Exact Signal Recovery from Sparsely Corrupted Measurements through the Pursuit of Justice

Jason N. Laska, Mark A. Davenport, Richard G. Baraniuk
Department of Electrical and Computer Engineering
Rice University
Houston, Texas, 77005

***Abstract*—Compressive sensing provides a framework for recovering sparse signals of length N from $M \ll N$ measurements. If the measurements contain noise bounded by ϵ, then standard algorithms recover sparse signals with error at most $C\epsilon$. However, these algorithms perform suboptimally when the measurement noise is also sparse. This can occur in practice due to shot noise, malfunctioning hardware, transmission errors, or narrowband interference. We demonstrate that a simple algorithm, which we dub *Justice Pursuit* (JP), can achieve exact recovery from measurements corrupted with sparse noise. The algorithm handles unbounded errors, has no input parameters, and is easily implemented via standard recovery techniques.**

I. Introduction

The recently developed *compressive sensing* (CS) framework enables acquisition of a signal $\mathbf{x} \in \mathbb{R}^N$ from a small set of M non-adaptive, linear measurements [1, 2]. This process can be represented as

$$\mathbf{y} = \Phi\mathbf{x} \tag{1}$$

where Φ is an $M \times N$ matrix that models the measurement system. The hope is that we can design Φ so that $\mathbf{x}$ can be accurately recovered even when $M \ll N$. While this is not possible in general, when $\mathbf{x}$ is K-sparse, meaning that it has only K nonzero entries, it is possible to exactly recover $\mathbf{x}$ using Φ with $M = O(K \log(N/K))$. Signal reconstruction can be performed using optimization techniques or greedy algorithms. The broad applicability of this framework has inspired research that extends the CS framework by proposing practical implementations for numerous applications, including sub-Nyquist sampling systems [3–5], compressive imaging architectures [6–8], and compressive sensor networks [9].

In practical settings, there may be many sources of noise, including noise present in the signal $\mathbf{x}$, noise caused by the measurement hardware, quantization noise, and transmission errors in the case where the measurements are sent over a noisy channel. Thus, it is typically more realistic to represent the measurement process as

$$\mathbf{y} = \Phi\mathbf{x} + \mathbf{e}, \tag{2}$$

This work was supported by the grants NSF CCF-0431150, CCF-0728867, CNS-0435425, CNS-0520280, and CCF-0926127, DARPA/ONR N66001-08-1-2065, ONR N00014-07-1-0936, N00014-08-1-1067, N00014-08-1-1112, and N00014-08-1-1066, AFOSR FA9550-07-1-0301, ARO MURIs W311NF-07-1-0185 and W911NF-09-1-0383, and the Texas Instruments Leadership University Program.
Email: {laska, md, richb}@rice.edu. Web: dsp.rice.edu.

where $\mathbf{e}$ is an $M \times 1$ vector that represents noise in the measurements. It has been shown that it is possible to reconstruct the signal with ℓ_2-error that is at most $C_0\|\mathbf{e}\|_2$, where $C_0 > 1$ is a small constant that depends on certain properties of Φ [10, 11]. Thus, CS systems are stable in the sense that if the measurement error is bounded, then the reconstruction error is also bounded.

This is a powerful result for dealing with noise that is evenly distributed across the measurements, such as i.i.d. Gaussian, thermal, or quantization noise. However, this is not representative of some common settings. For example, there may be short bursts of high noise, or certain measurements may be invalid because of defective hardware or spikes in the power supply. When measurements are sent over a network, some measurements may be lost altogether, or in a sensor network, malfunctioning sensors may regularly transmit corrupted measurements while the other sensors do not. In these cases the noise is sparse in the canonical basis. In other settings, the measurement noise may be sparse or compressible when represented in some transform basis Ω. For example, the measurements could be corrupted with 60Hz hum,[1] in which case the noise is sparse in the Fourier basis. Similarly, measurement noise from a DC bias that changes abruptly would be piecewise-smooth and thus sparse in a wavelet basis.

In these cases, $\|\mathbf{e}\|_2$ may be extremely large, and thus the resulting bound $C_0\|\mathbf{e}\|_2$ on the reconstruction error will also be large. However, one can hope to do much better. To see why, suppose that the measurement noise is sparse in the time domain so that only a few of the measurements are corrupted with large errors and that the remaining measurements are noise-free. Standard recovery algorithms will return a signal estimate $\widehat{\mathbf{x}}$ that satisfies only $\|\widehat{\mathbf{x}} - \mathbf{x}\|_2 \leq C_0\|\mathbf{e}\|_2$. However, if we knew which measurements were corrupted, then we could simply ignore them. If Φ is generated randomly with M sufficiently large, and if the locations of the corrupted measurements were known *a priori*, then the signal could be reconstructed exactly by using only the noiseless measurements [12]. The challenge is that it is typically not possible to know exactly which measurements have been corrupted.

In this paper, we consider a more general version of the measurement noise model (2), namely

$$\mathbf{y} = \Phi\mathbf{x} + \Omega\mathbf{e}, \tag{3}$$

[1] In some regions hum consists of a 50Hz sinusoid (and its harmonics).

where Ω is an $M \times L$ matrix with $L \leq M$ orthonormal columns, and the vector $\mathbf{e}$ is sparse. The matrix Ω represents the basis or subspace in which the noise is sparse; e.g., if the measurements are corrupted by hum, then Ω is the Fourier basis. We demonstrate that it is still possible to recover $\mathbf{x}$ exactly when the noise $\mathbf{e}$ is sparse via an algorithm we dub *Justice Pursuit* (JP), since it is able to accurately identify corrupted measurements. The JP algorithm, which has been previously proposed and analyzed in a different context in [13, 14], is described in detail in Section IV. The hallmarks of JP include:

1) *exact recovery* of the sparse signal $\mathbf{x}$;
2) *exact recovery* of the sparse noise term $\mathbf{e}$;
3) *blindness* to the locations and size of the measurement errors — thus, the corrupted measurements could be adversarily selected and the noise on the corrupted measurements can be arbitrarily large;
4) no user-defined parameters;
5) standard CS recovery algorithm implementations can be trivially modified, i.e., *justified*, to perform JP, so that optimized routines can be easily adapted to this setting.

This paper is organized as follows. In Section II, we review the key results of the CS framework and in Section III we introduce our algorithm for recovering sparse signals from corrupted measurements and prove the validity of this algorithm. In Section IV, we demonstrate, via simulation, the performance of our algorithm. In Section V, we discuss the implications of these results and propose future work.

II. Background

In the CS framework, there are typically two main theoretical questions: *(i)* What conditions must Φ satisfy to ensure that the measurement process is well-behaved? and *(ii)* How do we recover $\mathbf{x}$ from the measurements $\mathbf{y}$? The most common answer to the first question is the *restricted isometry property* (RIP), introduced by Candès and Tao [15]. We say that a matrix Φ satisfies the RIP of order K if there exists a constant, $\delta \in (0, 1)$, such that

$$(1-\delta)\|\mathbf{x}\|_2^2 \leq \|\Phi\mathbf{x}\|_2^2 \leq (1+\delta)\|\mathbf{x}\|_2^2, \tag{4}$$

holds for all $\mathbf{x}$ and $\|\mathbf{x}\|_0 := |\text{supp}(x)| \leq K$. In words, Φ acts as an approximate isometry on the set of vectors that are K-sparse. The RIP has been shown to be a sufficient condition on Φ to enable exact recovery of sparse signals, as well as stable recovery from noisy measurements of both sparse and *compressible* signals [16].

A key theoretical CS result is that by acquiring only $M = O(K \log(N/K))$ *random* measurements (by which we mean the entries of Φ are generated independently at random according to a sub-Gaussian distribution), we obtain a Φ that satisfies the RIP of order $2K$ [17]. Furthermore, it has been shown in [12] that the random constructions of Φ typically advocated in CS are *democratic*, meaning that not only does Φ satisfy the RIP, but all sufficiently large submatrices of Φ satisfy the RIP as well. This provides an important inspiration for this work, since it suggests that no matter which measurements are corrupted, the key information about $\mathbf{x}$ is preserved.

We now turn to the problem of recovering $\mathbf{x}$ from the measurements $\mathbf{y}$. In the case where the measurements are noise-free, as in (1), one can use the convex program commonly known as *Basis Pursuit* (BP):

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \ \|\mathbf{x}\|_1 \ \text{ s.t. } \Phi\mathbf{x} = \mathbf{y}. \tag{5}$$

Provided that Φ satisfies the RIP of order $2K$ with $\delta \leq \sqrt{2}-1$, BP is guaranteed to recover $\mathbf{x}$ exactly [16]. In the case where our measurements are contaminated with noise, as in (2), we can no longer expect to be able to recover $\mathbf{x}$ exactly in general. However, *Basis Pursuit De-Noising* (BPDN)

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \ \|\mathbf{x}\|_1 \ \text{ s.t. } \|\Phi\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon \tag{6}$$

will yield a recovered signal $\widehat{\mathbf{x}}$ that satisfies $\|\widehat{\mathbf{x}} - \mathbf{x}\|_2 \leq C_0\epsilon$ provided that $\|\mathbf{e}\|_2 \leq \epsilon$, where $C_0 > 1$ is a constant depending only on δ [10].

While convex optimization techniques like BP and BPDN are powerful methods for CS signal recovery, there also exist a variety of alternative algorithms that are used in practice and have comparable performance guarantees. Examples include iterative algorithms such as CoSaMP and iterative hard thresholding (IHT) [18, 19].

III. Exact signal recovery from measurements corrupted by sparse noise

Our goal is to design an algorithm that will recover both the signal and noise vectors by leveraging their sparsity. Towards this end, suppose that we acquire measurements of the form in (3) and that $\|\mathbf{x}\|_0 = K$ and $\|\mathbf{e}\|_0 = \kappa$. Note that the measurements can be expressed in terms of an $M \times (N + L)$ matrix multiplied by a $(K + \kappa)$-sparse vector:

$$\Phi\mathbf{x} + \Omega\mathbf{e} = [\Phi \ \Omega] \begin{bmatrix} \mathbf{x} \\ \mathbf{e} \end{bmatrix}. \tag{7}$$

We now introduce our reconstruction program, *Justice Pursuit* (JP):

$$\widehat{\mathbf{u}} = \arg\min_{\mathbf{u}} \|\mathbf{u}\|_1 \ \text{ s.t. } \ [\Phi \ \Omega]\mathbf{u} = \mathbf{y}, \tag{8}$$

where $\widehat{\mathbf{u}}$ is an intermediate $(N + L) \times 1$ recovery vector. The signal estimate $\widehat{\mathbf{x}}$ is obtained by selecting the first N elements of $\widehat{\mathbf{u}}$, i.e., $\widehat{x}_i = \widehat{u}_i, \ i = 1, \ldots, N$. Furthermore, an estimate of the noise vector $\widehat{\mathbf{e}}$ can be obtained by selecting the last L elements of $\widehat{\mathbf{u}}$, i.e., $\widehat{e}_i = \widehat{u}_{i+N}, \ i = 1, \ldots, L$. Note that one can adapt iterative algorithms such as CoSaMP and iterative hard thresholding (IHT) by simply replacing Φ with $[\Phi \ \Omega]$.

JP is essentially identical to a program proposed in [13, 14]. Note, however, that in [13, 14] the authors consider only Φ that are composed of a set of highly correlated training vectors and do not consider this program within the more traditional context of CS. Indeed, due to our differing assumptions on Φ, we can demonstrate stronger, nonasymptotic guarantees on the recovery of $\mathbf{x}$ and $\mathbf{e}$ provided by JP. The sparse noise model has also been considered in the context of CS

in [20], however the authors use a probabilistic approach for the analysis, a specialized measurement scheme, and propose a non-convex program with non-linear constraints for signal recovery, resulting in substantial differences from the results we present below. Note also that while [15] also considers the use of ℓ_1-minimization to mitigate sparse noise, this is in the context of error correction coding. In this framework the signal to be encoded is not necessarily sparse and $M > N$, resulting in a substantially different approach.

While JP is relatively intuitive, it is not clear that it will necessarily work. In particular, in order to analyze JP using standard methods, we must show that the matrix $[\Phi\ \Omega]$ satisfies the RIP. We now demonstrate that for random constructions of Φ with high probability $[\Phi\ \Omega]$ will satisfy the RIP for any Ω. To do so, we first establish the following lemma, which demonstrates that for any $\mathbf{u}$, if we draw Φ at random, then $\|[\Phi\ \Omega]\mathbf{u}\|_2$ is concentrated around $\|\mathbf{u}\|_2$.

Lemma 1. *Let Φ by an $M \times N$ matrix with elements ϕ_{ij} drawn i.i.d. according to $\mathcal{N}(0, 1/M)$ and let Ω be an $M \times L$ matrix with orthonormal columns. Furthermore, let $\mathbf{u} \in \mathbb{R}^{N+M}$ be an arbitrary vector with first N entries denoted by $\mathbf{x}$ and last L entries denoted by $\mathbf{e}$. Let $\delta \in (0, 1)$ be given. Then the matrix $[\Phi\ \Omega]$ satisfies*

$$\mathbb{E}\left(\|[\Phi\ \Omega]\mathbf{u}\|_2^2\right) = \|\mathbf{u}\|_2^2 \tag{9}$$

and

$$\mathbb{P}\left(\left|\|[\Phi\ \Omega]\mathbf{u}\|_2^2 - \|\mathbf{u}\|_2^2\right| \geq 2\delta\|\mathbf{u}\|_2^2\right) \leq 3e^{-M\delta^2/8}. \tag{10}$$

Proof: We first note that since $[\Phi\ \Omega]\mathbf{u} = \Phi\mathbf{x} + \Omega\mathbf{e}$,

$$\begin{aligned}\|[\Phi\ \Omega]\mathbf{u}\|_2^2 &= \|\Phi\mathbf{x} + \Omega\mathbf{e}\|_2^2 \\ &= (\Phi\mathbf{x} + \Omega\mathbf{e})^T(\Phi\mathbf{x} + \Omega\mathbf{e}) \\ &= \mathbf{x}^T\Phi^T\Phi\mathbf{x} + 2\mathbf{e}^T\Omega^T\Phi\mathbf{x} + \mathbf{e}^T\Omega^T\Omega\mathbf{e} \\ &= \|\Phi\mathbf{x}\|_2^2 + 2\mathbf{e}^T\Omega^T\Phi\mathbf{x} + \|\mathbf{e}\|_2^2. \end{aligned} \tag{11}$$

Since the entries ϕ_{ij} are i.i.d. according to $\mathcal{N}(0, 1/M)$, it is straightforward to show that $\mathbb{E}\left(\|\Phi\mathbf{x}\|_2^2\right) = \|\mathbf{x}\|_2^2$ (see, for example, [21]). Similarly, one can also show that $2\mathbf{e}^T\Omega^T\Phi\mathbf{x} \sim \mathcal{N}\left(0, 4\|\mathbf{x}\|_2^2\|\Omega\mathbf{e}\|_2^2/M\right)$, since the elements of $\Phi\mathbf{x}$ are distributed as zero mean Gaussian variables with variance $\|\mathbf{x}\|_2^2/M$. Thus, from (11) we have that

$$\mathbb{E}\left(\|[\Phi\ \Omega]\mathbf{u}\|_2^2\right) = \|\mathbf{x}\|_2^2 + \|\mathbf{e}\|_2^2,$$

and since $\|\mathbf{u}\|_2^2 = \|\mathbf{x}\|_2^2 + \|\mathbf{e}\|_2^2$, this establishes (9).

We now turn to (10). Using the arguments in [21], one can show that

$$\mathbb{P}\left(\left|\|\Phi\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2\right| \geq \delta\|\mathbf{x}\|_2^2\right) \leq 2e^{-M\delta^2/8}. \tag{12}$$

As noted above, $2\mathbf{e}^T\Omega^T\Phi\mathbf{x} \sim \mathcal{N}\left(0, 4\|\mathbf{x}\|_2^2\|\Omega\mathbf{e}\|_2^2/M\right)$. Note that since the columns of Ω are orthonormal, $\|\Omega\mathbf{e}\|_2^2 = \|\mathbf{e}\|_2^2$. Hence, we have that

$$\begin{aligned}\mathbb{P}\left(\left|2\mathbf{e}^T\Omega^T\Phi\mathbf{x}\right| \geq \delta\|\mathbf{x}\|_2\|\mathbf{e}\|_2\right) &= 2Q\left(\frac{\delta\|\mathbf{x}\|_2\|\mathbf{e}\|_2}{2\|\mathbf{x}\|_2\|\mathbf{e}\|_2/\sqrt{M}}\right) \\ &= 2Q(\sqrt{M}\delta/2),\end{aligned}$$

where $Q(\cdot)$ denotes the tail integral of the standard Gaussian distribution. From (13.48) of [22] we have that

$$Q(z) \leq \frac{1}{2}e^{-z^2/2}$$

and thus we obtain

$$\mathbb{P}\left(\left|2\mathbf{e}^T\Omega^T\Phi\mathbf{x}\right| \geq \delta\|\mathbf{x}\|_2\|\mathbf{e}\|_2\right) \leq e^{-M\delta^2/8}. \tag{13}$$

Thus, combining (12) and (13) we obtain that with probability at least $1 - 3e^{-M\delta^2/8}$ we have that both

$$(1-\delta)\|\mathbf{x}\|_2^2 \leq \|\Phi\mathbf{x}\|_2^2 \leq (1+\delta)\|\mathbf{x}\|_2^2 \tag{14}$$

and

$$-\delta\|\mathbf{x}\|_2\|\mathbf{e}\|_2 \leq 2\mathbf{e}^T\Omega^T\Phi\mathbf{x} \leq \delta\|\mathbf{x}\|_2\|\mathbf{e}\|_2. \tag{15}$$

Using (11), we can combine (14) and (15) to obtain

$$\begin{aligned}\|[\Phi\ \Omega]\mathbf{u}\|_2^2 &\leq (1+\delta)\|\mathbf{x}\|_2^2 + \delta\|\mathbf{x}\|_2\|\mathbf{e}\|_2 + \|\mathbf{e}\|_2^2 \\ &\leq (1+\delta)\left(\|\mathbf{x}\|_2^2 + \|\mathbf{e}\|_2^2\right) + \delta\|\mathbf{x}\|_2\|\mathbf{e}\|_2 \\ &\leq (1+\delta)\|\mathbf{u}\|_2^2 + \delta\|\mathbf{u}\|_2^2 \\ &= (1+2\delta)\|\mathbf{u}\|_2^2,\end{aligned}$$

where the last inequality follows from the fact that $\|\mathbf{x}\|_2\|\mathbf{e}\|_2 \leq \|\mathbf{u}\|_2\|\mathbf{u}\|_2$. Similarly, we also have that

$$\|[\Phi\ \Omega]\mathbf{u}\|_2^2 \geq (1-2\delta)\|\mathbf{u}\|_2^2,$$

which establishes (10). ∎

We note that the bound on the cross-term $\left|2\mathbf{e}^T\Omega^T\Phi\mathbf{x}\right|$ given by (13) is the only part of the proof that relies on the Gaussianity of ϕ_{ij}; similar results can be shown for the more general class of sub-Gaussian matrices.

Using Lemma 1, we now demonstrate that the matrix $[\Phi\ \Omega]$ satisfies the RIP provided that M is sufficiently large. This theorem follows immediately from Lemma 1 by using a proof identical to that given in [17], so we omit the proof for the sake of brevity.

Theorem 1. *Let Φ by an $M \times N$ matrix with elements $\phi_{i,j}$ drawn according to $\mathcal{N}(0, 1/M)$ and let Ω be an $M \times L$ matrix with orthonormal columns. If*

$$M \geq C_1(K+\kappa)\log\left(\frac{N+L}{K+\kappa}\right) \tag{16}$$

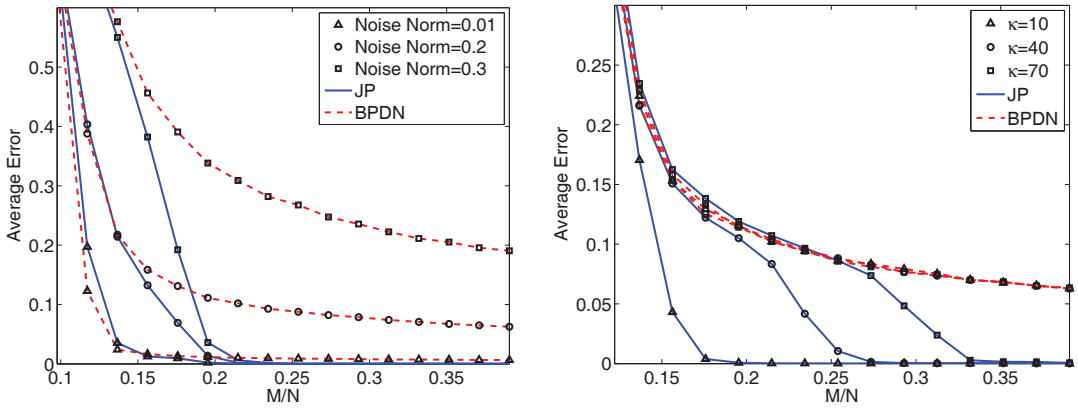
then $[\Phi\ \Omega]$ satisfies the RIP of order $(K+\kappa)$ with probability exceeding $1 - 3e^{-C_2M}$, where C_1 and C_2 are constants that depends only on the desired RIP constant δ.

Theorem 1 implies that when both $\mathbf{x}$ and $\mathbf{e}$ are sparse, JP recovers both $\mathbf{x}$ and $\mathbf{e}$ exactly. Thus, even if $\|\mathbf{e}\|_2$ is unbounded, in this setting JP achieves optimal performance.

In the case where $\mathbf{e}$ contains additional sources of noise that are not sparse, e.g., additive white Gaussian noise (AWGN) or quantization error in addition to hum, but has norm bounded by ϵ, we propose an algorithm we dub *Justice Pursuit De-Noising* (JPDN):

$$\widehat{\mathbf{u}} = \arg\min_{\mathbf{u}} \|\mathbf{u}\|_1 \quad \text{s.t.} \quad \|[\Phi\ \Omega]\mathbf{u} - \mathbf{y}\|_2 < \epsilon. \tag{17}$$

The performance guarantees of JPDN are analogous to those for BPDN, but a detailed analysis of this algorithm is beyond the scope of this paper.



(a) Fixed $\kappa = 10$ (b) Fixed $\|\mathbf{e}\|_2 = 0.1$

Fig. 1. *Comparison of average reconstruction error $\|\mathbf{x} - \widehat{\mathbf{x}}\|_2$ between JP (solid lines) and BPDN (dashed lines). All trials used parameters $N = 2048$ and $K = 10$. (a) Comparison with fixed $\kappa = 10$ and noise norms $\|\mathbf{e}\|_2 = 0.01,\ 0.2$, and 0.03, depicted by the triangle, circle, and square-marked lines, respectively. This plot demonstrates that while BPDN never achieves exact reconstruction, JP does. (b) Comparison with fixed noise norm $\|\mathbf{e}\|_2 = 0.1$ and $\kappa = 10,\ 40$, and 70, depicted by the triangle, circle, and square-marked lines, respectively. This plot demonstrates that JP performs similarly to BPDN until M is large enough to reconstruct κ noise entries.*

IV. Simulations

A. Average performance comparison

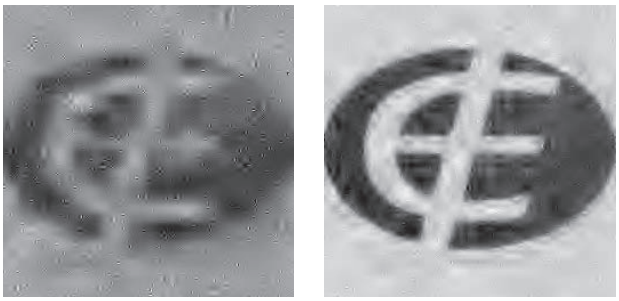
In Figure 1, we compare the average reconstruction error of JP (solid lines) against the average error of BPDN (dashed lines). We perform two experiments, each with parameters $N = 2048$, $K = 10$, and $\|\mathbf{x}\|_2 = 1$, with $M/N \in [0.1, 0.4]$, and record the average error $\|\mathbf{x} - \widehat{\mathbf{x}}\|_2$ over 100 trials.

In the first experiment, depicted in Figure 1(a), we fix $\|\mathbf{e}\|_0 = \kappa = 10$ and vary $\|\mathbf{e}\|_2$. We use the values $\|\mathbf{e}\|_2 = 0.01,\ 0.2$, and 0.03, depicted by the triangle, circle, and square-marked lines, respectively. We observe that the reconstruction error for BPDN does not decay to zero no matter how large we set M. Most representative of this is the dashed line marked by triangles denoting the $\|\mathbf{e}\|_2 = 0.01$ case. As M/N increases, this line reaches a minimum value greater than zero and does not decay further. In contrast, JP reaches exact recovery in all tests.

In the second experiment, depicted in Figure 1(b), we fix $\|\mathbf{e}\|_2 = 0.1$ and vary κ. We use the values $\kappa = 10,\ 40$, and 70, depicted by the triangle, circle, and square-marked lines, respectively. Again, the performance of BPDN does not decay to zero, and furthermore, the performance does not vary with κ on average. As expected the error of JP goes to zero and requires more measurements to do so as κ increases.

B. Reconstruction with hum

In this experiment we study the reconstruction performance from measurements corrupted by hum, meaning that we add a 60Hz sinusoid to the measurements. We use a 256×256 test image that is compressible in the wavelet domain, set the measurement ratio to $M/N = 0.2$, and set the measurement signal-to-noise ratio (SNR) to 9.3dB, where measurement SNR in dB is defined as $10\log_{10}(\|\Phi\mathbf{x}\|_2^2/\|\mathbf{e}\|_2^2)$. We recover using BPDN with $\epsilon = \|\mathbf{e}\|_2$ and using JP with the Fourier basis for Ω. Note that rather than choosing the entire Fourier basis, a matrix containing the 60Hz tone and its harmonics can be chosen to reduce the number of required measurements.



(a) (b)

Fig. 2. *Reconstruction of an image from CS measurements that have been distorted by an additive 60Hz sinusoid (hum). The experimental parameters are $M/N = 0.2$ and measurement SNR $= 9.3$dB. (a) Reconstruction using BPDN. (b) Reconstruction using JP. Spurious artifacts due to noise are present in the image in (a) but not in (b). Significant edge detail is lost in (a) but recovered in (b).*

Figure 2(a) depicts the reconstruction from BPDN and Figure 2(b) depicts the reconstruction from JP. Both images contain compression artifacts, such as "ringing," since the signal is not strictly sparse. However, the BPDN reconstruction contains spurious artifacts, due not to compression but to noise, while the JP reconstruction does not. Furthermore, significant edge detail is lost in the BPDN reconstruction.

C. Measurement denoising

In this experiment[2] we use our algorithm to denoise measurements $\mathbf{y}$ that have been acquired by the single-pixel camera [6]. The image dimensions are 256×256 and $M/N = 0.1$. The denoising procedure is as follows. First we reconstruct the image using JP with the Fourier basis for Ω. Second, because

[2]Thanks to Professor Kevin Kelly for providing the data used in this experiment. More information is available at dsp.rice.edu/cscamera.

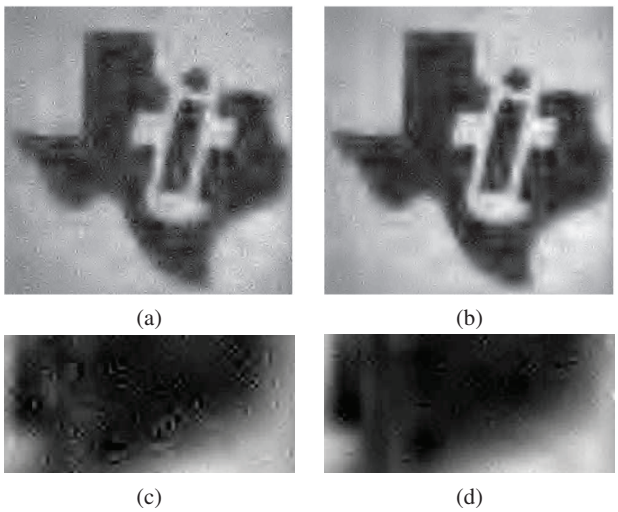


Fig. 3. *Reconstruction from CS camera data. (a) Reconstruction from CS camera measurements. (b) Reconstruction from denoised CS camera measurements. (c) and (d) depict zoomed sections of (a) and (b), respectively. Noise artifacts are removed without further smoothing of the underlying image.*

the measurement noise is not strictly sparse, we select the 15 largest terms from $\widehat{\mathbf{e}}$, denoted as $\widehat{\mathbf{e}}'$, and subtract their contribution from the original measurements, i.e.,

$$\mathbf{y}' = \mathbf{y} - \Omega\widehat{\mathbf{e}}'.$$

Third, reconstruction from $\mathbf{y}'$ is performed with BPDN using the parameter $\epsilon = 0.3$. To compare, we also reconstruct the image from the original measurements $\mathbf{y}$ using BPDN with the same ϵ. In general, this procedure can be performed iteratively, selecting several spikes from $\widehat{\mathbf{e}}$ at each iteration and subtracting their contribution from from the measurements.

Figure 3(a) depicts the reconstruction from $\mathbf{y}$ and Figure 3(b) depicts the reconstruction from $\mathbf{y}'$, and Figures 3(c) and 3(d) show a zoomed section of each, respectively. The reconstruction from the original measurements contains significantly more spurious artifacts, while the reconstruction from denoised measurements removes these artifacts without further smoothing of the underlying image.

V. Discussion

In this paper we have extended the ability of CS to reconstruct from inaccurate measurements. This was accomplished by introducing a new reconstruction algorithm, Justice Pursuit (JP), for sparse signals from CS measurements with sparse noise. We have proven that JP can recover sparse signals and sparse noise *exactly* and derived the number of measurements required to do so. JP requires no user-defined parameters and can be easily implemented by extending existing algorithms. Furthermore, we have demonstrated in simulation that JP can achieve exact recovery when conventional algorithms cannot and that JP can improve reconstruction performance from real data.

There are many topics that have not been fully explored in this paper. For instance, the noise could be compressible rather than strictly sparse, or could consist of low energy noise on all measurements in addition to the sparse noise. For example, measurements may be subject to both shot noise and quantization errors simultaneously. Analysis of JPDN and the related iterative methods explored in Section IV-C as applied to this problem remain a topic of ongoing work. Additionally, models can be employed to exploit additional noise structure and reduce the number of required measurements, or recover the signal with higher accuracy.

References

[1] E. Candès, "Compressive sampling," in *Proc. Int. Congress of Mathematics*, Madrid, Spain, Aug. 2006, pp. 1433–1452.

[2] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 6, no. 4, pp. 1289–1306, Apr. 2006.

[3] J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk, "Beyond Nyquist: Efficient sampling of sparse, bandlimited signals," *to appear in IEEE Trans. Inform. Theory*, 2009.

[4] J. Romberg, "Compressive sensing by random convolution," *to appear in SIAM J. Imaging Sciences*, 2009.

[5] J. A. Tropp, M. Wakin, M. F. Duarte, D. Baron, and R. G. Baraniuk, "Random filters for compressive sampling and reconstruction," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Toulouse, France, May 2006.

[6] M. F. Duarte, M. Davenport, D. Takhar, J. N. Laska, T. Sun, K. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 83–91, Mar. 2008.

[7] R. Robucci, L. K. Chiu, J. Gray, J. Romberg, P. Hasler, and D. Anderson, "Compressive sensing on a CMOS separable transform image sensor," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, Las Vegas, Nevada, Apr. 2008.

[8] R. Marcia, Z. Harmany, and R. Willett, "Compressive coded aperture imaging," in *Proc. IS&T/SPIE Symposium on Electronic Imaging: Computational Imaging VII*, vol. 7236, San Jose, California, Jan. 2009.

[9] D. Baron, M. F. Duarte, M. B. Wakin, S. Sarvotham, and R. G. Baraniuk, "Distributed compressive sensing," *Preprint*, 2005.

[10] E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[11] P. Wojtaszczyk, "Stability and instance optimality for Gaussian measurements in compressed sensing," *to appear in Found. Comp. Mathematics*, 2009.

[12] J. N. Laska, P. Boufounos, M. A. Davenport, and R. G. Baraniuk, "Democracy in action: Quantization, saturation, and compressive sensing," *Preprint*, 2009.

[13] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[14] J. Wright and Y. Ma, "Dense error correction via ℓ_1 minimization," *Preprint*, 2009.

[15] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[16] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes rendus de l'Académie des Sciences, Série I*, vol. 346, no. 9-10, pp. 589–592, May 2008.

[17] R. G. Baraniuk, M. A. Davenport, R. DeVore, and M. B. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[18] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comp. Harmonic Anal.*, vol. 26, no. 3, pp. 301–321, May 2009.

[19] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressive sensing," *Appl. Comp. Harmonic Anal.*, vol. 27, no. 3, pp. 265–274, Nov. 2009.

[20] R. E. Carrillo, K. E. Barner, and T. C. Aysal, "Robust sampling and reconstruction methods for sparse signals in the precense of impulsive noise," *Preprint*, 2009.

[21] S. Dasgupta and A. Gupta, "An elementary proof of the Johnson-Lindenstrauss lemma," *Tech. Report 99-006, U.C. Berkeley*, Mar. 1999.

[22] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions, Volume 1*. Wiley, 1994.