Fourteenth International Workshop on Algebraic and Combinatorial Coding Theory
September 7–13, 2014, Svetlogorsk (Kaliningrad region), Russia *pp. 181–185*

Double Sparse Compressed Sensing Problem[1]

Grigory Kabatiansky kaba@iitp.ru
Institute for Information Transmission Problems
Moscow, Russia
Cedric Tavernier ctavernier@assystem.com
Assystem, France

Abstract. We investigate a variant of the Compressed Sensing problem when the number of measurements corrupted by errors is upper bounded by some value l but l_2 norm of measurements' errors can be arbitrary large. We prove that in this case it is enough to make $2(t+l)$ measurements, where t is sparsity of original data. Moreover for this case a rather simple recovery algorithm is proposed. We derive an analog of Singleton bound from coding theory what proves optimality of the corresponding measurement matrices.

1 Introduction and Definitions

A vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ in n-dimensional vector space $\mathbb{R}^n$ called t-sparse if its Hamming weight $wt(x)$ or equivalently its l_0 norm $||x||_0$ is at most t, where by the definition $wt(x) = ||x||_0 = |\{i : x_i \neq 0\}|$. Let us recall that the Compressed Sensing (CS) Problem [1], [2] is a problem of reconstructing of an n-dimensional t-sparse vector x by a few (r) linear measurements $s_i = (h^{(i)}, x)$, assuming that measurements $(h^{(i)}, x)$ are known with some errors e_i, for $i = 1, \ldots, r$. Saying in other words, one needs to construct an $r \times n$ matrix H with minimal number of rows $h^{(1)}, \ldots, h^{(r)}$, such that the following equation

$$\hat{s} = Hx^T + e, \tag{1}$$

has either unique t-sparse solution or all such solutions are "almost equal". The compressed sensing problem was mainly investigated under assumption that the vector $e = (e_1, \ldots, e_r)$, called as an error vector, has relatively small Euclidean norm (length) $||e||_2$. We consider another problem's statement assuming that the error vector e is also sparse but its Euclidean norm can be arbitrary large. Saying in other words, we consider double sparse CS problem when $||x||_0 \leq t$ and $||e||_0 \leq l$. First time the assumption $||e||_0 \leq l$ was considered in [3] as a proper replacement for discrete version of CS-problem of usual assumption that

[1]The work of G. Kabatiansky was supported by RFBR grants 13-07-00978 and 13-01-12458.

an error vector e has relatively small Euclidean norm. Next definition is the corresponding modification of the definition given in [3].

Definition 1. *A real $r \times n$ matrix H called a (t, l)-compressed sensing (CS) matrix if*

$$||Hx^T - Hy^T||_0 \geq 2l + 1 \tag{2}$$

for any two distinct vectors $x, y \in \mathbb{R}^n$ such that $||x||_0 \leq t$ and $||y||_0 \leq t$.

This definition immediately leads (see [3]) to the following

Proposition 2. *A real $r \times n$ matrix H is a (t, l)-CS matrix iff*

$$||Hz^T||_0 \geq 2l + 1 \tag{3}$$

for any nonzero vector $z \in \mathbb{R}^n$ such that $||z||_0 \leq 2t$.

Our main result is an explicit and simple construction of (t, l)-CS matrices with $r = 2(t + l)$ for any n. We show this value of r is the minimal possible for (t, l)-CS matrices by proving an analog of well-known in coding theory Singleton bound for compressed sensing problem. Besides that we propose an efficient recovery (decoding) algorithm for the considered double sparse CS-problem.

2 Optimal Matrices for Double Sparse Compressed Sensing Problem

We start from constructing of (t, l)-CS matrices. Let a real $\tilde{r} \times n$ matrix $\tilde{H}$ be a parity-check matrix of an $(n, n - \tilde{r})$-code code over $\mathbb{R}$, correcting t errors, i.e. any $2t$ columns $\tilde{h}_{i_1}, \ldots, \tilde{h}_{i_{2t}}$ of $\tilde{H}$ are linear independent. And let G be a generator matrix of an $(r, \tilde{r})$-code over $\mathbb{R}$ of length r, correcting l errors. Let matrix H consists of columns $h_1, \ldots, h_n$, where

$$h_j^T = \tilde{h}_j^T G \tag{4}$$

and transposition ^{T} means, that vectors h_j and $\tilde{h}_j$ are considered in (4) as row vectors, i.e.

$$H = G^T \tilde{H} \tag{5}$$

Saying in words, we encode columns of parity-check matrix $\tilde{H}$, which already capable to correct t errors, by a code, correcting l errors, in order to restore correctly syndrom of $\tilde{H}$.

Theorem 3. *Matrix $H = G^T \tilde{H}$ is a (t, l)-CS matrix.*

Proof. According to Proposition 2 it is enough to prove that $||Hz^T||_0 \geq 2l+1$ for any nonzero vector $z \in \mathbb{R}^n$ such that $||z||_0 \leq 2t$. Indeed, $u = \tilde{H}z^T \neq 0$ since any $2t$ columns of $\tilde{H}$ are linear independent. Then $Hz^T = G^T\tilde{H}z^T = G^Tu = (u^TG)^T$ and u^TG is a nonzero vector of a code over $\mathbb{R}$, correcting l errors. Hence $||Hz^T||_0 = ||u^TG||_0 \geq 2l+1$. □

Now let us choose well known Reed-Solomon (RS) codes (which are a particular case of evaluation codes construction) as both constituent codes. The length of RS-code is restricted by the number of elements in the field so in the case of $\mathbb{R}$ the length of evaluation code can be arbitrary large. Indeed, consider the corresponding evaluation code $\mathbb{RS}_{(n,k)} = \{(f(a_1,\dots,f(a_n)) : \deg f(x) < k\}$, where $a_1,\dots,a_n \in \mathbb{R}$ are n different real numbers. The distance of $\mathbb{RS}_{(n,k)}$ code $d = n-k+1$ since the number of roots of a polynomial cannot exceed its degree and hence $d \geq n-k+1$,but, on the other hand, the Singleton bound states that $d \leq n-k+1$ for any code, see [4]. Therefore the resulting matrix H is a (t,l)-CS matrix with $r = 2(t+l)$. The next result, which is a generalization of the Singleton bound for double sparse CS problem, shows these matrices are optimal in the sense having the minimal possible number r of linear measurements.

Theorem 4. *For any (t,l)-CS $r \times n$-matrix*

$$r \geq 2(t+l). \tag{6}$$

Proof. Let H be a (t,l)-CS matrix of size $r \times n$, i.e., $||Hz^T||_0 \geq 2l+1$ for any nonzero vector $z \in \mathbb{R}^n : ||z||_0 \leq 2t$. And let H_{2t-1} be $(2t-1) \times n$ matrix consisting of first $2t-1$ rows of H. There exists a nonzero vector $\hat{z} = (\hat{z}_1,\dots,\hat{z}_{2t},0,0,\dots,0) \in \mathbb{R}^n$ such that $H\hat{z}^T = 0$ (a system of linear homogenious equations with the number of unknown variables larger than the number of equations has a nontrivial solution). Then $||H\hat{z}^T||_0 \leq r-(2t-1)$ and finally $r \geq 2t+2l$ since $||H\hat{z}^T||_0 \geq 2l+1$. □

3 Recovery Algorithm for Double Sparse Compressed Sensing Problem

Let us start from a simple remark that for $e = 0$ recovering of the original sparse vector x, i.e., solving the equation (1), is the same as syndrom decoding of some code (over $\mathbb{R}$) defined by matrix H as a parity-check matrix. In general, syndrom $s = Hx^T$ is known with some error, namely, as $\hat{s} = s+e$ and therefore we additionally encoded columns of H by some error-correcting code in order to recover the original syndrom s and then apply usual syndrom decoding algorithm. Therefore recovering, i.e., decoding algorithm for constructed in previous chapter optimal matrices is in some sense a “concatenation” of decoding algorithms of constituent codes.

Namely, first we decode vector $\hat{s} = s + e$ by a decoding algorithm of the code with generator matrix G. Since $||e||_0 \leq l$ this algorithm outputs the correct syndrome s. After that we form a syndrome $\tilde{s}$ by selecting first $\tilde{r}$ coordinates of s and then apply syndrom decoding algorithm (of the first code with parity-check matrix $\tilde{H}$) for the following syndrom equation

$$\tilde{s} = \tilde{H}x^T. \tag{7}$$

Now let us discuss a right choice of constituent codes. Surely, we stay in the class of Reed-Solomon codes over $\mathbb{R}$. There are well known algorithms of their decoding up to half of the code distance (bounded distance decoding, see [4]), for instance, Berlekamp-Massey algorithm, which in our case (codes over $\mathbb{R}$) is known also as Trench algorithm, see [5], [6]. Hence the total decoding complexity doesn't exceed $O((t + l)n)$ or not more than $O(n^2)$ operations over real numbers. Moreover we can even decode these codes over their half distances by application of Guruswami-Sudan list decoding algorithm [7].

We know from coding theory that encoding-decoding procedures of Reed-Solomon codes become more simple in the case of cyclic codes, when the set $a_1, \ldots, a_n$ is a cyclic group under multiplication. In order to do it let us consider $a_1, \ldots, a_n$ as complex roots of degree n and define our codes through their "roots", i.e. our codes consists of *polynomials* $f(x)$ over $\mathbb{R}$ such that $f(e^{2\pi i \frac{m}{n}}) = 0$ for $m \in \{-s, \ldots, -1, 0, +1, \ldots, +s\}$ with $s = t$ for first consituent code and $s = l$ for second. It easy to check that such codes achieve the Singleton bound with $d = 2s + 2$, so the corresponding double sparse code has redudancy $r = 2(t + l + 1)$ what is slightly larger than the corresponding Singleton bound, but in return these codes can be decoded via FFT.

4 Conclusion

It is a very natural question why do not use just a single RS-code with extra redudancy in order to correct possible errors in measurements, i.e., in its syndrom. In fact, it is rather old question, which goes back to time of the French Revolution, when R.Prony [8] asked how to reconstruct a polynomial of a given degree by its value in some points, when at most l of these values could be incorrect. The modern solution was given in [9], namely, it was shown that it is possible to solve equation (1) by RS-code *iff* its redudancy $r \geq 2t(2l + 1)$. We see that it is too much expensive solution for double sparse CS-problem.
Let us note that first papers on CS-problem gave a special attention to the case of $e = 0$ in order to say that RIP and other Compressed Sensing techniques can recover a sparse vector even if this vector can be arbitrary large. We show how to recover a sparse (but arbitrary large) vector x by the minimal number of linear measurements if errors in measurements are sparse also (but can be component-wise arbitrary large).

References

[1] D.L. Donoho, “Compressed sensing”, IEEE Transactions on Information Theory, v. 52 (4), pp. 1289-1306, 2006

[2] E.J. Candes, T. Tao, “Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies? ”, IEEE Transactions on Information Theory, v. 52 (12), pp. 5406 - 5425, 2006

[3] G.Kabatiansky, S.Vladuts, “What to do if syndromes are corrupte also”, in Proc. Int.Workshop Optimal Codes, Albena, Bulgaria, 2013.

[4] F.J. MacWilliams, N.J.A. Sloane, The Theory of Error-Correcting Codes (North-Holland Mathematical Library)

[5] W.R.Trench, “An algorithm for the inversion of finite Toeplitz matrices”, J. Soc. Indust.Appl. Math.,v.12, 1964.

[6] D.Knuth, The Art of Programming.

[7] V. Guruswami and M. Sudan, “Improved decoding of Reed-Solomon and algebraic-geometry codes”, IEEE Trans. on Information Theory, vol. 45, pp. 1757–1767, 1999.

[8] R.Prony, “Essai experimantal et analytique sur les lois del Dilatabilite de fluides” J. de Ecole Polytechnique 1, pp.24-76, 1795.

[9] M.T. Comer, E.L. Kaltofen, C.Pernet, “Sparse Polynomial Interpolationb and Berlekamp-Massey Algorithms That Correct Outlier Errors in Input Values”, in Proc. ISSAC 2011, pp. 138-145.