Robust Data Hiding Scheme for Compressively Sensed Signals

Mehmet YAMAÇ
Laboratory of Signal Processing
Tampere University of Technology
Tampere, Finland
Email: mehmet.yamac@tut.fi

Bülent Sankur
Electrical and Electronics
Engineering
Boğaziçi University
Bebek 34342, Istanbul, Turkey
Email: bulent.sankur@boun.edu.tr

Moncef Gabbouj
Laboratory of Signal Processing
Tampere University of Technology
Tampere, Finland
Email:moncef.gabbouj@tut.fi

***Abstract*—We consider the problem of linear data hiding or watermark embedding directly onto compressively sensed measurements (CSMs). In our encoding and decoding scheme, we seek exact recovery of concealed data and a small reconstruction error for a sparse signal under the additive noise model. We propose an efficient Alternating Direction of Methods of Multiplier (ADMM) based decoding algorithm and we show through experimental results that proposed decoding scheme is more robust against additive noise compared to competing algorithms in the literature.**

***Index Terms*—Compressive Sensing, Data Hiding, Watermarking, Image Encryption, Privacy Preserving**

I. Introduction

Compressive sensing [1] (CS) theory has emerged as a remedy for applications where data acquisition is costly, e.g., expensive sensor are required or when the resulting sampled data volume is impractically large. CS theory states a signal can be represented with far fewer samples compared to the Nyquist rate if in a proper domain it is sparse or at least compressible. There are already success stories of CS such in MRI imaging [2] where as a consequence the signal acquisition time has been significantly reduced, or in a health monitoring system where compressive sampling of streaming ECG signals increases the battery life span [3]. In addition, single-pixel cameras have been developed [4] using compressive image sensing for mobile phones [5].

Data hiding and watermarking technologies have witnessed tremendous developments in the last two decades, becoming thus a mainstream technology. A good illustrative case is a health monitoring system where patient biomedical data is transmitted to a health center. In this case, data hiding and watermarking enable [6] embedding of meta-data such as Electronic Health Records or patient's identity for identification and authentication purposes. The overwhelming majority of these attempts, however, use embedding medium signals sampled according to the Nyquist-Shannon theorem and can not be applied to CS acquisition systems such as [2], [3], [4], [5].

Although CS is conceived as a data acquisition method, a CS framework is also capable of inherently providing confidentiality with a reasonable level of security. Furthermore, this capability comes at practically no additional cost and data encryption [7] can be added right into the sampling process. Compressive sensing enables encryption via random or pseudo-random sampling matrices.

There have been few attempts in the literature [8] to embed metadata directly onto a compressively sensed signal. The benefits of data hiding in compressively sampled signals are as follows: First, the compressive samples can be used as a carrier for subliminal information, and such a scheme can perform data hiding at a low cost by implementing linear encoding and spreading the hidden message directly during sensing. Second, encryption is enabled for the sensed samples and the additional embedding makes it harder for malicious user to hack the signals. In this case, the hacker must obtain both the encryption matrix (i.e., sensing matrix), and the data hiding or encoding matrix.

In this work, we propose a new decoding strategy for the data embedding in which the meta data is spread directly onto CSMs (compressively sensed measurements). This new method builds upon our previous method for data hiding in CSMs [8]. The novelty of the work is that it addresses the extraction of the watermark/hidden data and recovery of the carrier message as a sparse signal reconstruction problem. To this effect, we use proximal calculus and ADMM: Alternating Direction Method of Multipliers based decoding method. We show that this novel signal recovery and hidden data extraction method is more robust to additive noise, Gaussian or non-Gaussian, compared to the scheme in [8].

The rest of the paper is organized as follows. Section II provides the notation and the mathematical preliminaries. In Section III, we briefly review the CS framework. Then, in Section IV, we explain the proposed scheme and give the details of the ADMM-based decoding algorithm. Finally, the performance of the proposed algorithm is analyzed and conclusions are drawn.

ISBN 978-90-827970-1-5 © EURASIP 2018

II. Preliminaries

We define the ℓ_p norm of any vector $x \in \mathbb{R}^N$ as $\|x\|_{\ell_p^N} = \left(\sum_{i=1}^{N} |x_i|^p\right)^{1/p}$ for $p \geq 1$. The ℓ_0-norm of the vector $x \in \mathbb{R}^N$ is given as $\|x\|_{\ell_0^N} = \lim_{p \to 0} \sum_{i=1}^{N} |x_i|^p = \#\{j : x_j \neq 0\}$. The indicator function, $i_c : \mathbb{R}^N \to \{0, \infty\}$, of a convex set C is also defined as

$$i_C(x) = \begin{cases} 0 & \text{if } x \in C \\ \infty & \text{if } x \notin C. \end{cases}$$

A signal S is said to be strictly k-sparse if the number of non-zero coefficients is less than a constant k when we represent this signal in a proper basis (or dictionary), Φ, i.e.,

$$\|x\|_{\ell_0^N} \leq k, \tag{1}$$

where $S = \Phi x \in \mathbb{R}^N$. Indeed, most signals we encounter in real-world applications exhibit a power law decay in some apropriate Φ. Consider the coefficients of x sorted in descending order in magnitude, i.e., $\left|x^1\right| \geq \left|x^2\right| \geq ... \geq \left|x^{N-1}\right| \geq \left|x^N\right|$, where x^i is the i.th largest coefficient of x (in magnitude). Let $\wedge^{(k)}$ be the set of indices corresponding to the k largest coefficients. Then, the signal of interest S is said to be approximately k-sparse if

$$\left\|S - \Phi x_{\wedge^{(k)}}\right\|_{\ell_2}^{N} \leq \kappa, \tag{2}$$

where κ is a small constant.

The Restricted Isometry Constant (RIC) of order k of a $m \times N$ matrix A is defined as the smallest constant, $\delta_k(A) \in (0, 1)$ satisfying

$$(1 - \delta_k(A)) \|x\|_{\ell_2^N} \leq \|Ax\|_{\ell_2^m} \leq (1 + \delta_k(A)) \|x\|_{\ell_2^N} \tag{3}$$

for all k-sparse $x \in \mathbb{R}^N$.

Since we use in this work an ADMM-based decoding scheme, a brief reminder on proximal method is in order. ADMM [9] is a special type of a proximal algorithm [10]. Proximal operator or proximal mapping [11] of a function f at a point z with a parameter $\gamma > 0$ is defined as

$$\text{prox}_{\gamma f}(z) = \arg\min_u \{f(u) + \frac{1}{2\gamma} \|u - z\|_{\ell_2^N}^2\}. \tag{4}$$

A proximal operator can be considered as a gradient descent step for the smooth approximation of f. This method proves to be very useful in optimization problems involving non-differentiable functions. It can be interpreted as a generalization of the projection operator [12]. For instance, when f is the indicator function, i_C, the proximal operator will simply be the projection operation onto C, i.e.,

$$\Pi_c(z) = \arg\min_{x \in C} \|x - z\|_{\ell_2^N}.$$

III. Compressive Sensing Review

In compressive sensing we have m measurements of an N-dimensional signal $S \in \mathbb{R}^{\mathbb{N}}$, i.e.,

$$y = \Psi S, \tag{5}$$

where Ψ is the $m \times N$ (typically, $m << N$) linear measurement matrix. Assuming that this signal is compressible (approximately k-sparse) in a proper sparsifying basis Φ, we can re-arrange the equation as

$$y = \Psi S = \Psi\Phi x_{\wedge^{(k)}} + \Psi\Phi x_{\Lambda^c} = Ax_{\wedge^{(k)}} + n^0, \tag{6}$$

where $\Psi\Phi$ and the complement of set $\wedge^{(k)}$, defined as $\Lambda^c = \{1, 2, 3, ..., N\} \setminus \wedge^{(k)}$, represents the indices of the non-compressible small magnitude part of x, and n^0 is the corresponding additive distortion due to the discarded measurements (it vanishes in strictly sparse case). In addition CSMs can be corrupted by channel errors during transmission, also modeled as additive noise, or by quantization errors. Then, the reconstruction algorithm must handle

$$y = Ax + n, \tag{7}$$

where n is a general additive noise and x is a sparse signal (hereafter, we use the notation of x instead of $x_{\wedge^{(k)}}$ for convenience). Even in the noise-free case when n vanishes, Equation (7) is an underdetermined system of linear equations and has infinitely many solutions. In this case, one may find the sparsest solution from infinitely many by solving

$$\hat{x} = \arg\min_x \ \|x\|_{\ell_0^N} + i_{\{Ax=y\}}(x). \tag{8}$$

The problem formulated in (8) is not convex, and hence one can relax and convexify it so that, e.g., the well-known Basis Pursuit problem can be applied:

$$\hat{x} = \arg\min_x \ \|x\|_{\ell_1^N} + i_{\{Ax=y\}}(x). \tag{9}$$

It is proven that the solution of (9) is unique given that A has certain properties (e.g., null space property) [13].

In the noisy case, (9) can be expressed in terms of the following Basis Pursuit Denoising problem [14],

$$\hat{x} = \arg\min_x \|x\|_{\ell_0^N} \quad \text{s.t} \quad \|y - Ax\|_{\ell_2^m} \leq \epsilon \tag{10}$$

The stability of the solution of (10) is well studied in the literature. For instance, given that $\|n\|_{\ell_2} \leq \epsilon$, if matrix A possesses RIC with $\delta_{2k} < \sqrt{2} - 1$, equation (10) approximates x with

$$\|x - \hat{x}\|_{\ell_2^N} \leq C_0 \epsilon, \tag{11}$$

where C_0 depends on $\delta_{2k}(A)$ [15]. As an example of the measurement matrix, A with i.i.d. elements $A_{i,j}$ drawn according to $\mathcal{N}\left(0, \frac{1}{m}\right)$, $m > k(\log(N/k))$ guarantees with high probability exact signal reconstruction when the noise n vanishes [15].

IV. Data Hiding

Let $x \in \mathbb{R}^N$ be a k-sparse signal and $w \in \{a, -a\}^M$ be an M-length binary hidden message (e.g., watermark) that we wish to linearly embed into CSMs of S. Recall that compressive sensing algorithms relocate the computational burden from the sampling side (e.g., transmitting end) to the reconstruction side (receiving end) by performing linear sampling and non-linear reconstruction. On the other hand, the M-long binary hidden message can be linearly spread directly onto CSMs via the encoding matrix $B \in \mathbb{R}^{m \times M}, M < m$, resulting in the marked signal $y = Ax + Bw$, where $A \in \mathbb{R}^{m \times N}$ is the measurement matrix.

Finally, the marked signal carrying the hidden data can be modified by an additive noise or as a consequence of some attack, yielding:

$$y = Ax + Bw + n. \tag{12}$$

A. Previous Decoding Scheme

In [8], the authors proposed a joint reconstruction and recovery algorithm as in Algorithm 1. Here $F \in \mathbb{R}^{p \times m}$ is the left annihilator matrix of B, i.e., $FB = 0$ with $p = m - M$. Briefly in this iterative method, we first try to remove the watermark via the annihilator matrix F, reconstruct the original signal x sparsely, deflate accordingly the received signal y, and then proceed to extract the binary watermark w via thresholding.

Algorithm 1 Algorithm 1 in [8]

```
**Input:** $y$, $A$, $B$;
**Determine:** $\epsilon$
**1.** Apply $F$ to $y : \tilde{y} = Fy$
**2.** Estimate $\tilde{x} : \tilde{x} = \arg\min_x \|x\|_{\ell_1^N}$  s.t.  $\|\tilde{y} - FAx\|_{\ell_2^p} \leq \epsilon$
**3.** Estimate $\tilde{w} : \tilde{w} = (B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}(y - A\tilde{x})$
**4.** Threshold $\tilde{w} : \hat{w}_i = a * \mathrm{sgn}(\tilde{w}_i)$
**5.** $\hat{x} = \arg\min_x \ \|x\|_{\ell_1^N}$  s.t.  $\|(y - B\hat{w}) - Ax\|_{\ell_2^m} \leq \epsilon$
**Return:** $\hat{x}$, $\hat{w}$
```

It is stated in [8] that the data embedding capacity depends on the restricted isometry constant of FA and the signal to noise ratio (SNR).

B. Proposed Robust Decoding Scheme

1) Problem Formulation: In this work, we formulate the joint estimation of embedded data and sparse signal recovery as an optimization problem:

$$(x^*, w^*) = \arg\min_{(x,w)} \{ \frac{\lambda_1}{2} \|y - (Ax + Bw)\|^2_{\ell_2^m} + \lambda_3 \|x\|_{\ell_1^N} + \frac{\lambda_2}{2} \|Fy - FAx\|^2_{\ell_2^p} + i_{\{-a,+a\}^M}(w) \}. \tag{13}$$

In (13), a second fidelity term, $\|Fy - FAx\|^2_{\ell_2^p}$ is added to increase the solution stability, as it was done in Algorithm 1 along with the first term $\|y - (Ax + Bw)\|^2_{\ell_2^m}$. While the second term increases the stability, the first term is instrumental for the ADMM formulation, and this can be seen as a feed-back mechanism. Finally, the last term represents the projection on a non-convex set and corresponds to the thresholding operation to extract the hidden binary message. The optimization problem (13) becomes a non-convex one due to last term, since w_i is an integer. We can solve this problem using the following ADMM strategy. In the following subsection, we first explain the primal-dual conversion of (13) and then proceed with the ADMM solution for (13).

2) From dual ascent to ADMM: The equivalent consensus form can be written as

$$(x^*, w^*, z_1^*, z_2^*) = \arg\min_{(x,w,z_1,z_2)} \{ \frac{\lambda_1}{2} \|y - (Ax + Bw)\|^2_{\ell_2^m} + \lambda_3 \|z_2\|_{\ell_1^N} + \frac{\lambda_2}{2} \|Fy - FAx\|^2_{\ell_2^p} + i_{\{-a,+a\}^M}(z_1) \} \quad \text{subject to} \ \ w = z_1, x = z_2. \tag{14}$$

The Augmented Lagrangian form for this problem in (14) can be cast as

$$L_{(\mu_1,\mu_2)}(\beta_1, \beta_2, x, w, z_1, z_2) = \frac{\lambda_1}{2} \|y - (Ax + Bw)\|^2_{\ell_2^m} + \lambda_3 \|z_2\|_{\ell_1^N} + \frac{\lambda_2}{2} \|Fy - FAx\|^2_{\ell_2^p} + i_{\{-a,+a\}^M}(z_1) + \langle \beta_1, (z_1 - w) \rangle + \langle \beta_2, (z_2 - x) \rangle + \frac{\mu_1}{2} \|w - z_1\|^2_{\ell_2^M} + \frac{\mu_2}{2} \|x - z_2\|^2_{\ell_2^N}, \tag{15}$$

where $\beta_1 \in \mathbb{R}^M, \beta_2 \in \mathbb{R}^N$ are called Lagrange multipliers or dual variables, and the last two terms are penalty terms with parameters, $\mu_1, \mu_2 > 0$, respectively. The corresponding dual function can be written as

$$g_{(\mu_1,\mu_2)}(\beta_1, \beta_2) = \inf_{(x,w,z_1,z_2)} L_{(\mu_1,\mu_2)}(\beta_1, \beta_2, x, w, z_1, z_2). \tag{16}$$

Instead of the primal problem in (13), we can approximate the optima of the primal function (with some duality gap due to the non-convex term) by maximizing the dual function which is

$$(\beta_1^*, \beta_2^*) = \arg\max_{\beta_1,\beta_2} \left\{ g_{(\mu_1,\mu_2)}(\beta_1, \beta_2) \right\}. \tag{17}$$

Then, we can approximate the primal optimal points by solving following problem:

$$(x^*, w^*, z_1^*, z_2^*) = \arg\min_{(x,w,z_1,z_2)} L_{(\mu_1,\mu_2)}(\beta_1^*, \beta_2^*, x, w, z_1, z_2). \tag{18}$$

In maximizing Problem (17), the primal values are updated jointly. This is called augmented Lagrangian method or method of multipliers [16], which has the following iterative form for our problem,

$$(x, w, z_1, z_2)^{k+1} \leftarrow \arg\min_{(x,w,z_1,z_2)} L_{(\mu_1,\mu_2)}(\beta_1^k, \beta_2^k, .., z_1, z_2)$$

$$\beta_1^{k+1} \leftarrow \beta_1^k + \mu_1(z_1^{k+1} - w^{k+1})$$

$$\beta_2^{k+1} \leftarrow \beta_2^k + \mu_2(z_2^{k+1} - x^{k+1}),$$

where the last two terms come from the gradient ascent step for the dual function and the specific choice of the ascent step

μ_1, μ_2 (these dual variable updates can be done independently since $\nabla_{(\beta_1,\beta_2)} g_{(\mu_1,\mu_2)}(\beta_1, \beta_2)$ are separable in β_1 and β_2).

The joint optimization stated in Eq. (18) can be solved conveniently for the x, w, z_1, z_2 variables using ADMM [17] (see the recent review [9] for details.) We can derive an algorithm in which the primal and dual variables are updated independently in an alternating manner. The general structure of ADMM algorithm for Problem (13) is given in Algorithm 2.

3) Primal Variable Updates: Let us start from the update of z_2,

$$z_2^{k+1} = \arg\min_{z_2} \{\lambda_2 \|z_2\|_{\ell_1^N} + \langle \beta_2^k, (z_2 - x^{k+1}) \rangle + \frac{\mu_2}{2} \|x^{k+1} - z_2\|^2_{\ell_2^N}\} \quad (19)$$

which is actually equivalent to

$$z_2^{k+1} = \arg\min_{z_2} \left\{ \lambda_2 \|z_2\|_{\ell_1^N} + \frac{\mu_2}{2} \left\| z_2 - \left(x^{k+1} - \frac{\beta_2^k}{\mu_2} \right) \right\|^2_{\ell_2^N} \right\}. \quad (20)$$

By the definition in (4), one can easily see that it is nothing but the the proximity operator of $f(x) = \|x\|_{\ell_1^N}$ with the parameter $(\frac{\lambda_2}{\mu_2})$

$$z_2^{k+1} = prox_{(\frac{\lambda_2}{\mu_2})\ell_1^N} \left(x^{k+1} - \frac{\beta_2^k}{\mu_2} \right). \quad (21)$$

Using the separable sum property for proximal maps [11], one can obtain the proximal operator of $f(x) = \|x\|_{\ell_1^N}$ with parameter γ at vector z as follows,

$$prox_{\gamma f}(z_i) = \begin{cases} z_i + \gamma & \text{if } z_i \leq -\gamma \\ 0 & \text{if } -\gamma \leq z_i \leq +\gamma \\ z_i - \gamma & \text{if } z_i \geq \gamma \end{cases}$$

where z_i is the i-th element of the vector z. This operation is actually the well-known soft thresholding. Similarly, the z_1 update can be performed by solving

$$z_1^{k+1} = \arg\min_{z_1} \left\{ i_C(z_1) + \frac{\mu_1}{2} \left\| z_1 - \left(w^{k+1} - \frac{\beta_1^k}{\mu_1} \right) \right\|^2_{\ell_2^M} \right\} \quad (22)$$

Algorithm 2 ADMM for Problem

```
repeat
    Primal Updates
    $x^{k+1} \leftarrow \arg\min_x L_{(\mu_1,\mu_2)}(\beta_1^k, \beta_2^k, x, w^k, z_1^k, z_2^k)$
    $w^{k+1} \leftarrow \arg\min_x L_{(\mu_1,\mu_2)}(\beta_1^k, \beta_2^k, x^{k+1}, w, z_1^k, z_2^k)$
    $z_1^{k+1} \leftarrow \arg\min_x L_{(\mu_1,\mu_2)}(\beta_1^k, \beta_2^k, x^{k+1}, w^{k+1}, z_1, z_2^k)$
    $z_2^{k+1} \leftarrow \arg\min_x L_{(\mu_1,\mu_2)}(\beta_1^k, \beta_2^k, x^{k+1}, w^{k+1}, z_1^{k+1}, z_2)$
    Dual Updates:
    $\beta_1^{k+1} \leftarrow \beta_1^k + \mu_1(z_1^{k+1} - w^{k+1})$
    $\beta_2^{k+1} \leftarrow \beta_2^k + \mu_2(z_2^{k+1} - x^{k+1})$
until Convergence
return $\hat{x}, \hat{w}$
```

and we can easily see that it is the following operator

$$z_1^{k+1} = proj_{\{\{-a,+a\}^M\}} \left(w^{k+1} - \frac{\beta_1^k}{\mu_1} \right) \quad (23)$$

where we can approximate the projection onto the set $C = \{-a, +a\}^M$ as the simple thresholding operator

$$proj_{\{\{-a,+a\}^M\}}(z_i) \approx a * \text{sgn}(z_i) \quad (24)$$

for the i-th element of a vector z. We can see that the operation is a relax-and-round heuristic for integer valued non-convex optimization problem (22) by relaxing the set $\{-a, +a\}$ to $(-a, +a)$, solving the corresponding convex problem and rounding the solution to the nearest integer $-a$ or a [18].

Update of the primal variable x can be done by solving

$$x^{k+1} = \arg\min_x \{ \frac{\lambda_1}{2} \|y - (Ax + Bw^k)\|^2_{\ell_2^m} + \langle \beta_2, (x - z_2^k) \rangle + \frac{\lambda_2}{2} \|Fy - FAx\|^2_{\ell_2^p} + \frac{\mu_2}{2} \|x - z_2^k\|^2_{\ell_2^N} \}. \quad (25)$$

Since the right hand side is differentiable, the update equation can be cast as solving the linear equation $\nabla_x L(.) = 0$, which reduces to

$$x^{k+1} = (\lambda_1 A^T A + \lambda_2 A^T F^T F A + I\mu_2)^{-1} (\lambda_1 A^T (y - Bw^k) + \lambda_2 (A^T F^T F y) + \beta_2^k + \mu_2 z_2^k) \quad (26)$$

Similarly, the update of primal variable w can be achieved by solving

$$w^{k+1} = (\arg\min_w \frac{\lambda_1}{2} \|y - (Ax^{k+1} + Bw)\|^2_{\ell_2^m} + \langle \beta_1^k, (z_1^k - w) \rangle + \frac{\mu_1}{2} \|w - z_1^k\|^2_{\ell_2^M}) \quad (27)$$

which yields

$$w^{k+1} = (\lambda_1 B^T B + \mu_1 I)^{-1} \left[\lambda_1 B^T (y - Ax^{k+1}) + \beta_1^k + \mu_1 z_1^k \right] \quad (28)$$

by solving $\nabla_w L(.) = 0$. In addition to these variables, robustness parameters (μ_1, μ_2) can also be updated,

$$(\mu_1^{k+1}, \mu_2^{k+1}) \leftarrow \left(\rho_1 \mu_1^k, \rho_2 \mu_2^k \right). \quad (29)$$

V. Simulation Results

We generate a $k = \frac{m}{5}$-sparse $N = 512$ length synthetic signal. A and F are chosen as explained in Section IV. The M-long watermark, w, is generated with $\|w\|_{\ell_2^M} = \frac{\|Ax\|_{\ell_2^m}}{4}$ so that the embedded-data-to document ratio is -6 dB and the marked measurements are contaminated with AWGN with different signal-to-noise ratios (SNR). We define the SNR as $20 \log_{10} \left(\frac{\|Ax+Bw\|_{\ell_2^m}}{\|n\|_{\ell_2^m}} \right)$. Each experiment is conducted 250 times and the average performance results are reported. Similar experiments are also conducted with different sparsity level and SNR, but for the sake of brevity we report only the cases for 32 dB and 24 dB in Figure 1 and Figure 2, respectively.

ISBN 978-90-827970-1-5 © EURASIP 2018

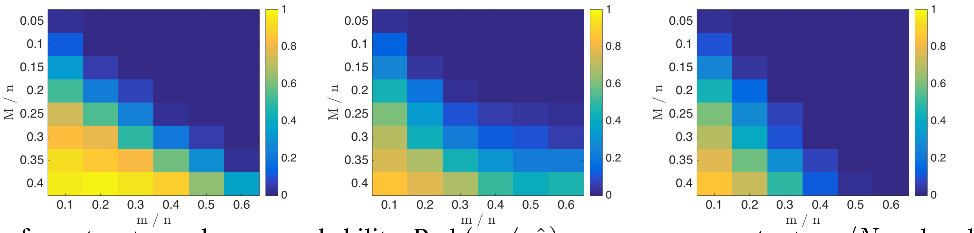

Fig. 1: Heat-maps of exact watermark error probability, Prob$(w \neq \hat{w})$ over measurement rate m/N and embedding rate M/m under AWGN at 32 db SNR. (a) Algorithm 1. (b) The proposed method without the term $\|Fy - FAx\|^2_{\ell_2^p}$. (c) The proposed method.

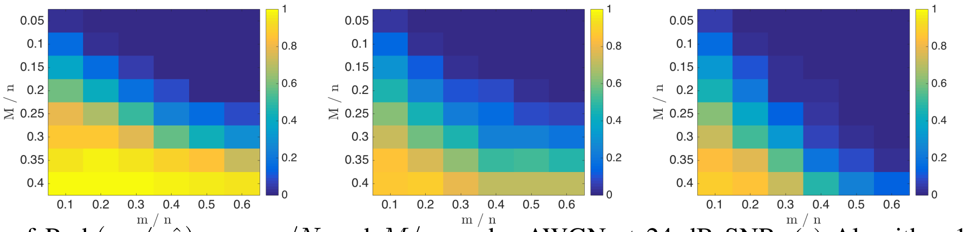

Fig. 2: Heat-maps of Prob$(w \neq \hat{w})$ over m/N and M/m under AWGN at 24 dB SNR. (a) Algorithm 1. (b) The proposed method without the term $\|Fy - FAx\|^2_{\ell_2^p}$. (c) The proposed method.

Similar performance was observed in the experiments not reported here. We compare the performance results of Algorithm 1, the proposed Algorithm with and without the data fidelity term $\|Fy - FAx\|^2_{\ell_2^p}$ for different embedding rates, $\frac{M}{m}$ and the different measurement rates $\frac{m}{N}$. From Figure 1-(b), (c) and Figure 2-(b), (c), it can be clearly seen that modeling the problem as joint optimization problem in (13) which includes an extra data fidelity term $\|Fy - FAx\|^2_{\ell_2^p}$, which is in dimensionality reduced measurement domain clearly surpasses the performance of modeling without it. It is also apparent from the Figure 1-(a), (c) and Figure 2-(a), (c) that the performance of final ADMM based solution to (13) exceeds the previous state of art [8].

In our experiments we set $\lambda_1 = \lambda_2 = 1$, $\lambda_3 = 1 \times 10^{-2}$, $\mu_1 = 3.3 \times 10^{-6}$, $\mu_2 = 8 \times 10^{-3}$, $\rho_1 = 1$, $\rho_2 = 1.035$.

ℓ_1-magic [19] solver (it was observed that different solvers, such as CVX, resulted in a similar performance) is used to conduct Algorithm Algorithm 1 with $\epsilon \approx \sqrt{m\sigma_n^2}$.

VI. Conclusion

In this work we have proposed a new iterative decoding strategy for joint watermark extraction and signal recovery in compressively sampled signals. The new approach boosts the watermark capacity of the compressively sampled signals and improves also its noise robustness. We plan to extend this framework to embed meta-data onto CSMs to structurally sparse signals such as group sparse ones.

References

[1] Candes, Emmanuel J. "Compressive sampling." Proceedings of the international congress of mathematicians. Vol. 3. 2006.

[2] Lustig, Michael, David Donoho, and John M. Pauly. "Sparse MRI: The application of compressed sensing for rapid MR imaging." Magnetic resonance in medicine 58.6 (2007): 1182-1195.

[3] Mamaghanian, Hossein, et al. "Compressed sensing for real-time energy-efficient ECG compression on wireless body sensor nodes." IEEE Transactions on Biomedical Engineering 58.9 (2011): 2456-2466.

[4] Takhar, Dharmpal, et al. "A new compressive imaging camera architecture using optical-domain compression." Computational Imaging IV. Vol. 6065. International Society for Optics and Photonics, 2006.

[5] Oike, Yusuke, and Abbas El Gamal. "CMOS image sensor with per-column ADC and programmable compressed sensing." IEEE Journal of Solid-State Circuits 48.1 (2013): 318-328.

[6] Kozat, Suleyman S., et al. "Embedding and retrieving private metadata in electrocardiograms." Journal of medical systems 33.4 (2009): 241-259.

[7] Orsdemir, Adem, et al. "On the security and robustness of encryption via compressed sensing." Military Communications Conference, 2008. MILCOM 2008. IEEE. IEEE, 2008.

[8] Yamaç Mehmet, Çağatay Dikici, and Bülent Sankur. "Hiding data in compressive sensed measurements: A conditionally reversible data hiding scheme for compressively sensed measurements." Digital Signal Processing 48 (2016): 188-200.

[9] Boyd, Stephen, et al. "Distributed optimization and statistical learning via the alternating direction method of multipliers." Foundations and Trends in Machine Learning 3.1 (2011): 1-122.

[10] Combettes, Patrick L., and Jean-Christophe Pesquet. "Proximal splitting methods in signal processing." Fixed-point algorithms for inverse problems in science and engineering. Springer New York, 2011. 185-212.

[11] Parikh, Neal, and Stephen P. Boyd. "Proximal Algorithms." Foundations and Trends in optimization 1.3 (2014): 127-239.

[12] Moreau, Jean-Jacques. "Proximité et dualité dans un espace hilbertien." Bull. Soc. Math. France 93.2 (1965): 273-299.

[13] Cohen, Albert, Wolfgang Dahmen, and Ronald DeVore. "Compressed sensing and best -term approximation." Journal of the American mathematical society 22.1 (2009): 211-231.

[14] Chen, Scott Shaobing, David L. Donoho, and Michael A. Saunders. "Atomic decomposition by basis pursuit." SIAM review 43.1 (2001): 129-159.

[15] Candes, Emmanuel J. "The restricted isometry property and its implications for compressed sensing." Comptes rendus mathematique 346.9-10 (2008): 589-592.

[16] Hestenes, Magnus R. "Multiplier and gradient methods." Journal of optimization theory and applications 4.5 (1969): 303-320.

[17] Gabay, Daniel, and Bertrand Mercier. "A dual algorithm for the solution of nonlinear variational problems via finite element approximation." Computers & Mathematics with Applications 2.1 (1976): 17-40.

[18] Wu, Baoyuan, and Bernard Ghanem. "ℓ_p-Box ADMM: A Versatile Framework for Integer Programming." arXiv preprint arXiv:1604.07666 (2016).

[19] Candes, Emmanuel, and Justin Romberg. "l1-magic: Recovery of sparse signals via convex programming." URL: www. acm. caltech. edu/l1magic/downloads/l1magic. pdf 4 (2005): 14.