When Compressive Sensing Meets Data Hiding

Guang Hua, *Member, IEEE*, Yong Xiang, *Senior Member, IEEE*, and Guoan Bi, *Senior Member, IEEE*

Abstract

We present a novel framework of performing multimedia data hiding using an over-complete dictionary, which brings compressive sensing to the application of data hiding. Unlike the conventional orthonormal full-space dictionary, the over-complete dictionary produces an underdetermined system with infinite transform results. We first discuss the minimum norm formulation (ℓ_2-norm) which yields a closed-form solution and the concept of watermark projection, so that higher embedding capacity and an additional privacy preserving feature can be obtained. Furthermore, we study the sparse formulation (ℓ_0-norm) and illustrate that as long as the ℓ_0-norm of the sparse representation of the host signal is less than the signal's dimension in the original domain, an informed sparse domain data hiding system can be established by modifying the coefficients of the atoms that have not participated in representing the host signal. A single support modification based data hiding system is then proposed and analyzed as an example. Several potential research directions are discussed for further studies. More generally, apart from the ℓ_2- and ℓ_0-norm constraints, other conditions for reliable detection performance are worth of future investigation.

Index Terms

Compressive sensing, compressed sensing, sparsity, data hiding, watermarking.

I. Introduction

CONSIDER a generic signal model for conventional data hiding in multimedia content, in which the original domain host signal in vector form, $\mathbf{x} \in \mathbb{R}^{M\times 1}$, is first expressed by a unique

This work is supported by Ministry of Education (MOE), Singapore, under the research grant MOE2014-T2-1-079.

G. Hua is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore (e-mail:ghua@ntu.edu.sg).

Y. Xiang is with the School of Information Technology, Deakin University, Australia (e-mail:yong.xiang@deakin.edu.au).

G. Bi is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore (e-mail:egbi@ntu.edu.sg).

transform domain vector, $\mathbf{z} \in \mathbb{C}^{M\times 1}$, via an orthonormal full-space transform dictionary, $\mathbf{H} \in \mathbb{C}^{M\times M}$, i.e.,

$$\mathbf{x} = \mathbf{H}\mathbf{z}, \tag{1}$$

where $\mathbf{H}$ can take from the inverse forms of the widely used discrete Fourier transform (DFT), discrete cosine transform (DCT), or modulated complex lapped transform (MCLT) matrix, etc. Then, the hidden data, also referred to as watermarks, are embedded into $\mathbf{z}$. Specifically, denote the information bit to be embedded as $b \in \{-1, +1\}$, then it is modulated by a random sequence $\mathbf{m} \in \mathbb{C}^{M\times 1}$, and the watermarked signal in the original domain is obtained by

$$\tilde{\mathbf{x}} = \mathbf{H}(\mathbf{z} + \alpha b\mathbf{m}), \tag{2}$$

where α controls the embedding strength. The hidden information bit is detected (extracted) via the following hypothesis test

$$r = \mathrm{sgn}\langle \mathbf{m}, \mathbf{H}^{-1}\widetilde{\mathbf{x}}\rangle = \mathrm{sgn}\left(\langle \mathbf{m}, \mathbf{z}\rangle + \alpha b \left\|\mathbf{m}\right\|_2^2\right), \tag{3}$$

where $\mathrm{sgn}(\cdot)$ is the sign function, $\|\cdot\|_p$ denotes ℓ_p-norm, and $\langle\cdot,\cdot\rangle$ represents inner product. It should be noted that sometimes b is omitted and $\mathbf{m}$ individually serves as the watermark information. In such cases, the hidden data are like a copyright mark, and watermark detection (3) is modified to a correlation based detection problem. If $\mathbf{m}$ is embedded without considering the properties of $\mathbf{z}$, then (1)-(3) represent the classical *non-informed* spread spectrum watermarking system [1], where the host signal is considered as inteference during watermark detection. On the other hand, if $\mathbf{z}$ is exploited to facilitate the embedding of $\mathbf{m}$ and reduce the interference term $\langle \mathbf{m}, \mathbf{z}\rangle$, then an *informed* watermarking system is established [2], [3]. While rich literature can be found to explore specific designs of (2) and (3) for different multimedia formats and designing criteria, the transform matrix $\mathbf{H}$ is usually predefined and treated as a constant parameter.

In this letter, we recast the design of the above data hiding system from reconsidering the transform model in (1) by replacing the square orthonormal transform matrix $\mathbf{H}$ with an over-complete dictionary $\mathbf{D}$ with full row rank. Without loss of generality, we assume all the quantities in this letter are real-valued, thus $\mathbf{D} \in \mathbb{R}^{M\times N}$, where $M < N$, and the columns have unit norm,

$$\|\mathbf{d}_i\|_2^2 = 1, \quad i = 0, 1, \dots, N-1. \tag{4}$$

Such a modification is in fact a relaxation of the transform matrix, and it also connects the concept of compressive sensing with the application of data hiding. Let $\mathbf{y} \in \mathbb{R}^{N\times 1}$ be the transform domain representation of $\mathbf{x}$ using $\mathbf{D}$, then we have

$$\mathbf{x} = \mathbf{D}\mathbf{y}. \tag{5}$$

A primary problem arising from the new transformation is that $\mathbf{y}$ is not unique. Let the watermark sequence for model (5) be $\mathbf{w} \in \mathbb{R}^{N\times 1}$, then the watermarked signal is given by

$$\tilde{\mathbf{x}} = \mathbf{D}(\mathbf{y} + \alpha b\mathbf{w}). \tag{6}$$

Note that the data hiding process (6) is analogous to a lossy compression, where the information about $\mathbf{w}$ and the distinction between $\mathbf{y}$ and $\mathbf{w}$ may be lost. In addition, $\mathbf{w}$ can take arbitrary forms, which is even less constrained than $\mathbf{y}$. Then, it is unrealistic to directly solve $\tilde{\mathbf{y}}$ in

$$\tilde{\mathbf{x}} = \mathbf{D}\tilde{\mathbf{y}}, \tag{7}$$

and expect that the solution of $\tilde{\mathbf{y}}$ contains the component $\mathbf{w}$. Therefore, extra conditions are to be imposed to (7) to ensure the uniqueness of the solution. The uncertainty of underdetermined systems actually yields more flexible designs of data hiding systems, which brings in new features unavailable in conventional systems.

In section II, we formulate the system design as a minimum norm problem (ℓ_2-norm) to ensure the one-to-one mapping between the forward and inverse transforms using dictionary $\mathbf{D}$. The solution of this problem introduces the concept of watermark projection which can add the privacy preserving feature to the system. In section III, we provide some preliminary results based on a sparse formulation (ℓ_0-norm) of the system, and describe a design example that modifies a single support in the sparse representation of the host signal. Other design possibilities are also addressed.

II. Minimum Norm Solution - Watermark Projection

The first attempt to enable the traceability of $\mathbf{y}$ and $\mathbf{w}$ is via a minimum norm formulation, which leads to a closed-form solution and the concept of watermark projection. We then demonstrate that a privacy preserving system can be achieved via the minimum norm formulation. Let the minimum norm solution of $\mathbf{y}$ be $\mathbf{y}_{\ell_2}$, then (5) can be neatly solved via pseudo inverse,

$$\mathbf{y}_{\ell_2} = \mathbf{D}^T(\mathbf{D}\mathbf{D}^T)^{-1}\mathbf{x} \triangleq \mathbf{D}^\dagger \mathbf{x}, \tag{8}$$

where $\{\cdot\}^T$ is the transpose operator. Furthermore, to ensure successful watermark detection, a similar minimum norm formulation is applied to solve the transform of $\tilde{\mathbf{x}}$, i.e.,

$$\hat{\mathbf{y}} = \mathbf{D}^\dagger \tilde{\mathbf{x}} = \mathbf{D}^\dagger \mathbf{D}(\mathbf{y}_{\ell_2} + \alpha b \mathbf{w}) = \mathbf{y}_{\ell_2} + \alpha b \mathbf{w}_{\mathrm{P}}, \tag{9}$$

where we further define a decomposition pair of $\mathbf{w}$ as

$$\mathbf{w}_{\mathrm{P}} = \mathbf{D}^\dagger \mathbf{D}\mathbf{w}, \tag{10}$$

$$\mathbf{w}_{\mathrm{O}} = \mathbf{w} - \mathbf{w}_{\mathrm{P}}. \tag{11}$$

It is then indicated that only if $\mathbf{w}$ also lies within the row space of $\mathbf{D}$, i.e., $\mathbf{w}_{\mathrm{P}} = \mathbf{w}$, can the full knowledge of $\mathbf{w}$ be contained in (9). However, the two most common forms of $\mathbf{w}$, i.e., a pseudorandom sequence or a vectorized pattern signal, are generated without considering the above condition. In fact, none of the existing data hiding systems considers the relationship between the transform dictionary and the watermark, because this is a trivial problem when $\mathbf{H}$ is an orthonormal dictionary. The hidden information bit is extracted via

$$r = \operatorname{sgn}\left\langle \mathbf{w}_{\mathrm{P}}, \mathbf{D}^\dagger \tilde{\mathbf{x}} \right\rangle = \operatorname{sgn}\left(\langle \mathbf{w}_{\mathrm{P}}, \mathbf{y}_{\ell_2} \rangle + \alpha b \left\| \mathbf{w}_{\mathrm{P}} \right\|_2^2 \right), \tag{12}$$

which is similar to (3). Comparing conventional data hiding system using $\mathbf{H}$ and the alternative system using $\mathbf{D}$, we see very similar performance in terms of embedding distortion and detection robustness, which are characterized by the second and first terms at the right hand side of (3) and (12) respectively. However, in terms of embedding capacity, since $\operatorname{card}(\mathbf{w}) = N > M = \operatorname{card}(\mathbf{m})$, more data can be hidden in the alternative system. Note that N can be arbitrarily chosen while M is determined by the host signal.

Now, we can think of a privacy preserving system by omitting b and let $\mathbf{w}$ be the vectorized version of the binary mark pattern in Fig. 1 (a). A data owner wants to share some data with others. The owner embeds the identification pattern's projection data as shown in Fig. 1 (b) or (c) into the host data before sharing. In the meantime, for some users, the owner wants to authorize them the copyright but not willing to let them know the embedded information. To ensure the privacy, the owner passes $\tilde{\mathbf{x}}$, $\mathbf{D}$, and $\mathbf{w}_{\mathrm{P}}$, to them. Therefore, these users can use the provided

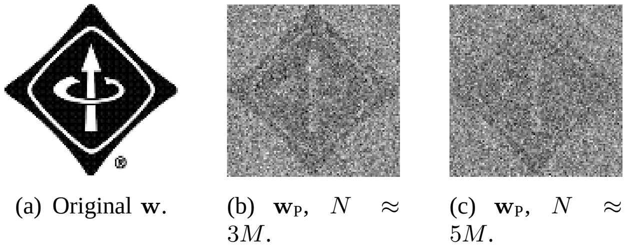

(a) Original $\mathbf{w}$. (b) $\mathbf{w}_{\mathrm{P}}$, $N \approx 3M$. (c) $\mathbf{w}_{\mathrm{P}}$, $N \approx 5M$.

Fig. 1. An illustration of watermark projection and comprehensibility.

information to claim that the copies are authorized, but they only know the incomprehensible $\mathbf{w}_{\mathrm{P}}$ instead of $\mathbf{w}$. For the other users, the owner has no problem of sharing the mark pattern, thus, $\tilde{\mathbf{x}}$, $\mathbf{D}$, and $\mathbf{w}$, are passed to these users. In this example, we can see that the dictionary $\mathbf{D}$ has become a component of the copyright mark, which is not applicable in conventional system. More importantly, the decomposed watermark $\mathbf{w}_{\mathrm{P}}$ enables an effective means to deal with the first group of users, which is also not possible in conventional cases. It can also be seen from Fig. 1 that increasing the ratio of M/N could reduce the comprehensibility of the watermark pattern, but this is achieved by increasing the computation load because M and N determine the dimension of $\mathbf{D}$. However, since the embedding and extraction of hidden data could be performed off-line, it is reasonable to conclude that the privacy preserving feature is more important than computational efficiency. Next, we discuss the system design in a compressive sensing framework.

III. Sparse Solution

A. Feasibility and Applicability

Let the sparse solution of (5) be $\mathbf{y}_{\ell_0}$, then

$$\mathbf{y}_{\ell_0} = \arg\min_{\mathbf{y}} \|\mathbf{y}\|_0 \quad \text{s.t.} \quad \mathbf{x} = \mathbf{D}\mathbf{y}. \tag{13}$$

Note that finding the sparsest solution to the above problem is not easy, which involves the selection of $\mathbf{D}$ (e.g., [4], [5]) and what iterative algorithm (e.g., [6], [7], or ℓ_1-norm relaxed [8], [9]) is used to find $\mathbf{y}_{\ell_0}$. However, a data hiding system based on the above model is not quite interested in whether $\mathbf{y}_{\ell_0}$ is globally sparsest. Instead, it is more important here to ensure that $\mathbf{w}$ can be unambiguously recovered during watermark detection. If we assume that $\mathbf{w}$ is also sparse, and expect that it can be recovered from

$$\tilde{\mathbf{x}} = \mathbf{D}(\mathbf{y}_{\ell_0} + \alpha b \mathbf{w}) \tag{14}$$

using the same algorithm that solves (13) or another one if necessary, then we are facing the stringent conditions for the exact recovery of $\mathbf{y}_{\ell_0} + \alpha b \mathbf{w}$ from $\tilde{\mathbf{x}}$ [10]–[12]. Specifically, the two features, i.e., mutual incoherence property (MIP) [6] and restricted isometry property (RIP) [13], are widely used to analyze the exact recovery performance. Here, we use the more computationally efficient MIP, which is defined by [6]

$$\mu \triangleq \max_{i \neq j} |\langle \mathbf{d}_i, \mathbf{d}_j \rangle|, \quad i, j = 0, 1, \ldots, N-1. \tag{15}$$

The condition for the exact recovery of $\mathbf{y}_{\boldsymbol{\ell}_\mathbf{0}}$ from $\mathbf{x}$ is then given by [6]

$$K \triangleq \operatorname{card}(\mathbf{y}_{\ell_0}) < \frac{\mu + 1}{2\mu}. \tag{16}$$

If we use a dictionary of independent and identically distributed (i.i.d.) Gaussian variables with $M = 128$ and $N = 256$ (a very good condition), then the condition for $\mathbf{y}_{\ell_0} + \alpha b \mathbf{w}$ to be exactly recovered from $\tilde{\mathbf{x}}$ is $\operatorname{card}(\mathbf{y}_{\ell_0} + \alpha b \mathbf{w}) = 1$, which is impossible. Therefore, watermark detection via sparse support recovery is generally inapplicable to the system. However, we will show that watermark detection can be reliably achieved via an alternative means that exploits the sparsity of $\mathbf{y}_{\ell_0}$, and the only condition required on $\mathbf{y}_{\ell_0}$ is

$$K < M. \tag{17}$$

This condition is somewhat trivial because if $\operatorname{card}(\mathbf{y}_{\ell_0}) = M$, then (13) reduces to (1), and $\operatorname{card}(\mathbf{y}_{\ell_0}) > M$ never holds for a full row rank dictionary. Condition (17) also implies that all the existing sparse recovery algorithms can be applied to the data hiding system. We illustrate in the next subsection how to make use of (17) to design an effective data hiding system.

B. A Single Support Modification Design Example

Let $\mathbf{k} \triangleq [k_0, k_1, \ldots, k_{K-1}]$ be the set of column indices of $\mathbf{D}$ that correspond to the nonzero elements of $\mathbf{y}_{\ell_0}$, then $\mathbf{x} = \mathbf{D}\mathbf{y}_{\ell_0}$ can be rewritten in a compact form

$$\mathbf{x} = \boldsymbol{\Phi}\dot{\mathbf{y}}, \tag{18}$$

Algorithm 1: Single Support Modification - Embedding

```
  Input: $\mathbf{x}$, $\mathbf{D}$, $\alpha$, $b = \pm 1$;
  Output: $\tilde{\mathbf{x}}$;
1 Initialization: $\mathbf{w} = [w_0, w_1, \dots, w_{N-1}]^T \leftarrow \mathbf{0}$;
2 $\mathbf{y}_{\ell_0} \leftarrow \arg\min_{\mathbf{y}} \|\mathbf{y}\|_0 \quad \text{s.t.} \quad \mathbf{x} = \mathbf{D}\mathbf{y}.$ ;
3 $\mathbf{\Phi} \leftarrow [\mathbf{d}_{k_0}, \dots, \mathbf{d}_{k_{K-1}}]$, $\mathbf{\Psi} \leftarrow [\mathbf{d}_{l_0}, \dots, \mathbf{d}_{l_{N-K-1}}]$;
4 $\mathbf{\Delta} \leftarrow \text{null}(\mathbf{\Phi}^T) \quad \text{s.t.} \quad \mathbf{\Delta}^T\mathbf{\Delta} = \mathbf{I}$;
5 $l_w \leftarrow \arg\max_{l_i} \|\mathbf{\Delta}^\dagger \mathbf{\Delta}^T \mathbf{d}_{l_i}\|_2^2, \quad i = 0, 1, \dots, N-K-1$;
6 $w_{l_w} \leftarrow \alpha b$;
7 $\tilde{\mathbf{x}} \leftarrow \mathbf{D}(\mathbf{y}_{\ell_0} + \mathbf{w})$;
```

where

$$\mathbf{\Phi} \triangleq [\mathbf{d}_{k_0}, \mathbf{d}_{k_1}, \dots, \mathbf{d}_{k_{K-1}}] \in \mathbb{R}^{M \times K}, \tag{19}$$

and $\dot{\mathbf{y}} \in \mathbb{R}^{K \times 1}$ is composed by selecting the nonzero elements of $\mathbf{y}_{\ell_0}$. According to (17), $\mathbf{\Phi}$ is a tall matrix with full column rank. Otherwise (18) can be further compressed. The columns of $\mathbf{D}$ that have not participated in the above linear combination is then denoted by $\mathbf{\Psi}$,

$$\mathbf{\Psi} \triangleq [\mathbf{d}_{l_0}, \mathbf{d}_{l_1}, \dots, \mathbf{d}_{l_{N-K-1}}] \in \mathbb{R}^{M \times (N-K)}, \tag{20}$$

where $\mathbf{l} \triangleq [l_0, l_1, \dots, l_{N-K-1}]$ does not overlap with $\mathbf{k}$. Considering $\mathbf{\Phi}$, $\mathbf{\Psi}$, and $\mathbf{D}$, as different sets of $\mathbf{d}_i$, then it follows $\mathbf{\Phi} \cup \mathbf{\Psi} = \mathbf{D}$ and $\mathbf{\Phi} \cap \mathbf{\Psi} = \emptyset$. Because $\text{rank}(\mathbf{D}) = M$, $\text{rank}(\mathbf{\Phi}) = K$, and $K < M$, it is indicated that the column space of $\mathbf{\Psi}$ intersects with the null space of $\mathbf{\Phi}^T$. We use $\mathbf{\Delta} \in \mathbb{R}^{M \times (M-K)}$ to denote the orthonormal basis of the null space of $\mathbf{\Phi}^T$ such that $\mathbf{\Phi}^T\mathbf{\Delta} = \mathbf{0}$. Using the above parameters, the watermark embedding and detection algorithms for a single support modification system can be designed as Algorithm 1 and Algorithm 2 respectively.

In Algorithm 1, a single watermark chip b is inserted into index l_w which has not been used to represent $\mathbf{x}$. In this setup, conventional modulation sequences, i.e., random realizations of $\mathbf{w}$ are not needed. Instead, αb is absorbed in $\mathbf{w}$ such that

$$\mathbf{w} = [0, 0, \dots, w_{l_w}(= \alpha b), \dots, 0]^T. \tag{21}$$

Note that the atom at the selected index has the strongest projection in $\mathbf{\Delta}$ (Step 5). In fact, multiple chips can be inserted as long as the indices for insertion have not been used to represent $\mathbf{x}$. In addition, the insertion formula, i.e., Step 6 of Algorithm 1 can also take other forms. It is then suggested that the flexibility of choosing insertion locations and inserting formula can

Algorithm 2: Single Support Modification - Detection

```
**Input**: $\tilde{\mathbf{x}}$, $\mathbf{D}$, $\mathbf{k}$;
**Output**: $\hat{b}$;
1 $\boldsymbol{\Phi} \leftarrow [\mathbf{d}_{k_0}, \ldots, \mathbf{d}_{k_{K-1}}]$;
2 $\boldsymbol{\Delta} \leftarrow \mathrm{null}(\boldsymbol{\Phi}^T)$ s.t. $\boldsymbol{\Delta}^T \boldsymbol{\Delta} = \mathbf{I}$;
3 $\tilde{\mathbf{x}}_{\mathrm{P}} \leftarrow \boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T \tilde{\mathbf{x}}$;
4 $\hat{\mathbf{w}} \leftarrow \arg\min_{\tilde{\mathbf{y}}} \|\tilde{\mathbf{y}}\|_0$ s.t. $\mathbf{x}_{\mathrm{P}} = \boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T \mathbf{D} \tilde{\mathbf{y}}$;
5 $\hat{b} \leftarrow \mathrm{sgn}(\hat{w}_j)$, $j = \arg\min_i i$ s.t. $\hat{w}_i \neq 0$;
```

improve the system performance, but here, we focus on the simplest design to illustrate a simple, feasible, and effective data hiding system in sparse domain, whose crucial component is the detection algorithm.

In Algorithm 2, the watermarked signal $\tilde{\mathbf{x}}$ is first projected into the column space of $\boldsymbol{\Delta}$ via Step 3. Due to $\boldsymbol{\Delta}^T \boldsymbol{\Phi} = \mathbf{0}$, we can have the following explicit expression

$$\tilde{\mathbf{x}}_{\mathrm{P}} = \boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T \mathbf{D}(\mathbf{y}_{\ell_0} + \mathbf{w}) = \boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T \mathbf{D} \mathbf{w}, \tag{22}$$

meaning that the projection procedure removes the host interference and forms a new underdertermined linear system with a modified dictionary $\boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T \mathbf{D}$, whose RIP is rigorously discussed in [14]. Note that the problem of identifying $\mathbf{w}$ from $\tilde{\mathbf{x}}_{\mathrm{P}}$ can always be efficiently solved because $\mathrm{card}(\mathbf{w}) = 1$, which satisfies the MIP condition (16). Therefore, we can take the simplest method to implement Step 4 of Algorithm 2, e.g., orthogonal matching pursuit (OMP) [6] with single iteration. Mathematically, the single iteration OMP first calculates

$$\begin{aligned}\tilde{\mathbf{x}}_{\mathrm{P}}^T \boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T \mathbf{D} =& \mathbf{w}^T \mathbf{D}^T \boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T \boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T \mathbf{D} \\ =& \alpha b \mathbf{d}_{l_w}^T \boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T [\mathbf{d}_0, \ldots, \mathbf{d}_{l_w}, \ldots, \mathbf{d}_{N-1}],\end{aligned} \tag{23}$$

where $(\boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T)(\boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T) \ldots (\boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T) = \boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T$ is the property of a projection matrix, and it is easy to detect l_w since

$$\max(\tilde{\mathbf{x}}_{\mathrm{P}}^T \boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T \mathbf{D}) = w_{l_w} \mathbf{d}_{l_w}^T \boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T \mathbf{d}_{l_w}, \tag{24}$$

thus the strongest supporting atom is $\hat{\mathbf{d}} \triangleq \boldsymbol{\Delta}^\dagger \boldsymbol{\Delta}^T \mathbf{d}_{l_w}$. Then, the OMP algorithm projects $\tilde{\mathbf{x}}_{\mathrm{P}}$ onto this atom and obtains the single support value of $\hat{\mathbf{w}}$, and we have $\hat{w} = \mathrm{sgn}(\alpha \bar{w}) = \mathrm{sgn}(b)$, an exact recovery of the information bit. Such system can be considered as a (semi-) informed system, where partial information about the host signal, i.e., $\mathbf{k}$, is needed during the detection phase. The “space” for data hiding is discovered in (17) which leads to the construction of the null space basis $\mathbf{\Delta}$. Note that this is inapplicable in conventional cases where the exact complete dictionary $\mathbf{H}$ does not have such a “space”. Note that if we also make l_w available at detection phase, then a more robust detection can be achieved because possible error in detecting l_w can be avoided in noisy conditions.

C. Noisy Case

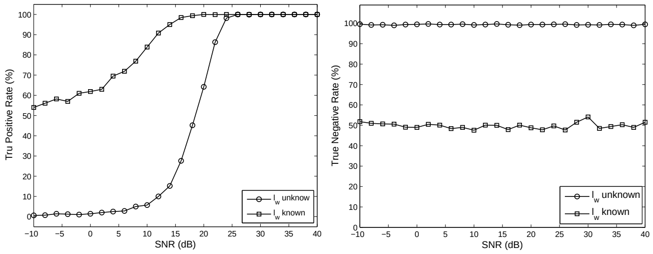
In noise-free situations, host interference is rejected by the use of $\mathbf{\Delta}$, and successful detection is guaranteed for an arbitrary $\alpha > 0$. However, in a noisy condition, (22) becomes

$$\tilde{\mathbf{x}}_{\mathrm{P}} = \mathbf{\Delta}^{\dagger}\mathbf{\Delta}^{T}[\mathbf{D}(\mathbf{y}_{\ell_0} + \mathbf{w}) + \mathbf{v}] = \mathbf{\Delta}^{\dagger}\mathbf{\Delta}^{T}\mathbf{D}\mathbf{w} + \mathbf{v}_{\mathrm{P}}, \tag{25}$$

where $\mathbf{v}$ is additive white Gaussian noise (AWGN) and $\mathbf{v}_{\mathrm{P}}$ is its projection on $\mathbf{\Delta}$. If the detection is performed without the knowledge of l_w, then Step 4 of Algorithm 2 will be affected by $\mathbf{v}_{\mathrm{P}}$, and l_w could be wrongly identified. However, if l_w is known, then the scalar projection of (25) onto $\hat{\mathbf{d}}$ yields

$$\hat{w} = \mathrm{sgn}\left(\alpha b + \frac{\mathbf{d}_{l_w}^{T}\mathbf{\Delta}^{\dagger}\mathbf{\Delta}^{T}\mathbf{v}}{\mathbf{d}_{l_w}^{T}\mathbf{\Delta}^{\dagger}\mathbf{\Delta}^{T}\mathbf{d}_{l_w}}\right), \tag{26}$$

where the term at the right hand side in the parenthesis is the interference. Fig. 2 shows the synthetic analysis results of the system in noisy environments obtained by averaging 1000 realizations of AWGN, where the host signal $\mathbf{x}$ and dictionary $\mathbf{D}$ are both generated by i.i.d. Gaussian random variables with normal distribution, having $K \approx 100$. The advantage of using l_w during the detection is evidenced in Fig. 2 (a). However, we can observe from Fig. 2 (b) that detection without using l_w has nearly 100% accuracy in detecting non-watermarked signals. In other words, it is almost free from false positive decisions. In contrast, the detection using l_w yields a random guess when examining signals that have not been watermarked. Generally, Fig. 2 (b) reflects system performance when watermark is not embedded. In this case, signal at the known index is not marked at all (αb does not exist in (26) anymore), and the corresponding value is solely governed by noise. Therefore, the detection performance turns out to be like random guesses with around 50% correct rate. In contrast, if the index is unknown, then Steps 4 and 5 of Algorithm 2 are solely governed by noise, and the strongest support becomes random. Thus, the possibility of the strongest support index being coincident with the known index becomes



(a) Sensitivity.

(b) Specificity.

Fig. 2. Sensitivity (a) and specificity (b) study of the single support modification system under different SNR values, where $M = 128$, $N = 1024$, $\alpha = 0.5$, $K \approx 100$, and the resultant embedding distortion is 27 ± 0.5 dB.

extremely low (roughly speaking, less than $1/(N - K)$), and the true negative rate approaches 100%.

D. Variations

The system presented and analyzed above is the first attempt to validate the concept of data hiding in sparse domain. Based on the framework described in this letter, several variations can be considered for alternative design purposes. First, to enhance the security of the system, one can employ multiple dictionaries. For example, a series of dictionaries can be obtained by $\tilde{\mathbf{D}}_i = \underbrace{\mathbf{R} \cdots \mathbf{R}}_{i} \mathbf{D}, \quad i = 1, 2, 3, \cdots$, where $\mathbf{R} \in \mathbb{R}^{M \times M}$ is an orthogonal matrix. It can be easily verified that $\forall i < \infty$, each column of $\tilde{\mathbf{D}}_i$ has unit norm. Using this setting, only two matrices need to be stored, but many dictionaries can be generated, and it becomes impossible to extract the hidden information without both knowledge of $\mathbf{D}$ and $\mathbf{R}$. Apart from the flexibility in choosing the dictionaries, it is also possible to design a spread spectrum like data hiding mechanism with the inclusion of watermark projection discussed in Section II. Specifically, it can be seen from (20) and (22) that data can be hidden at multiple locations irrelevant to $\mathbf{x}$, indexed by $[w_{l_0}, w_{l_1}, \ldots, w_{l_{N-K-1}}]$. This indicates that $N - K$ samples can be used for data embedding, which can usually be greater than M. We can also observe from (22) that the extracted portion is a projection of the original domain watermark, i.e., $\mathbf{D}\mathbf{w}$, onto the column space of $\mathbf{\Delta}$.

IV. Conclusion

The use of an over-complete dictionary to perform forward and inverse transforms in multimedia data hiding systems forms an underdetermined linear function with infinite solutions, which opens new design possibilities. We have reported in this letter some preliminary results in designing such a system, showing that i) an ℓ_2-norm based formulation introduces the concept of watermark projection, enabling a privacy preserving feature, and ii) the feasibility of sparse domain data hiding is guaranteed with a loose condition $K < M$. A single support modification system is described as an unique example for sparse domain data hiding. Further research efforts could be devoted into more advanced designs based on the proposed framework including the aspects in trading off embedding distortion, robustness, capacity, and security, etc. In addition, it is also worthy to investigate the possibility of applying the proposed framework to reversible data hiding applications [15], [16]. Generally, it is important to note that the ℓ_2- and ℓ_0-norm solutions are only two special solutions for an underdertermined linear system. Thus, it is possible to look for other conditions with more specific properties in a data hiding system so that alternative solutions could be approached.

References

[1] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, “Secure spread spectrum watermarking for multimedia,” *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1673–1687, Dec. 1997.

[2] H. Malvar and D. Florencio, “Improved spread spectrum: A new modulation technique for robust watermarking,” *IEEE Trans. Signal Process.*, vol. 51, no. 4, pp. 898–905, Apr. 2003.

[3] B. Chen and G. W. Wornell, “Quantization index modulation: A class of provably good methods for digital watermarking and information embedding,” *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.

[4] M. Aharon, M. Elad, and A. Bruckstein, “K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation,” *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[5] M. Sadeghi, M. Babaie-Zadeh, and C. Jutten, “Learning overcomplete dictionaries based on atom-by-atom updating,” *IEEE Trans. Signal Process.*, vol. 62, no. 4, pp. 883–891, Feb. 2014.

[6] J. A. Tropp, “Greed is good: algorithmic results for sparse approximation,” *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[7] W. Dai and O. Milenkovic, “Subspace pursuit for compressive sensing signal reconstruction,” *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.

[8] S. J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, “An interior-point method for large-scale ℓ_1-regularized least squares,” *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 606–617, Dec. 2007.

[9] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, “A fast approach for overcomplete sparse decomposition based on smoothed ℓ^0 norm,” *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 289–301, Jan. 2009.

[10] T. T. Cai, L. Wang, and G. Xu, “New bounds for restricted isometry constants,” *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4388–4394, Sep. 2010.

[11] J. Lin, S. Li, and Y. Shen, “New bounds for restricted isometry constants with coherent tight frames,” *IEEE Trans. Signal Process.*, vol. 61, no. 3, pp. 611–621, Feb. 2013.

[12] R. Wu, W. Huang, and D. R. Chen, “The exact support recovery of sparse signals with noise via orthogonal matching pursuit,” *IEEE Signal Process. Lett.*, vol. 20, no. 4, pp. 403–406, Apr. 2013.

[13] E. J. Candes and T. Tao, “Decoding by linear programming,” *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[14] L.-H. Chang and J.-Y. Wu, “An improved rip-based performance guarantee for sparse signal recovery via orthogonal matching pursuit,” *IEEE Trans. Inf. Theory*, vol. 60, no. 9, pp. 5702–5715, Sep. 2014.

[15] L. An, X. Gao, X. Li, D. Tao, C. Deng, and J. Li, “Robust reversible watermarking via clustering and enhanced pixel-wise masking,” *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3598–3611, Aug. 2012.

[16] X. Gao, L. An, Y. Yuan, D. Tao, and X. Li, “Lossless data embedding using generalized statistical quantity histogram,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 8, pp. 1061–1070, Aug. 2011.