Abstract—Compressive sensing is a new signal acquisition technology with the potential to reduce the number of measurements required to acquire signals that are sparse or compressible in some basis. Rather than uniformly sampling the signal, compressive sensing computes inner products with a randomized dictionary of test functions. The signal is then recovered by a convex optimization that ensures the recovered signal is both consistent with the measurements and sparse. Compressive sensing reconstruction has been shown to be robust to multi-level quantization of the measurements, in which the reconstruction algorithm is modified to recover a sparse signal consistent to the quantization measurements. In this paper we consider the limiting case of 1-bit measurements, which preserve only the sign information of the random measurements. Although it is possible to reconstruct using the classical compressive sensing approach by treating the 1-bit measurements as $±1$ measurement values, in this paper we reformulate the problem by treating the 1-bit measurements as sign constraints and further constraining the optimization to recover a signal on the unit sphere. Thus the sparse signal is recovered within a scaling factor. We demonstrate that this approach performs significantly better compared to the classical compressive sensing reconstruction methods, even as the signal becomes less sparse and as the number of measurements increases.

I. INTRODUCTION

Compressive sensing is a new low-rate signal acquisition method for signals that are sparse or compressible [1]–[5]. The fundamental premise is that certain classes of signals, such as natural images or communications signals, have a representation in terms of a sparsity inducing basis (or sparsity basis for short) where most of the coefficients are zero or small and only a few are large. For example, smooth signals and piecewise smooth signals are sparse in a Fourier and wavelet basis, respectively.

Recent results [1], [5] demonstrate that sparse or compressible signals can be directly acquired at a rate significantly lower than the Nyquist rate. The low-rate acquisition process projects the signal onto a small set of vectors, a dictionary, that are incoherent with the sparsity basis. The signals can subsequently be recovered using a greedy algorithm or a linear program that determines the sparsest representation consistent with the acquired measurements. The quality of the reconstruction depends on the compressibility of the signal, the choice of the reconstruction algorithm, and the incoherence of the sampling dictionary with the sparsity basis. One of the most useful results is that randomly generated dictionaries are universal in the sense that, with very high probability, they are incoherent with any fixed sparsity basis. This property makes such dictionaries very desirable for compressive sensing applications.

To date most of the compressive sensing literature does not explicitly handle quantization of the compressive sensing measurements. Most of the literature models quantization as norm-limited additive noise on the measurements. Although this treatment is sufficient in the case of high-rate quantization, it is extremely conservative if the measurements are coarsely quantized to very low bit-rates.

In this paper we examine the reconstruction from compressive sensing measurements quantized to one bit per measurement. Quantization to 1-bit measurements is particularly appealing in hardware implementations. The quantizer takes the form of a comparator to zero, an extremely inexpensive and fast hardware device. Furthermore, 1-bit quantizers do not suffer from dynamic range issues. If the analog side of the measurement system is properly implemented then the sign of the measurement remains valid even if the quantizer saturates. The appeal of 1-bit quantization is evident in the success of 1-bit Sigma-Delta converters, which use such quantizers at the expense of very high sampling rate [6], [7].

The main contribution of this paper is treating the quantized measurements as sign constraints on the measurements as opposed to values to be matched in the reconstruction in a mean squared sense. Since the signs of the measurements do not provide amplitude information of the signal, any positive scalar multiple of the reconstructed signal, including the zero signal, is consistent with the measurements. The signal can only be recovered within that scalar factor. The reconstruction formulation we provide resolves the ambiguity by imposing a unit energy constraint on the reconstructed signal and recovering the signal on the unit sphere. This constraint significantly reduces the reconstruction search space, thus significantly improving the reconstruction results.

An important principle we use is consistent reconstruction which states that the reconstructed signal should be consistent with the quantized measurements. In other words, if the reconstructed signal is to be measured and quantized with the same system then it should produce the same measurements as the ones at hand. Consistent reconstruction is very effective in increasing the reconstruction performance for quantized frame expansions [8]. We demonstrate that the same principle significantly improves the reconstruction from compressive sensing measurements.
The next section provides a brief background on compressive sensing, quantization, and consistent reconstruction. The aim is to establish the notation and serve as a quick reference. Section III-A introduces the 1-bit measurement model and Section III-B formulates the consistent reconstruction problem. Section III-C presents an algorithm to compute the solution and reconstruct the signal. Section IV presents simulation results that validate the formulation and demonstrate significant performance increase over classical compressive sensing reconstruction. Conclusions are presented in Section V.

II. BACKGROUND

A. Compressive Sensing

Compressive sensing is a new sampling and reconstruction method for signals that are known to be sparse or compressible in some basis [1], [5]. Without loss of generality, we assume a signal $x$ in an $N$-dimensional vector space. The signal is $K$-sparse in a sparsity-inducing basis $\{b_i\}$ if there are at most $K$ non-zero coefficients $\{\alpha_i\}$ in the basis expansion $x = \sum_i \alpha_i b_i$—compactly denoted $B\alpha$. The signal is $K$-compressible if it is well approximated by the $K$ most significant coefficients in the expansion.

The signal is sampled using $M$ measurements with the measurement vectors $\phi_i$, $i = 1, \ldots, M$:

$$y_i = \langle x, \phi_i \rangle. \quad (1)$$

We compactly denote (1) using $y = \Phi x = \Phi B\alpha$, where $y$ is the vector of measurements and $\Phi$ is the measurement operator which models the measurement system.

For the remainder of this paper we assume the signal is sparse or compressible in the canonical basis, i.e., $B = I$. It is straightforward to apply the subsequent results to signals sparse in any basis by substituting $\Phi = \Phi B$ as the measurement system and treating $\alpha$, instead of $x$, as the sparse signal to be reconstructed.

Classical sampling theory dictates that for robust linear reconstruction of any signal $x$, the set $\phi_i$ should form a Riesz basis or a frame. This implies that at least $N$ measurements are necessary to recover the signal. Compressive sensing, on the other hand, allows the use of only $M = O(K\log(N/K))$ non-adaptive measurements to robustly acquire and reconstruct $K$-sparse or compressible signals. The reconstruction in this case ceases to be linear.

The reconstruction from $y$ amounts to determining the sparsest signal that explains the measurements $y$. The strictest measure of sparsity is the $\ell_0$ pseudonorm of the signal, defined as the number of non-zero coefficients of the signal. Unfortunately the $\ell_0$ pseudonorm is combinatorially complex to optimize for. Instead compressive sensing enforces sparsity by minimizing the $\ell_1$ norm of the reconstructed signal, $\|x\|_1 = \sum_i |x_i|$. In several cases, including the classical compressive sensing reconstruction methods, minimizing the $\ell_1$ norm has been theoretically proven equivalent to minimizing the $\ell_0$ pseudonorm of the signal [9], [10].

Reconstruction from compressive sensing measurements is thus achieved by solving the minimization problem:

$$\hat{x} = \arg\min_x \|x\|_1 \text{ s.t. } y = \Phi x, \quad (2)$$

where $\|x\|_1 = \sum_i |x_i|$ is the $\ell_1$ norm of the signal. In [3] it is shown that with proper selection of the measurement system $\Phi$, (2) exactly recovers the signal.

Specifically, exact recovery requires that the measurement vectors $\{\phi_i\}$ are sufficiently incoherent with the sparsity basis $\{b_i\}$. Incoherence, as defined in [11], can be guaranteed with very high probability if the measurement matrix $\Phi$ is drawn randomly from a variety of possible distributions. The number of measurements necessary to guarantee recovery using a random measurement system is $M = O(K\log(N/K))$.

B. Measurement Quantization

Quantization is usually modeled as measurement noise, denoted using $\mathbf{n}$, added to the measurements:

$$y = Q(\Phi x) = \Phi x + \mathbf{n}, \quad (3)$$

where $Q(\cdot)$ is the quantizer and $\mathbf{n}$ is energy-limited to some $\epsilon$ depending on the quantization accuracy:

$$\|\mathbf{n}\|_2 = \left(\sum_i |n_i|^2\right)^{1/2} \leq \epsilon. \quad (4)$$

For a uniform linear quantizer with quantization interval $\Delta$, $\epsilon \leq \sqrt{M\Delta^2/12}$.

In the presence of norm-limited measurement noise such as quantization, it has been shown that robust reconstruction can be achieved by solving:

$$\hat{x} = \arg\min_x \|x\|_1 \text{ s.t. } \|y - \Phi x\|_2 \leq \epsilon. \quad (5)$$

In this case, the reconstruction error norm is bounded by $\|x - \hat{x}\|_2 \leq C\epsilon$, where the constant $C$ depends on the properties of the measurement system $\Phi$ but not on the signal [4].

In practice, the optimization in (5) is often relaxed to:

$$\hat{x} = \arg\min_x \|x\|_1 + \frac{\lambda}{2} \|y - \Phi x\|_2^2, \quad (6)$$

which is often more efficient to solve algorithmically (for some examples, see [12]–[15]). The solution path of (5) as $\epsilon$ decreases is the same as the solution path of (6) as $\lambda$ increases. However, the exact correspondence of $\lambda$ and $\epsilon$—and, therefore, the appropriate value for $\lambda$ in (6)—cannot be known in advance before the solution is obtained.

C. Consistent Reconstruction

Consistent reconstruction enforces the requirement that the solution should be consistent with all our knowledge about the signal and measurement process. In the case of quantized measurements, this implies that if the reconstructed signal is re-measured using the measurement system $\Phi$ and quantized at the same accuracy then the measurements should be exactly the same as the original measurements used to reconstruct the
signal. In [8] it is shown that consistent reconstruction significantly improves the reconstruction performance in quantized frame representations.

Although in the general case of norm-limited measurement noise the reconstruction in (5) is consistent with the measurements, this is not the case if the measurement noise is due to quantization. Specifically, in the case of uniform linear quantization, all noise components have magnitude \(|n_i| \leq \Delta/2\), which implies that consistent reconstruction should produce a signal that satisfies:

\[
|\langle \Phi \hat{x} - y \rangle_i| \leq \frac{\Delta}{2}. \tag{7}
\]

In the case of 1-bit quantization, the quantizer is most often implemented as a comparator to a voltage level \(\ell\), usually zero. In this case consistent reconstruction should require that measurements of the reconstructed signal should be on the same side of the voltage level as the measurements obtained from the measurement system:

\[
(\Phi \hat{x})_i y_i = \begin{cases} +1 & \text{if } \hat{x}_i \geq \ell \\ \pm \ell & \text{if } \|x\|_2 < \ell \\ -1 & \text{if } \hat{x}_i < \ell \end{cases}, \tag{8}
\]

where \(y_i = \pm 1\) is the quantized measurement. This is equivalent to:

\[
\text{sign} \left( (\Phi \hat{x})_i - \ell \right) = y_i. \tag{9}
\]

The remainder of this paper examines how the principle of consistent reconstruction can be applied to significantly improve the reconstruction from compressive sensing measurements quantized to one bit. In our development we assume \(\ell = 0\).

### III. 1-BIT COMPRESSIVE SENSING MEASUREMENTS

#### A. Measurement Model

As we describe above, each measurement is the sign of the inner product of the sparse signal with a measurement vector \(\phi_i\):

\[
y_i = \text{sign}(\langle \phi_i, x \rangle). \tag{10}
\]

It follows that the product of each quantized measurement with the measurement is always non-negative:

\[
y_i \text{sign}(\langle \phi_i, x \rangle) \geq 0. \tag{11}
\]

The measurements are compactly expressed using:

\[
y = \text{sign}(\Phi x), \tag{12}
\]

where \(y\) is the vector of measurements, \(\Phi\) is a matrix representing the measurement system and the 1-bit quantization function \(\text{sign}(\cdot)\) is applied element-wise to the vector \(\Phi x\). Using matrix notation, (11) is compactly expressed using:

\[
Y \Phi x \geq 0, \tag{13}
\]

where \(Y = \text{diag}(y)\) and the inequality is applied element-wise.

#### B. Consistent Reconstruction

For consistent reconstruction from 1-bit measurements we treat the measurements as sign constraints that we enforce in the reconstruction to recover the signal. In the reconstruction we enforce the model using the \(\ell_1\) norm as a sparsity measure.

If \(x\) is consistent with the measurements then so is \(ax\) for all \(0 \leq a < 1\). Since \(\|ax\|_1 = a\|x\|_1 < \|x\|_1\), a minimization-based reconstruction algorithm that only requires consistency with the measurements will drive the solution to \(x = 0\). To enforce reconstruction at a non-trivial solution we need to artificially resolve the amplitude ambiguity. Thus, we impose an energy constraint that the reconstructed signal lies on the unit \(\ell_2\)-sphere:

\[
\|x\|_2 = \left( \sum_i x_i^2 \right)^{1/2} = 1. \tag{14}
\]

Note that this constraint significantly reduces the optimization search space. This reduction plays an important role in improving the reconstruction performance.

The sparsest signal on the unit sphere that is consistent with the measurements, as expressed in (13), is the solution to:

\[
\tilde{x} = \arg\min_x \|x\|_1 \quad \text{s.t. } Y \Phi x \geq 0 \quad \text{and } \|x\|_2 = 1. \tag{15}
\]

To enforce the constraint we relax the problem using a cost function \(f(x)\) that is positive for \(x < 0\) and zero for \(x \geq 0\) and a relaxation parameter \(\lambda\):

\[
\tilde{x} = \arg\min_x \|x\|_1 + \lambda \sum_i f(\langle Y \Phi x \rangle_i) \tag{16}
\]

\[
\text{s.t. } \|x\|_2 = 1,
\]

Assuming that the original problem (15) is feasible, as \(\lambda\) tends to infinity (15) and (16) have the same solution.

The algorithm we introduce in this paper minimizes (16) for \(f(x) = \frac{x^2}{2} \cdot u(-x)\), where \(u(x)\) is the unit step function. In other words, \(f(x)\) is a one-sided quadratic penalty if \(x\) is negative and zero otherwise:

\[
f(x) = \begin{cases} \frac{x^2}{2}, & x < 0 \\ 0, & x \geq 0. \end{cases} \tag{17}
\]

The convexity and smoothness of this function allows the use of gradient descent and fixed-point methods to perform the minimization.

For notational convenience, in the remainder of this paper we use \(g(x) = \|x\|_1\) to denote the \(\ell_1\) norm part of the cost function, and \(f(Y \Phi x)\) to denote the one-sided quadratic penalty:

\[
\bar{f}(x) = \sum_i f(x_i), \tag{18}
\]

such that the cost function is equal to:

\[
\text{Cost}(x) = g(x) + \lambda \bar{f}(Y \Phi x). \tag{19}
\]
C. Reconstruction Algorithm

We employ a variation of the fixed point continuation (FPC) algorithm introduced in [13]. Specifically, we introduce two modifications. The first modifies the computation of the gradient descent step in [13] such that it computes the gradient of the one-sided quadratic penalty in (17) projected on the unit sphere \( \|x\|_2 = 1 \). The second introduces a renormalization step after each iteration of the algorithm to enforce the constraint that the solution lies on the unit sphere.

These modifications are similar to the ones introduced in [16] to stabilize the reconstruction of sparse signals from their zero crossings. The similarity is not coincidental. Both sign measurements and zero crossings information eliminate amplitude information from the signal. The main difference between the two problems is that measurements of zero crossings are signal-dependent, whereas compressive measurements are signal-independent. Although it is possible to reformulate the reconstruction from zero-crossings as reconstruction from 1-bit measurements, this reformulation is beyond the scope of this paper.

The algorithm computes and follows the gradient of the cost function in (19). If the minimization is not constrained on the sphere, then the gradient of the cost at the minimum is 0:

\[
\text{Cost}'(x) = 0 = g'(x) + \lambda (Y \Phi)^T j'(Y \Phi x)
\]

\[
\Rightarrow g'(x) = - (Y \Phi)^T j'(Y \Phi x),
\]

where

\[
(g'(x))_i = \begin{cases} 
-1, & x_i < 0 \\
-1, & x_i = 0 \\
+1, & x_i > 0 
\end{cases}
\]

and

\[
(f'(x))_i = \begin{cases} 
-x_i, & x_i \leq 0 \\
0, & x_i > 0 
\end{cases}
\]

It follows that if the sphere constraint is introduced then the gradient of the cost function at the minimum is orthogonal to the sphere. Thus, a gradient descent algorithm followed by renormalization has the minimum of (19) on the unit sphere as a fixed point.

The iterative steps to reconstruct the signal are presented in Algorithm 1. The algorithm is seeded with an initial signal estimate \( \bar{x}_0 \) and a gradient descent step size \( \delta/\lambda \). At every iteration the algorithm computes the gradient of the one-sided quadratic in Step 3, projects it on the sphere in Step 4 and descends on that gradient in Step 5. Step 6 is a shrinkage step using the soft threshold shrinkage function shown in the solid line in Figure 1. Step 7 renormalizes the estimate to have unit magnitude and the algorithm iterates from Step 2 until the solution converges.

As discussed in [13], the shrinkage Step 6 is interpreted as a gradient descent on the \( \ell_1 \)-norm component of the cost function. Specifically, for \( |x_i| \geq \delta/\lambda \) the magnitude of the coefficient is reduced by \( \delta/\lambda \), which is the expected behavior of a gradient descent. For \( |x_i| \leq \delta/\lambda \) the discontinuity at 0 zero makes the gradient descent set the coefficient to 0.

The reconstruction algorithm should be executed with \( \lambda \) large enough such that the relaxed minimization (16) converges to the constrained minimization in (15). Unfortunately, the larger the value of \( \lambda \), the smaller the descent step \( \delta/\lambda \). Furthermore, the value of \( \lambda \) that is sufficiently large is not known in advance of the algorithm.

Both issues are resolved by wrapping the algorithm in an outer iteration loop that executes the algorithm using a small value \( \lambda_0 \) until convergence and then restarts the algorithm with a higher value \( \lambda_i = c \lambda_{i-1} \), \( c > 1 \) using the previous estimate as a seed for the next execution. The outer loop terminates once the solution from the current iteration is not significantly different from the solution of the previous iteration.

Since the minimization is performed on the unit sphere, the problem is not convex. Therefore, the algorithm cannot be guaranteed to converge to the global minimum. Our simulations have demonstrated that a good heuristic is to initialize the algorithm using

\[
\bar{x}_0 = \Phi^\dagger y,
\]

Algorithm 1 Renormalized Fixed Point Iteration

1) Initialization:

\( \bar{x}_0 \), s.t. \( \|\bar{x}_0\|_2 = 1 \),

Descent Step Size: \( \delta \),

Counter: \( k \leftarrow 0 \)

2) Counter Increase:

\( k \leftarrow k + 1 \)

3) One-sided Quadratic Gradient:

\( \Gamma_k \leftarrow - (Y \Phi)^T j'(Y \Phi x_{k-1}) \)

4) Gradient Projection on Sphere Surface:

\( f_k \leftarrow \frac{\Gamma_k}{\|\Gamma_k\|_2} \)

5) One-sided Quadratic Gradient Descent:

\( h \leftarrow \frac{\bar{x}_{k-1} - \delta f_k}{\|f_k\|_2} \)

6) Shrinkage (\( \ell_1 \) gradient descent):

\( (u)_i \leftarrow \text{sign}(h_i) \max \left\{ (h_i) - \frac{\delta}{\lambda}, 0 \right\} \), for all \( i \)

7) Normalization:

\( \bar{x}_k \leftarrow \frac{u}{\|u\|_2} \)

8) Iteration: Repeat from 2 until convergence.

Fig. 1: \( \ell_1 \) shrinkage function (soft threshold).
where $\Phi^\dagger$ is the pseudo-inverse of $\Phi$. The solution to
\[
\widehat{x}_0 = \arg\min_x \|x\|_1 + \frac{\lambda}{2} \|\Phi x - y\|_2^2
\]
for some small value of $\lambda$, also proved to be a good initialization heuristic. Random initialization of the algorithm also converges with extremely high probability.

IV. SIMULATION RESULTS

In this section we present simulation results that demonstrate the performance of the algorithm. For the remainder of this section we use randomly generated sparse signals of length $N = 512$. The signals have varying sparsity $K$, with the $K$ nonzero coefficients uniformly selected among all the possible $N$ signal coefficients. Each nonzero coefficient is drawn from a standard normal distribution. The measurement matrix $\Phi$ has i.i.d. coefficients also drawn from a standard normal distribution. The results we present are robust to variations of the parameters.

Figure 2 demonstrates the performance of our algorithm compared to the optimization in (5), implemented using the LARS algorithm with the LASSO modification [12]. The output of both algorithms was normalized to have unit power, and compared to the normalized original signal. The results of our algorithm are presented using solid curves, labeled ‘1-bit CS’. The results of the classical compressive sensing reconstruction are presented using dashed curves. The figure plots the reconstruction error in dB as a function of the number of measurements—which is also the number of bits used to represent the signal—for various values of $K$.

The plots demonstrate the performance advantage of consistent reconstruction on the unit sphere. At low sparsity rates the improvements are significant, reaching 20dB for very sparse signals. Even as the density of the signals increases, reconstruction on the unit sphere outperforms classical compressive sensing, especially as the number of measurements increases.

Figure 2: Reconstruction performance from 1-bit measurements.

Note that we performed the simulations even for $M > N$. This regime transcends the classical compressive sensing goal of few measurements. Still, it is important to study the performance since additional measurements improve performance in the presence of quantization. This is especially useful if the cost of the system is in the quantization accuracy instead of the number of measurements.

V. CONCLUSIONS

Our results demonstrate that reconstruction from 1-bit compressive sensing measurements can be significantly improved if the appropriate measurement model is used in the reconstruction. Specifically, 1-bit measurements eliminate amplitude information, and therefore the signal can only be recovered within a positive scalar factor. Constraining the reconstruction to be on the unit sphere resolves this ambiguity and significantly reduces the reconstruction search space. Similarly, treating each measurement as a constraint instead of a value to be matched in a mean-squared sense allows us to exploit consistent reconstruction principles. Our results demonstrate that both contributions significantly improve the reconstruction performance from 1-bit measurements.

A significant advantage of using the measurements as constraints is that the formulation of the reconstruction becomes entirely non-parametric. Specifically, the minimization programs in (5) and (6) require knowledge of $\epsilon$ or $\lambda$ at the initialization of the problem. On the other hand, (15) contains no such parameter. Although we introduce a parameter $\lambda$ when relaxing the problem in (16), this parameter in never explicitly required in the execution of the algorithm. Instead the algorithm increases $\lambda$ until the solution converges to the solution of (15).

Last, but not least, we note that the focus of our formulation is on number of bits instead of number of measurements. Compressive sensing systems reduce the number of measurements required to recover the signal. In the presence of coarsely quantized measurements, the performance of a compressive sensing system might not be sufficient for the application. Performance can be increased by incorporating more measurements or using a more precise quantizer. The tradeoff depends on the cost of more refined quantization versus the cost of additional measurements. Our approach is mostly applicable in the case where measurements are inexpensive whereas precision quantization is expensive.

REFERENCES


