# Hybrid Compressed Sensing

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 $k$-sparse signal ($x$) from hybrid (complex and real), noiseless compressive 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 ($k/n$), 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

Consider a $k$-sparse real-valued signal $x$ defined in an $n$-dimensional space. By $k$-sparse we mean that $x$ is non-zero only in $k$ coordinates: $k = \|x\|_0 := \#\{i : x_i \neq 0\}$. The main idea of Compressed Sensing (CS) [1]-[2] is that instead of sensing $n$ samples, the encoder projects $x$ onto $m$ compressive samples $y$ according to a projection matrix $P$, i.e. $y = Px$. Generally the under-determined system of equations $y = Px$ has an infinite set of solutions. However, if certain properties outlined in [1]-[2] are met then it is possible to find a unique solution using the following optimization problem:

$$x = \arg\min \|\hat{x}\|_0 = \arg\min \|\hat{x}\|_1 : y = P\hat{x} \qquad (1)$$

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 ($\ell_1$ optimization) algorithms [6] recover the signal with fewer samples but they are extremely complex ($\Omega(n^3)$) relative to greedy algorithms. This led to many efforts with promising results that demonstrated improvements over traditional CS approaches [3], [7]-[10]. More recently, there has been some interest in pursuing sparse projections that lead to low complexity and low measurement requirements [8]-[10]. 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 $O(n \log(n/k))$ [9]. Despite the clear benefits of sparse projections, one needs to use rather strict structures (e.g., expander graphs) to achieve these bounds simultaneously [7]-[9]. Even then, low values for the "big $O$" constant of the measurement bound $m = O(k \log(n/k))$ cannot be guaranteed. In this context, using traditional dense Gaussian matrices requires minimal measurements, however, at the expense of a high complexity $\ell_1$ 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]-[9]. 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 ($k/n$). 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

A sparse-projection based method naturally projects a subset $x_{M_i}$ of the signal $x$ ($x_{M_i} \subset \{x_1, \ldots, x_n\}$) into an observations $y_i$, where $y_i \in \{y_1, \ldots, y_m\}$ 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 signal[1] $x$. 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]-[10]. 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 $i \subseteq [b] = \{1, \ldots, b\}$ and $j \subseteq [c]$. $A^T$ represents the transpose of $A$. By $A_{b,}$ and $A_{,c}$ 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 $[d] \backslash e$) to zero.

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

### A. HCS Projection Matrices

The projection matrix ($P$) of HCS for a length $n$ signal is composed of two sub-matrices ($P^{(c)}$) and ($P^{(r)}$):

$$P = \begin{bmatrix} P^{(c)} \\ P^{(r)} \end{bmatrix} \quad (2)$$

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 $m_r$). 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 $M = \{M_1, \ldots, M_{m_c}\}$ be a set of equal size, random subsets over $[n] = 1, 2, \ldots, n$ such that $M$ forms a *partition*. Thus: $\forall i, j \in [n], i \neq j : |M_i| = w, M_i \cap M_j = \emptyset, \cup_{i=1}^{m_c} M_i = [n]$. 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 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: $\forall i \in [m_c] : P^{(c)}_{i,M_i} = f^{(i)}$ and $P^{(c)}_{i,[n]\backslash M_i} = 0$. Thus the $i^{th}$ 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 $P^{(c)}_{i,M_i} = f^{(i)} = [\, e^{j\phi_1^{(i)}} \ldots e^{j\phi_w^{(i)}} \,]$ be the corresponding vector of complex exponentials (on the unit circle) such that the phases of $f^{(i)}$ are distinct in the range of $[0, \pi)$, $\forall l, t \in [w], l \neq t : \phi_l^{(i)} \in [0, \pi), \phi_t^{(i)} \neq \phi_l^{(i)}$. If $x_{M_i}$ is non-zero in at most one index ($\|x_{M_i}\|_0 \leq 1$) then having $y_i = P^{(c)}_{i,} x = f^{(i)} x_{M_i}$ 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 $f^{(i)}$ ($\{\phi_1^{(i)}, \ldots, \phi_w^{(i)}\}$), 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 $f^{(i)} = P^{(c)}_{i,M_i}$, $w = |M_i|$ and $\angle f^{(i)} = [\angle f^{(i)}_1 \ \ldots \ \angle f^{(i)}_w]$. 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:

$$\angle y_i = \angle f^{(i)}_t \Leftrightarrow x_l = \left\{ \begin{array}{cc} \|y_i\| & l = M_{i,t} \\ 0 & l \in M_i \backslash t \end{array} \right\} \quad (3)$$

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:

$$I' = \left\{ i \in [m_c] : y_i = 0 \ or \ \angle y_i \in \angle f^{(i)} \right\} \quad (4)$$

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 [12] 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 [12]. Hence probability of the event that a complex compressive sample (say $y_i, i \in [m_c]$) spans $j$ non-zero(s) is:

$$\Pr\left(\|x_{M_i}\|_0 = j\right) = \binom{k}{j} \left(1 - \frac{1}{m_c}\right)^{k-j} \left(\frac{1}{m_c}\right)^j \quad (5)$$

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

$$E(|I'|) = m_c \Pr\left(\|x_{M_i}\|_0 \leq 1\right) \approx m_c \left(1 + \frac{k}{m_c}\right) e^{-\frac{k}{m_c}} \quad (6)$$

Since each complex sample spans $w = n/m_c$ coefficients, the number of identified coefficients ($I$) on average is $E(|I|) = wE(|I'|) = n(1 + k/m_c)\exp\left(-k/m_c\right)$. 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 $m_r \geq n - E(|I|) - 2\left(m_c - I'\right)$ 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:

$$\underset{m_c, m_r}{\arg\min} \ 2m_c + m_r \ s.t. \ m_r \geq n - E(|I|) - 2\left(m_c - I'\right) \quad (7)$$

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.)

**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 $n\left(2(\alpha - \alpha^2)^2 \ln k\right)^3 \ll k^4$ 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 $f^{(i)} = [\exp(j\phi^{(i)}_1) \ \exp(j\phi^{(i)}_2) \ \ldots \ \exp(j\phi^{(i)}_w)]$ are sorted based on phases: $a, b \in [w], a > b \Leftrightarrow \angle f^{(i)}_a = \phi^{(i)}_a > f^{(i)}_b = \angle\phi^{(i)}_b$. 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 $m_r + m_c - |I'| \leq 2m_c$ equations. Let $g(\beta)$ be the complexity of linear solver employed in the second stage of HCS solver to solve a full rank system of $\beta$ equations. Then the complexity of the second stage of HCS is $O\left(g(m_r + m_c - |I'|)\right) \leq O\left(g(2m_c)\right)$. Finally the complexity of HCS solver is $O\left(\max\{m_c \log(n/m_c), g(2m_c)\}\right)$.

## IV. Simulation Results

We tested the proposed HCS on a large number of standard signals from SparseLab [11]. Specifically we compared the performance of HCS (in terms of quality of the recovered signal as a function of the number of compressive samples and the required time for recovery) with the popular Gaussian random projection matrix and dominant Basis Pursuit (BP) and OMP solvers. In our simulations, we counted each complex valued compressive sample as two samples. In all plots $m$, $n$ and $k$ represent the total number of (real) compressive samples, the length of the signal and the number of non-zero DCT or wavelet coefficients, respectively. In all simulations, we have assumed that for a given total 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.

We performed our simulations under various configurations of signal length and sparsity ratio which we present some of them in Fig.1. In all these scenarios $m \approx 4k$ samples were sufficient for the perfect recovery under HCS. Although BP (except Fig.1a) achieved virtually perfect reconstruction, but clearly the complexity of BP is much higher compared to HCS.

## V. Conclusion

In this paper, we considered the problem of recovering a $k$-sparse signal $x$ from a limited number of linear samples ($y = Px$) 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 $O\left(k \log(n/k)\right)$, 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. [13]) 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 $\eta$. 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: $x_{M_i} = \mathbf{0} \Leftarrow y_i = 0$ 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 $f^{(i)}$ ($\angle f^{(i)} = \{\phi_1^{(i)}, \ldots, \phi_w^{(i)}\}$). 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:

$$\angle y_i = \angle f_l^{(i)} \Leftrightarrow x_u = \left\{ \begin{array}{ll} \|y_i\| & u = M_{i,l} \\ 0 & u \in M_i \backslash M_{i,l} \end{array} \right\} \quad (8)$$

Finally if $y_i \neq 0$ and the phase of $y_i$ 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:

$$\arg\min_{m_c, m_r} 2m_c + m_r \;\; s.t. \;\; m_r \geq n - E(|I|) - 2\left(m_c - I'\right) \quad (9)$$

One can use the Lagrangian multiplier to solve (9). Define $\alpha = E(|I'|)/m_c < 1$, then the objective function ($L$) is:

$$L = 2m_c + m_r + \lambda\left(\alpha n + 2(1-\alpha)m_c + m_r - n\right) \quad (10)$$

Taking derivatives respect to three parameters $\lambda$, $m_c$ and $m_r$ gives us three equations:

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

$$\frac{\partial L}{\partial m_c} = 0 \Rightarrow 2m_c^3 + 2km_c^2 + 2k^2 m_c = nk^2 \quad (12)$$

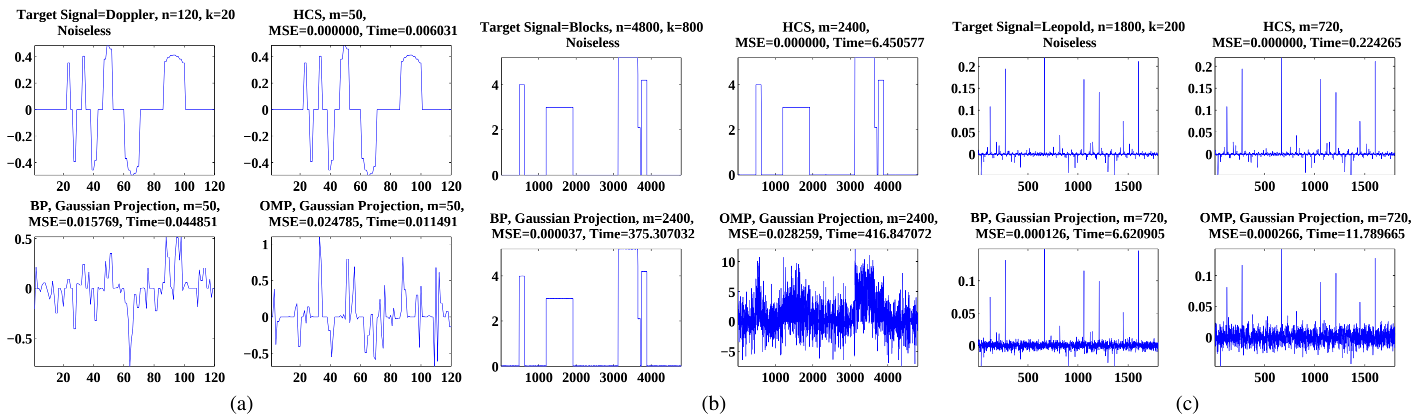


Fig. 1: $m$, $n$ and $k$ represent number of samples, signal length and signal sparsity respectively.

Define $l = n/k$. Let us assume $l^2 \gg l$, then solving (12) yields:

$$m_c = k\left(\sqrt[3]{\frac{l}{2}} - \sqrt[3]{\frac{16}{729l}} - \frac{1}{3}\right) < k\sqrt[3]{\frac{l}{2}} \tag{13}$$

Hence we have:

$$m_r = n - n\alpha - 2\alpha m_c \tag{14}$$

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 [12] 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 $\mu$, 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 $\gamma$ 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 $\gamma$ from $\mu$. Then we show that under a realistic presupposition, we have: $t \ll k$. 4) Applying Corollary 5.11 from [12] 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 $k' = k'_1, \ldots, k'_{m_c}$ represents the number of non-zeros spanned by each complex compressive sample in Poisson case. Let the binary random variable $\nu_i$ 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 [12] 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:

$$t \approx \sqrt{2\sigma^2 m_c \ln k} \tag{15}$$

It is straightforward to see that when $n\left(2\sigma^2 \ln k\right)^3 \ll k^4$ then we have $t \ll k$ and this completes the proof. ■

## Acknowledgment

This work was supported in part by NSF.

## References

[1] Emmanuel Cands, Justin Romberg, and Terence Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information". IEEE Trans. on Information Theory, Feb. 2006.
[2] David Donoho and Yaakov Tsaig, "Extensions of compressed sensing". (Signal Processing, 86(3), pp. 533-548, March 2006).
[3] David L. Donoho, Yaakov Tsaig, Iddo Drori, and Jean-Luc 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. Rezaiifar, 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. 1993.
[6] S. Chen, , D. Donoho, and M. Saunders. "Atomic decomposition by basis pursuit". Tech. Report 479, Dept. of Statistics, Stanford Univ., May 1995.
[7] Radu Berinde and Piotr Indyk, "Sparse recovery using sparse random matrices", Preprint, 2008.
[8] Piotr Indyk and R. Berinde "Sequential Sparse Matching Pursuit", Allerton, 2009.
[9] Sina Jafarpour, Weiyu Xu, Babak Hassibi, and Robert Calderbank, "Efficient compressed sensing using high-quality expander graphs". to appear in IEEE Trans Info Theory, 2009.
[10] Dror Baron, Shriram Sarvoham, Richard G. Baraniuk, "Bayesian compressive sensing via belief propagation". (To appear in IEEE Trans. Signal Processing, 2009)
[11] http://sparselab.stanford.edu
[12] "Probability and Computing, Randomized Algorithms and Probabilistic Analysis", Michael Mitzenmacher and Eli Upfal, Cambridge University Press, 2005.
[13] "Iterative Methods for Solving Linear Systems", Anne Greenbaum, Society for Industrial and Applied Mathematics, 1997.