

Hybrid Compressed Sensing of Images

Abdolreza Abdolhosseini Moghadam and Hayder Radha

Dep. of Electrical and Computer Engineering, Michigan State University, Email: {abdolhos,radha}@msu.edu

Abstract—We consider the problem of recovering a signal/image (x) with a k -sparse representation, from hybrid (complex and real), noiseless linear samples (y) using a mixture of complex-valued sparse and real-valued dense projections within a single matrix. The proposed Hybrid Compressed Sensing (HCS) employs the complex-sparse part of the projection matrix to divide the n -dimensional signal (x) into subsets. In turn, each subset of the signal (coefficients) is mapped onto a complex sample of the measurement vector (y). Under a worst-case scenario of such sparsity-induced mapping, when the number of *complex sparse* measurements is sufficiently large then this mapping leads to the isolation of a significant fraction of the k non-zero coefficients into different complex measurement samples from y . Using a simple property of complex numbers (namely complex phases) one can identify the isolated non-zeros of x . After reducing the effect of the identified non-zero coefficients from the compressive samples, we utilize the real-valued dense submatrix to form a full rank system of equations to recover the signal values in the remaining indices (that are not recovered by the sparse complex projection part). We show that the proposed hybrid approach can recover a k -sparse signal (with high probability) while requiring only $m \approx 3k \sqrt[3]{n/2k}$ real measurements (where each complex sample is counted as two real measurements). We also derive expressions for the optimal mix of complex-sparse and real-dense rows within an HCS projection matrix. Further, in a practical range of sparsity ratio (k/n) suitable for images, the hybrid approach outperforms even the most complex compressed sensing frameworks (namely basis pursuit with dense Gaussian matrices). The theoretical complexity of HCS is less than the complexity of solving a full-rank system of m linear equations. In practice, the complexity can be lower than this bound.

Index Terms—Compressed sensing, sparse projections, iterative decoding algorithms

I. INTRODUCTION

Over the past few years, Compressed Sensing [1], [2] has been introduced as an alternative, efficient method for sampling signals which have sparse representations with respect to some bases. Consequently, some compressive sampling schemes for sensing images [12], [13] have been proposed, exploiting the fact that images have sparse or compressible representations in some known bases (such as DCT, Contourlets and so on [14]-[16]). Consider a n -dimensional, real-valued signal (e.g. an n pixel image) x which has a k -sparse representation θ in the transform domain ψ , that is:

$$x \in \mathbb{R}^n, x = \psi\theta, k = \|\theta\|_0 := \{i : \theta_i \neq 0\} \quad (1)$$

The main idea of Compressed Sensing (CS) [1], [2] is that instead of sensing n samples x_i , the encoder projects x onto m compressive samples y according to a projection matrix P , i.e. $y = Px$, where $k < m < n$. Accordingly, the decoder recovers the sparse representation θ by utilizing the compressive samples y and the *effective* projection matrix $P\psi$

and thereafter reconstructs x . Generally the under-determined system of equations $y = Px = P\psi\theta$ has an infinite set of solutions. However, if certain properties outlined in [1], [2] are met then it is possible to find θ (and consequently $x = \psi\theta$) uniquely using the following optimization problem:

$$\theta = \arg \min \|\hat{\theta}\|_0 = \arg \min \|\hat{\theta}\|_1 : y = P\psi\hat{\theta} \quad (2)$$

Under early CS solutions, the complexity of the solver has an inverse relationship to the number of compressive samples required. For instance, greedy algorithms [3]-[5] have low complexities but require relatively large numbers of samples; on the other hand, convex relaxation (ℓ_1 optimization) algorithms [6] recover the signal (or transform coefficients of an image) with fewer samples but they are extremely complex ($\Omega(n^3)$) relative to greedy algorithms. Furthermore, in case of images, these CS methods suffer from what is known as *phase transition* [19]. In words, early CS solutions recover a sparse/compressible signal when the sparsity ratio (k/n) is in some specific range. Outside this range, the probability of the signal recovery (under that specific solver) drops significantly. Typically for early CS solutions, this ratio is small. Although, images may have sparse/compressible representations in some bases, in practice, the sparsity ratio of an image is (usually) outside the optimal range of these CS solutions. Hence, for images, a CS framework capable of operating under moderate sparsity is desirable. This led to many efforts with promising results that demonstrated improvements over traditional CS approaches [3], [7]-[11].

More recently, there has been some interest in pursuing sparse projections that lead to low complexity and low measurement requirements [8]-[11]. Overall, state of the art sparse approaches have shown to achieve recovery (with high probability) using $m = O(k \log(n/k))$ measurements and with time complexity $O(n \log(n/k))$ [10]. Despite the clear benefits of sparse projections, one needs to use rather strict structures (e.g., expander graphs) to achieve these bounds simultaneously [7]-[10]. Even then, low values for the constant (c) of the measurement bound $m = O(k \log(n/k)) = ck \log(n/k)$ cannot be guaranteed. In this context, using traditional dense Gaussian matrices requires minimal measurements, however, at the expense of a high complexity ℓ_1 minimization decoding algorithm.

In this paper, we present a Hybrid Compressed Sensing (HCS) framework, which employs two classes of (sparse and dense) sensing projections to solve the problem of compressive sampling in a novel way. More specifically, the HCS projection matrix P is composed of two sub-matrices: a sparse, complex-valued sub-matrix $P^{(c)}$ and a random dense, real-

valued sub-matrix $P^{(r)}$. As demonstrated below, the sparse complex-valued projection provides a natural mechanism for lowering the solver complexity. Meanwhile, the dense real-valued projection can facilitate the lowering of the sampling requirement for the recovery of the signal. Hence, in the context of the proposed hybrid sparse-dense framework, we approach compressed sensing as an optimization problem by identifying the best mix of sparse-dense random projections that minimize the total number of samples required for the recovery of a k -sparse signal x .

The sparse complex-valued sub-matrix $P^{(c)}$ of HCS can employ any underlying structure for its non-zero entries, including structures that are based on combinatorial designs or expander graphs [8]-[10]. Here, we focus on a worst case scenario (in terms of required measurements) by employing a simple structure that divides the signal coefficients into a number of non-overlapping sub-signals (i.e. *partitions*). In Section II, we show, that when the number of sub-signals (or equivalently the number of rows of $P^{(c)}$) is sufficiently large, then this partition (of coefficients) leads to the isolation of most of the k -non-zero transform coefficients into different compressive samples. Hence, one can solve for the non-zero coefficients of x in a “divide-and-conquer” strategy. After recovering these isolated non-zero transform coefficients (utilizing a simple property of complex numbers), one might use basic linear algebra methods to recover the remaining transform coefficients utilizing the random dense sub-matrix $P^{(r)}$. In Section III, we show that (with high probability) HCS requires $m \approx (3 + \delta)k \sqrt[3]{n/k}$ compressive samples to recover a k -sparse signal of length n where $\delta \ll 1$. Although HCS is not optimal (in terms of sample requirements for the perfect recovery), it can reconstruct a signal from fewer samples compared to the most complex CS framework (i.e., basis pursuit with dense Gaussian matrices) in a practical range of the sparsity ratio suitable for images. Moreover the complexity of HCS is low due to the solver simplicity. Specifically, the complexity of the HCS solver (in the worst case) is less than the complexity of solving a (full-rank) system of m linear equations. Section IV presents our simulation results and Section V concludes this paper.

II. HYBRID SPARSE-DENSE PROJECTIONS

For the simplicity of notations and without loss of generality, throughout this paper, we assume that the signal/image x is sparse itself, that is the sparsifying domain ψ in (1) is the identity matrix (hence $\theta = x$). Similarly, it would be beneficial if consider the vectorized form of an image. By vectorized form, we mean that we stack the columns of the image on top of each others, that is:

$$l = (j - 1)n_1 + i : x_l := x_{i,j} \quad (3)$$

where $x_{i,j}$ is the pixel located at the Cartesian location (i, j) . A sparse-projection based method naturally projects a subset x_{M_i} of the signal x ($x_{M_i} \subset \{x_1, \dots, x_n\}$) into an observations y_i , where $y_i \in \{y_1, \dots, y_m\}$ and M_i is the

set of column indices for the non-zero entries of the i^{th} row of a sparse projection matrix P . For example, if the i^{th} row of a sparse projection matrix P has zero entries everywhere except for non-zero entries in some locations (column indices), say $M_i = \{3, 8, 31\}$, then we have a corresponding subset $x_{M_i} = \{x_3, x_8, x_{31}\}$. Consequently, the i^{th} compressive sample y_i is a function of the subset x_{M_i} ($y_i = g_i(x_{M_i})$), and y_i does not depend on the values of other elements of the signal¹ x . Overall, the natural division of the signal x (into some subsets) by a sparse projection leads to low-complexity recovery algorithms since one can begin by recovering one or few subsets of x from a corresponding subset of observations, and then progressively recover more subsets of x . Due to their attributes, sparse projections have been receiving more attention recently [7]-[11]. One of the key contributions of this work is extending the utility of sparse projections by 1) using complex-valued entries with simple structures instead of binary-valued projections with rather complicated and restrictive structures and 2) complimenting them with important aspects of dense sensing as explained further below. More importantly, as opposed to most existing CS solutions, which rely on solely sparse or merely dense matrices, our proposed method employs the features of these two classes of matrices simultaneously to design a simple yet efficient solution for CS.

The following notation will be used in this paper. Let A be a $b \times c$ matrix and let $i \subseteq [b] = \{1, \dots, b\}$ and $j \subseteq [c]$. A^T represents the transpose of A . By A_b , and $A_{\cdot c}$ we respectively mean submatrices of A given by restricting A on rows and columns indexed by those members of b and c . Moreover let D be a $d \times 1$ vector and let $e \subseteq [d]$. Then \hat{D}_e is a $d \times 1$ vector given by keeping the entries of D indexed by e and setting the rest (indices of $[d] \setminus e$) to zero.

A. HCS Projection Matrices

The projection matrix (P) of HCS for a length n signal is composed of two sub-matrices ($P^{(c)}$) and ($P^{(r)}$): $P^T = [(P^{(c)})^T \ (P^{(r)})^T]$, where $P^{(r)}$ is a $m_r \times n$ real-valued, dense random matrix and $P^{(c)}$ is a $m_c \times n$ sparse matrix with random complex entries (in section III, we find the optimal values for m_c and m_r under an extreme case to be explained). If we count each complex valued compressive sample as two real samples, then the projection matrix is sensing $2m_c + m_r$ measurements. For simplicity of analysis and without loss of generality assume m_c counts n . Now let $M = \{M_1, \dots, M_{m_c}\}$ be a set of equal size, random subsets over $[n] = 1, 2, \dots, n$ such that M forms a *partition*. Thus: $\forall i, j \in [m_c], i \neq j : |M_i| = w, M_i \cap M_j = \emptyset, \cup_{i=1}^{m_c} M_i = [n]$. Then the i^{th} row of $P^{(c)}$ is non-zero, only in the column indices of M_i . Also let $f^{(i)}$ be a $1 \times w$ row vector with random complex entries having distinct phases in the range of $[0, \pi)$. Then we insert the entries of $f^{(i)}$ in the i^{th} row of $P^{(c)}$ in the column indices of M_i . This can be written as: $\forall i \in [m_c] : P_{i, M_i}^{(c)} = f^{(i)}$ and $P_{i, [n] \setminus M_i}^{(c)} = 0$. Thus the i^{th}

¹We say that the i^{th} measurement sample spans the indices of M_i .

sample (where $i \in [m_c]$) is a function of the signal coefficients only in the indices of M_i : $y_i = f^{(i)}x_{M_i}$. Appending m_r real-valued, dense random rows to this matrix yields the final projection matrix.

B. HCS Solver

We first highlight an important (yet arguably clear) proposition, which is critical for the successful working of the HCS solver: if a subset x_{M_i} of the coefficients of x has at most one non-zero coefficient, then only one complex-valued compressive sample y_i is sufficient for the recovery of that subset.

Proposition 1. *Let $M_i \subset [n], |M_i| = w$ be the (ordered) set of indices of non-zero coefficients for the i^{th} row of a sparse matrix $P^{(c)}$. And let $P_{i,M_i}^{(c)} = f^{(i)} = [e^{j\phi_1^{(i)}} \dots e^{j\phi_w^{(i)}}]$ be the corresponding vector of complex exponentials (on the unit circle) such that the random phases of $f^{(i)}$ are distinct in the range of $[0, \pi)$, $\forall l, t \in [w], l \neq t : \phi_l^{(i)} \in [0, \pi), \phi_l^{(i)} \neq \phi_t^{(i)}$. If x_{M_i} is real-valued and non-zero in at most one index ($\|x_{M_i}\|_0 \leq 1$) then having $y_i = P_{i,M_i}^{(c)}x = f^{(i)}x_{M_i}$ and $f^{(i)}$, one can recover x_{M_i} deterministically. Moreover having y_i and $f^{(i)}$, we can detect the event that x_{M_i} is non-zero in at least two indices. Specifically if $y_i = 0$ then $x_{M_i} = \mathbf{0}$. If the phase of y_i equals to the phase of l^{th} entry of $f^{(i)}$ ($\angle y_i = \angle f_l^{(i)} = \phi_l^{(i)}$) then x_{M_i} is non-zero only in the index of $M_{i,l}$ and has the value of $x_{M_{i,l}} = y_i/f_l^{(i)} = \|y_i\|$. If the phase of y_i is not among the phases of $f^{(i)}$ ($\{\phi_1^{(i)}, \dots, \phi_w^{(i)}\}$), then x_{M_i} is non-zero in at least two indices. The proof is presented in the Appendix.*

The HCS solver consists of two major stages: 1) finding and recovering subsets of the signal coefficients spanning at most one non-zero coefficient and 2) forming a full-rank system of linear equations and solving it. For the first m_c complex valued compressive samples ($y_i, i \in [m_c]$), the solver looks for the phase of each compressive sample ($\angle y_i$) among the available phases in the associated random vector. As before let $f^{(i)} = P_{i,M_i}^{(c)}$, $w = |M_i|$ and $\angle f^{(i)} = [\angle f_1^{(i)} \dots \angle f_w^{(i)}]$. Assume the phase of sample (y_i) equals to the phase of t^{th} entry of $f^{(i)}$. Then by Proposition 1, x_{M_i} is zero in all indices except in the index of $M_{i,t}$. This can be written as:

$$\angle y_i = \angle f_t^{(i)} \Leftrightarrow x_l = \begin{cases} \|y_i\| & l = M_{i,t} \\ 0 & l \in M_i \setminus t \end{cases} \quad (4)$$

Moreover, if the compressive sample y_i is zero, then x_{M_i} is a zero vector. Otherwise (if the sample is neither zero nor its phase is found in $\angle f^{(i)}$), this subset of the coefficients (x_{M_i}) spans at least two non-zeros and Proposition 1 is not applicable to this subset. Define I' as the set of sample indices, such that each of these samples is either zero or the phase of that sample can be found among the available phases in the respective random vector:

$$I' = \left\{ i \in [m_c] : y_i = 0 \text{ or } \angle y_i \in \angle f^{(i)} \right\} \quad (5)$$

Thus the first stage of the HCS solver determines signal values on indices of $I = \cup_{i \in I'} M_i$. Consequently $P\hat{x}_I$ is the effect

of identified coefficients on the compressive samples y ; and by subtracting $P\hat{x}_I$ from the compressive samples y we are nullifying such effect: $y - P\hat{x}_I = P\hat{x}_{I^c}$ where I^c is the set complement of I ($I \cup I^c = [n]$). If we consider the n signal coefficients as unknowns and compressive samples as equations, the problem of $y = Px$ can be viewed as a system of n unknowns and $m = 2m_c + m_r$ equations where $|I|$ coefficients have already been recovered (in the first stage of the solver). Having $n - |I|$ further (independent) equations one can recover the remaining unknowns. The random dense submatrix ($P^{(r)}$) and complex samples spanning at least two non-zeros provide us such equations.

III. ANALYSIS OF THE HCS SOLVER

In this section, we derive the sample requirements of HCS (under perfect recovery) and also the complexity of such procedure. To that end, we first find the average number of samples required for the perfect recovery and then using the tools from the concentration of measurement [17] we show that our derived bound is tight.

Recall that complex compressive samples span disjoint subset of coefficients. Hence it is straightforward to see that the distribution of number of non-zeros spanned by each sample is multinomial. The problem of finding this multinomial distribution can be casted as the classical ‘‘balls into bins’’ problem [17]. Hence probability of the event that a complex compressive sample (say $y_i, i \in [m_c]$) spans j non-zero(s) is:

$$\Pr(\|x_{M_i}\|_0 = j) = \binom{k}{j} \left(1 - \frac{1}{m_c}\right)^{k-j} \left(\frac{1}{m_c}\right)^j \quad (6)$$

Consequently, the expected number of samples spanning at most one non-zero coefficient ($|I'|$) for large values of k is:

$$E(|I'|) = m_c \Pr(\|x_{M_i}\|_0 \leq 1) \approx m_c \left(1 + \frac{k}{m_c}\right) e^{-\frac{k}{m_c}} \quad (7)$$

Since each complex sample spans $w = n/m_c$ coefficients, the number of identified coefficients (I) on average is $E(|I|) = wE(|I'|) = n(1 + k/m_c) \exp(-k/m_c)$. Note that we have not utilized $m_c - I'$ complex samples in the first stage of the algorithm (since these samples span at least two non-zeros). Counting each complex sample as two equations, we need at least $m_r \geq n - E(|I|) - 2(m_c - I')$ further equations to form a full rank system of equations to recover the remaining coefficients. To identify the required number of measurements, this problem can be casted as the following optimization problem:

$$\arg \min_{m_c, m_r} 2m_c + m_r \text{ s.t. } m_r \geq n - E(|I|) - 2(m_c - I') \quad (8)$$

The following two lemmas outline key results for solving this problem and determine the sample requirements of the HCS solver. The first lemma provides expressions for the numbers of complex samples (m_c) and real samples (m_r) that are required for the recovery of x under HCS. The second lemma simplifies these expressions and show that the HCS measurement bound is tight. A corollary is also stated below. (See the appendix for a sketch of the proof.)

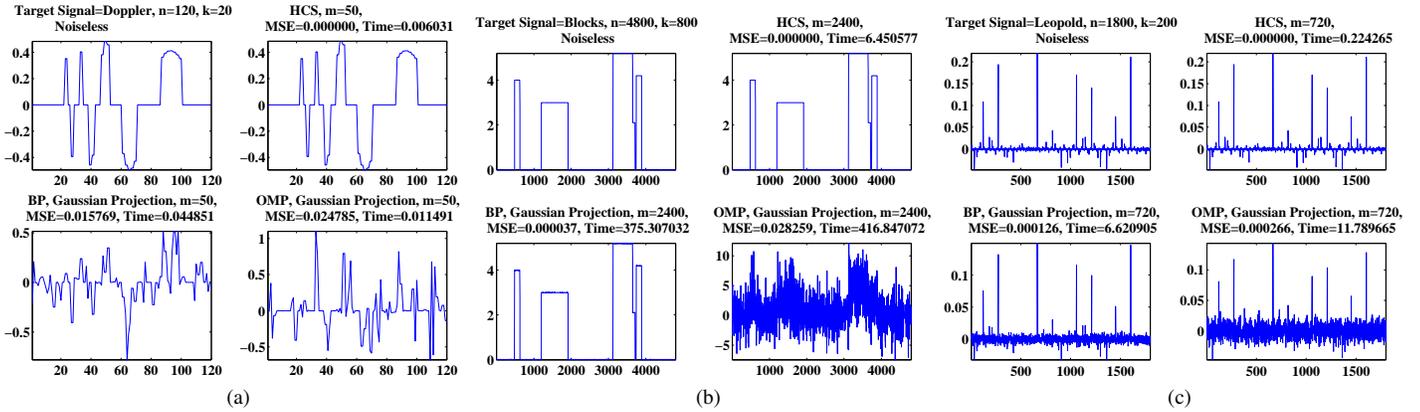


Fig. 1: m , n and k represent number of samples, signal length and signal sparsity respectively.

Lemma 1. Consider a k -sparse signal x of length n and define $l = n/k$. Then on average, HCS requires $m_c \approx k \left(\sqrt[3]{\frac{l}{2}} - \sqrt[3]{\frac{16}{729l}} - \frac{1}{3} \right)$ complex-valued samples and $m_r = n - n\alpha - 2\alpha m_c$ real-valued samples for the perfect recovery of x where $\alpha = (1 + k/m_c) \exp(-k/m_c)$.

Lemma 2. If $n(2(\alpha - \alpha^2)^2 \ln k)^3 \ll k^4$ then with a probability of at least $1 - \frac{2}{k}$, HCS requires $m = 2m_c + m_r$ compressive samples for the perfect recovery of x where $m_c = k(1 + \delta) \sqrt[3]{\frac{n}{2k}}$, $m_r = n - n\alpha - 2\alpha m_c$ and $0 < \delta \ll 1$.

Corollary 1. If $\sqrt[3]{2l^2} \leq 2 + 1/(1 - \alpha)$ then $m_r \leq m_c$.

Although the HCS sampling bound is not optimal for all values of k and n , HCS outperforms even the most complex CS solutions over a wide range of typical and practical values of k and n (e.g., as long as k/n is not excessively small). Specifically let $m_{opt} = Ck \log n/k$ be the sample requirement for an optimal CS framework where C is a constant only depending on the projection matrix and the solver. If $\sqrt[3]{n/k} < 1.25C \log(n/k)$, then HCS requires fewer samples for perfect recovery when compared with the sampling requirement of the hypothetical optimal solver. Now let us compute the complexity of the HCS solver. Recall that the HCS decoding process consists of two stages. In the first stage, for each complex valued compressive sample ($y_i, i \in [m_c]$), we look for the phase of that sample ($\angle y_i$) among the available phases in the respective random vector $\angle f^{(i)}$. Assume the entries of $f^{(i)} = [\exp(j\phi_1^{(i)}) \exp(j\phi_2^{(i)}) \dots \exp(j\phi_w^{(i)})]$ are sorted based on phases: $a, b \in [w], a > b \Leftrightarrow \angle f_a^{(i)} = \phi_a^{(i)} > \angle f_b^{(i)} = \phi_b^{(i)}$. Then the complexity of this search for each compressive sample is only $\log(w)$. Since we have to repeat this search for all complex valued compressive samples, thus the complexity of the first stage of the solver is $O(m_c \log(w))$.

As before, let $|I'| \geq 0$, denotes the number of compressive samples spanning at most one non-zero coefficient, then we have not utilized $m_c - |I'|$ complex sample from the first stage of the solver. The second stage of the HCS solver finds the solution of a full rank system of $m_r + m_c - |I'| \leq 2m_c$ equa-

tions. Let $g(\beta)$ be the complexity of linear solver employed in the second stage of HCS solver to solve a full rank system of β equations. Then the complexity of the second stage of HCS is $O(g(m_r + m_c - |I'|)) \leq O(g(2m_c))$. Finally the complexity of HCS solver is $O(\max\{m_c \log(n/m_c), g(2m_c)\})$.

IV. SIMULATION RESULTS

We tested the proposed HCS framework on a large number of standard compressed sensing signals and images. Here, we present the results for images of Monalisa and Lenna and compare our results with two dominant solvers/projection matrices (Fig. 2 and Fig. 3). We performed compressive sampling on the whole image of Monalisa, while for Lenna, we applied a block based compressed sensing. More specifically for Lenna, the target image is formed by keeping the largest $k = 8$ DCT coefficients in all 8×8 blocks and set the rest of DCT coefficients to zero. For Monalisa image, a total of $k = 57979$ DCT coefficients (with the largest absolute value) out of $n = 283554$ coefficients have been kept and we set the remaining coefficients to zero. In our simulations, we counted each complex valued compressive sample as two real samples. In these figures m , t_{bp} , t_{hcs} and t_{omp} represent the number of (real) compressive samples and the decoding time required for BP, HCS and OMP respectively. In all simulations, we have assumed that for a given number of samples m , the ratio of complex valued compressive samples and real valued dense samples (for HCS) is approximately one: ($m_c = \lceil m/3 \rceil$ and $m_r = m - 2m_c$). It is important to note that this sample assignment is not optimal and HCS needs fewer samples for the perfect recovery of the signal (see Lemma 1). However in some real world application, the number of non-zero coefficients might not be known beforehand. Although BP achieved virtually perfect reconstruction, but clearly the complexity of BP is much higher compared to HCS. Further simulation results in case of 1D signals are presented in Fig. 1.

V. CONCLUSION

In this paper, we considered the problem of recovering a k -sparse signal x from a limited number of linear samples

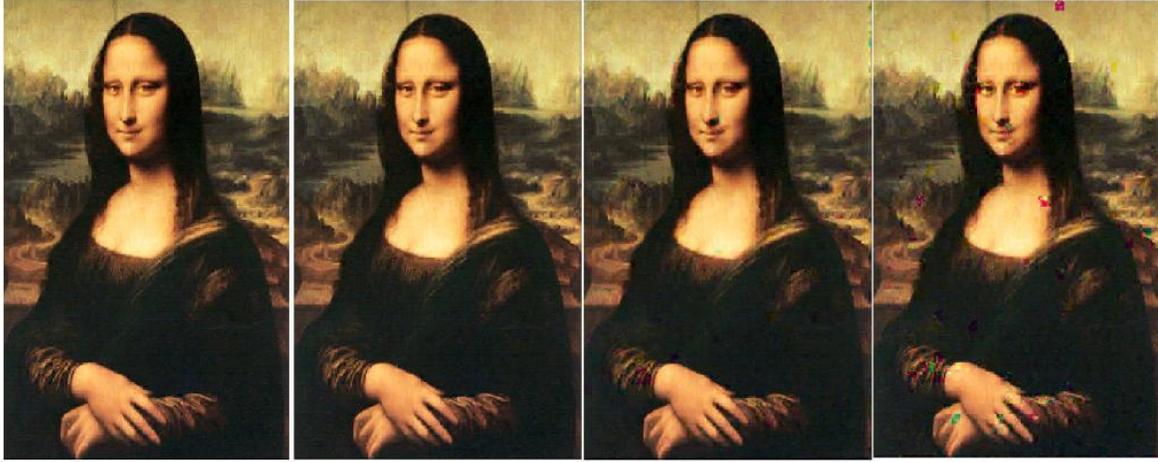


Fig. 2: From left to right, the k -sparse image of Monalisa, reconstructed by applying HCS, BP and OMP respectively to the whole image ($m = 150746$). The decoding times are $t_{bp} = 109.5077$, $t_{hcs} = 16.1704$ and $t_{omp} = 15.1670$ seconds.



Fig. 3: From left to right, the k -sparse image of Lenna, reconstructed with $m = 28$ compressive samples by applying HCS, BP and OMP respectively on blocks of size 8×8 pixel. The decoding times are $t_{bp} = 41.0283$, $t_{hcs} = 4.4735$ and $t_{omp} = 1.1512$ seconds.

($y = Px$) and proposed Hybrid Compressed Sensing (HCS) to solve this problem. We have shown that HCS recovers x from $m \approx (3 + \delta)k \sqrt[3]{n/2k}$ real measurements where $\delta \ll 1$. Although this bound is not optimal when compared to the well-known bound of $O(k \log(n/k))$, in a practical range of sparsity ratio, it outperforms the most complex CS approach that utilizes dense Gaussian projections and which is known to require the minimal number of measurements. To conclude this paper, we should highlight that the simulation results presented for HCS can be improved (arguably significantly). First, in all simulations, a non-iterative (full rank) linear equation solver has been deployed in the second phase of HCS. We can reduce the complexity of HCS furthermore by utilizing iterative linear equation solver (such as Jacobi method, Successive Over Relaxation method, etc. [18]) in the second stage of HCS. Moreover, the performance of HCS can be improved further by allowing overlaps and more restrictive structures in the sparse complex-valued part $P^{(c)}$ of P . As stated earlier, in this paper, we focused on the worst-case (partition) scenario for $P^{(c)}$.

VI. APPENDIX

Proof of Proposition 1 Clearly if x_{M_i} is a zero vector then $y_i = 0$. Now assume $y_i = 0$. We show that with high probability $x_{M_i} = \mathbf{0}$. Note that if x_{M_i} is non-zero only in index of $M_{i,l}$ ($\|x_{M_i}\|_0 = 1$ and $x_{M_{i,l}} \neq 0$) then $y_i = x_{M_{i,l}} e^{j\phi_l^{(i)}} \neq 0$. Now suppose x_{M_i} is non-zero in at least two indices ($\|x_{M_{i,l}}\|_0 \geq 2$). Hence y_i is a linear combination of at least two independent complex numbers. Suppose the values of all entries of x_l are bounded by η . Then the magnitude of y_i is a continuous random variable (at least in the range of $[0, \eta]$). Thus the probability of the event that the magnitude of y_i takes *exactly* the value of zero, is zero. Hence with high probability: $x_{M_i} = \mathbf{0} \Leftrightarrow y_i = 0$ and therefore $x_{M_i} = \mathbf{0} \Leftrightarrow y_i = 0$. Now assume x_{M_i} is non-zero only in the index of $M_{i,l}$. Hence: $y_i = x_{M_{i,l}} e^{j\phi_l^{(i)}}$. Note that in this case, the phase of y_i (i.e. $\angle y_i = \phi_l^{(i)}$) can be found in the ordered set of available phases in $f^{(i)}$ ($\angle f^{(i)} = \{\phi_1^{(i)}, \dots, \phi_w^{(i)}\}$). In other words, knowledge of the phase $\angle y_i = \phi_l^{(i)}$ translates into knowledge of the index l associated with the location of the (single) non-zero coefficient $x_{M_{i,l}}$. Now we prove (by

contradiction) that if the phase of y_i can be found in $\angle f^{(i)}$ then with high probability x_{M_i} is non-zero exactly in one index. Suppose x_{M_i} is non-zero in at least two indices and $\angle y_i$ can be found among $\angle f^{(i)}$. By the problem setup, $\angle y_i$ is a continuous random variable in the range of $[0, \pi)$. However, we know that the probability of the event that a continuous random variable ($\angle y_i$) attains the values of a finite set of numbers ($\angle f^{(i)}$), is zero and this is a contradiction. Thus:

$$\angle y_i = \angle f_l^{(i)} \Leftrightarrow x_u = \begin{cases} \|y_i\| & u = M_{i,l} \\ 0 & u \in M_i \setminus M_{i,l} \end{cases} \quad (9)$$

Finally if $y_i \neq 0$ and $\angle y_i$ is not found in the ordered set $\angle f^{(i)}$, then x_{M_i} is non-zero at least in two indices. ■

Proof of Lemma 1.

To that end we solve the following optimization problem:

$$\arg \min_{m_c, m_r} 2m_c + m_r \text{ s.t. } m_r \geq n - E(|I|) - 2(m_c - I') \quad (10)$$

One can use the Lagrangian multiplier to solve (10). Define $\alpha = E(|I'|)/m_c < 1$, then the objective function (L) is:

$$L = 2m_c + m_r + \lambda(\alpha n + 2(1 - \alpha)m_c + m_r - n) \quad (11)$$

Taking derivatives respect to three parameters λ , m_c and m_r gives us three equations:

$$\lambda = -1, \quad \alpha n + 2(1 - \alpha)m_c + m_r = n \quad (12)$$

$$\frac{\partial L}{\partial m_c} = 0 \Rightarrow 2m_c^3 + 2km_c^2 + 2k^2m_c = nk^2 \quad (13)$$

Define $l = n/k$. Let us assume $l^2 \gg l$, then solving (13) yields:

$$m_c = k \left(\sqrt[3]{\frac{l}{2}} - \sqrt[3]{\frac{16}{729l}} - \frac{1}{3} \right) < k \sqrt[3]{\frac{l}{2}} \quad (14)$$

Hence we have:

$$m_r = n - n\alpha - 2\alpha m_c \quad (15)$$

Assuming $\sqrt[3]{2l^2} < 2 + 1/(1 - \alpha)$ implies $m_r < m_c$. ■

Proof of Lemma 2.

We show that deviations from our derived bound is insignificant by following these steps: 1) We use Corollary 5.11 in [17] which in case of our problem, states that the distribution of number of non-zeros spanned by each sample is approximately Poisson with mean k/m_c . 2) Assuming independent Poisson distributions of non-zero coefficients in compressive samples, we compute the probability that a given sample spans more than one non-zero and find its variance and mean. Further one can easily find μ , the average number of samples spanning at least two non-zero coefficients. 3) Assuming independent Poisson distributions of non-zeroes in the compressive samples, let γ be the number of samples spanning at least two non-zero coefficients, we find the minimum t such that: $\Pr(\gamma - \mu \geq t) < 1/k$. In words, t (with high probability) is the maximum deviation of γ from μ . Then we show that under a realistic presupposition, we have: $t \ll k$. 4) Applying Corollary 5.11 from [17] to the value of t computed in the step

3, we have: in the actual distribution of non-zero coefficients among the partitions (i.e. multinomial), the maximum number of samples spanning at least two non-zeros is less than $\mu + t$ with a probability of higher than $1 - \frac{2}{k} = 1 - O(1/k)$.

Let us assume $k' = k'_1, \dots, k'_{m_c}$ represents the number of non-zeros spanned by each complex compressive sample in Poisson case. Let the binary random variable ν_i be one when $k'_i > 1$ and zero otherwise. Then $\Pr(\nu_i = 1) = 1 - \alpha$. Hence $E(\nu_i) = 1 - \alpha$ and $\sigma^2(\nu_i) = \alpha(1 - \alpha)$. Define $\sigma^2 = \sum(\sigma_i^2)/m_c = \sigma_i^2$. By Bennet's inequality [17] we have:

$$\Delta = \Pr \left(\sum_{i=1}^{m_c} (\nu_i - E(\nu_i)) > t \right) \leq \exp \left(-m_c \sigma^2 h \left(\frac{t}{m_c \sigma^2} \right) \right)$$

where $h(u) = (1+u) \log(1+u) - u > \frac{3u^2}{6+2u}$ for $u \geq 0$. Solving $\Delta \leq 1/k$ gives $t \approx \sqrt{2\sigma^2 m_c \ln k}$. It is straightforward to see that when $n(2\sigma^2 \ln k)^3 \ll k^4$ then we have $t \ll k$ and this completes the proof. ■

ACKNOWLEDGMENT

This work was supported in part by NSF.

REFERENCES

- [1] E. Cands, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information". IEEE Trans. on IT, Feb. '06.
- [2] D. Donoho and Y. Tsaig, "Extensions of compressed sensing". (Signal Processing, 86(3), pp. 533-548, March 2006).
- [3] D. L. Donoho, Y. Tsaig, I. Drori, and J. Starck, "Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit". Preprint, 2007.
- [4] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples". Preprint, 2008.
- [5] Y. C. Pati R. Rezaifar, and P. S. Krishnaprasad, "Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition", Proceedings of the 27 th Annual Asilomar Conference on Signals, Systems, and Computers, Nov. '93.
- [6] S. Chen, , D. Donoho, and M. Saunders. "Atomic decomposition by basis pursuit". Tech. Report 479, Dept. of Stat., Stanford Univ., May '95.
- [7] R. Berinde and P. Indyk, "Sparse recovery using sparse random matrices", Preprint, 2008.
- [8] P. Indyk and R. Berinde "Sequential Sparse Matching Pursuit", Allerton, 2009.
- [9] A. A. Moghadam and H. Radha, "Complex Sparse Projections for Compressed Sensing," in the Proceedings CISS 10, Princeton.
- [10] S. Jafarpour, W. Xu and et. al, "Efficient compressed sensing using high-quality expander graphs". IEEE Trans. IT, 2009.
- [11] Dror Baron, Shriram Sarvoham, Richard G. Baraniuk, "Bayesian compressive sensing via belief propagation". (To appear in IEEE Trans. Signal Processing, 2009).
- [12] M. Duarte, M. Davenport and et. al "Single-pixel imaging via compressive sampling." IEEE SP Magazine, 25(2) pp. 83-91, March 08.
- [13] P. Nagesh and B. Li, "Compressive Imaging of Color Images". IEEE ICASSP, Taipei, Taiwan, 09.
- [14] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," IEEE Trans. on IP, 14(12):2091-2106.
- [15] R. Eslami and H. Radha, "New Image Transforms using Hybrid Wavelets and Directional Filter Banks: Analysis and Design," ICIP, Sep. 2005.
- [16] I. W. Selesnick, R. G. Baraniuk and N. G. Kingsbury, "The Dual-Tree Complex Wavelet Transform," IEEE SP Magazine, 22(6):123-151, 2005.
- [17] "Probability and Computing, Randomized Algorithms and Probabilistic Analysis", Michael Mitzenmacher and Eli Upfal, Cambridge University Press, 2005.
- [18] "Iterative Methods for Solving Linear Systems", Anne Greenbaum, Society for Industrial and Applied Mathematics, 1997.
- [19] David L. Donohoa, A. Maleki and A. Montanari, "Message-passing algorithms for compressed sensing", preprint, 2009.