < 1. \tag{4}$$

This holds with high probability when $\mathbf{\Phi}$ is generated with zero-mean independent identically distributed (i.i.d.) Gaussian entries with variance $1/M$. Note that (3) and (4) are sufficient but not necessary conditions, and rather conservative conditions indeed, as shown in [6]. Conditions (3) and (4) lead to the following sufficient amount of measurements M for Gaussian measurement matrices $\mathbf{\Phi}$ [7]:

$$M \geq CK \log\left(\frac{N}{M}\right), \tag{5}$$

where C is a fairly small constant which can be calculated as a function of M/N [5].

Given the measurements $\mathbf{y}$, the unknown sparse vector $\mathbf{x}$ can be reconstructed by solving the following convex optimization problem [3, Section 4]:

$$\hat{\mathbf{x}} = \underset{\mathbf{u}:\|\mathbf{y}-\mathbf{A}\mathbf{u}\|_2 \leq \varepsilon}{\operatorname{argmin}} \|\mathbf{u}\|_1, \tag{6}$$

where the fidelity constraint $\|\mathbf{y}-\mathbf{A}\mathbf{u}\|_2 \leq \varepsilon$ ensures consistency with the observed measurements to within some margin of error, ε, which is chosen sufficiently large to accommodate the error $\mathbf{n}$ and/or approximation error in the case of compressible signals. The form of the optimization problem in (6) is known as Least Absolute Shrinkage and Selection Operator (LASSO) [8] or Basis Pursuit De-Noising (BPDN) [9] and also comes in other variants such as the Dantzig selector [10]. In addition to the convex optimization approach to reconstruction in compressed sensing, there exist several iterative/greedy algorithms such as Iterative Hard Thresholding (IHT) [11], or Subspace Pursuit (SP) [12] and Compressive Sampling Matching Pursuit (COSAMP) [13] as well as the more generalized incarnation of the two latter, Two-Stage Thresholding (TST) [14]. We generally refer to such convex or greedy approaches as reconstruction algorithms. The reconstruction algorithms generally assume the noise to be white and independent of the measurements before noise $\bar{\mathbf{y}} = \mathbf{A}\mathbf{x}$. In particular, to the best of the authors' knowledge, the case of measurement noise being linearly correlated with the measurements has not been treated in the existing literature. Such correlation arises in for example the case of low-resolution quantization. As we demonstrate in Section 2, this case poses a problem for the accuracy of the found solution $\hat{\mathbf{x}}$. More special cases of correlated noise arising from Poisson measurements or quantization of measurements have, however, been treated in for example [15–17].

In this paper, we propose a simple yet efficient approach to alleviating the problem of linear correlation between the measurements before noise $\bar{\mathbf{y}}$ and the noise $\mathbf{n}$. Our proposal boils down to a simple scaling of the solution $\hat{\mathbf{x}}$. Through numerical experiments we demonstrate how linearly correlated measurements and noise adversely affect the reconstruction error and demonstrate how our proposal improves the estimates considerably.

As an application example, we demonstrate the proposed approach in the case of low-rate scalar quantization of the measurements $\bar{\mathbf{y}}$ which can be observed to introduce the mentioned linearly correlated measurement noise. We demonstrate how a well-known linear model used for modeling such correlation in scalar quantization is equivalent to the model of correlated measurement noise considered in this work.

The paper is structured as follows: Section 2 introduces the considered model of linear correlation between compressed measurements and noise and proposes a solution to enhance reconstruction under these conditions, Section 3 describes simulations conducted to evaluate the performance of the proposed approach compared to a traditional approach, Section 4 presents the results of these numerical simulations, Section 5 provides discussions of some of the presented results, and Section 6 concludes the paper.

2. Methodology

2.1. Correlated measurements and noise

We consider additive measurement noise $\mathbf{n}$ which is correlated with the measurements before noise $\bar{\mathbf{y}}$. We model the correlation by the linear model:

$$\mathbf{y} = \alpha\mathbf{A}\mathbf{x} + \mathbf{w}, \tag{7}$$

where $\mathbf{w}$ is assumed an additive white noise uncorrelated with $\mathbf{x}$ and $0 < \alpha \leq 1$ where $\alpha = 1$ covers the ordinary case of uncorrelated measurement noise. $\mathbf{A}$ is the product of a measurement matrix $\mathbf{\Phi}$ with i.i.d. Gaussian entries $\sim \mathcal{N}(0, 1/M)$ and an orthonormal dictionary matrix $\mathbf{\Psi}$. The model (7) results in the following additive noise term:

$$\mathbf{n} = \mathbf{y} - \bar{\mathbf{y}} = \alpha\mathbf{A}\mathbf{x} + \mathbf{w} - \mathbf{A}\mathbf{x} = (\alpha - 1)\mathbf{A}\mathbf{x} + \mathbf{w} \tag{8}$$

We define $\bar{\mathbf{y}} = \mathbf{A}\mathbf{x}$ to signify the measurements before introduction of additive noise. It is readily seen from (8) that $\mathbf{n}$ is correlated with $\mathbf{x}$. The noise variance is

$$\sigma_n^2 = \frac{1}{M}\mathbb{E}\left[\mathbf{n}^T\mathbf{n}\right] = \frac{1}{M}((\alpha-1)^2\mathbb{E}[\bar{\mathbf{y}}^T\bar{\mathbf{y}}] + \mathbb{E}[\mathbf{w}^T\mathbf{w}]), \tag{9}$$

which can be calculated by assuming that $\sigma_{\bar{y}}^2 = (1/M)\mathbb{E}[\bar{\mathbf{y}}^T\bar{\mathbf{y}}]$ and $\sigma_w^2 = (1/M)\mathbb{E}[\mathbf{w}^T\mathbf{w}]$ are known or can be estimated. For example, we show an example for σ_w^2 in the case of quantization in Section 2.5, (21).

The specific problem caused by correlated measurements and noise as modeled by (7) is that the noise itself is partly sparse in the same dictionary as the signal of interest, $\mathbf{x}$. Intuitively, this causes a solution $\hat{\mathbf{x}}$ as given by, e.g., (6) to adapt to part of the noise *as well as* the signal of interest, unless steps are taken to mitigate this effect.

2.2. Proposed approach

Using the model in (7), we propose the following reconstruction of the sparse vector $\mathbf{x}$ instead of the standard approach in (6). Eq. (7) motivates replacing the system matrix $\mathbf{A}$ by its scaled version $\alpha\mathbf{A}$. We exemplify this approach by applying it in the BPDN reconstruction formulation as below. Replacing $\mathbf{A}$ by $\alpha\mathbf{A}$ in the standard