SIGNAL RECONSTRUCTION FROM NOISY RANDOMIZED PROJECTIONS WITH APPLICATIONS TO WIRELESS SENSING

Jarvis Haupt and Robert Nowak

Department of Electrical and Computer Engineering
University of Wisconsin-Madison

ABSTRACT

Recent results show that a relatively small number of random projections of a signal can contain most of its salient information. It follows that if a signal is compressible in some orthonormal basis, then a very accurate reconstruction can be obtained from projections of the signal onto random basis elements. We extend this type of result to show that compressible signals can be accurately recovered from random projections contaminated with noise, in many cases much more accurately than is possible using an equivalent number of conventional point samples. We also investigate the application of this new sampling and reconstruction procedure to remote wireless sensing.

1. INRODUCTION

This paper consists of two main contributions. First, we develop theory for reconstructing signals from random projections and show that this approach has significant advantages over more conventional sampling schemes. Second, we apply our new results to the problem of wireless sensing.

Recent theory informs us that a relatively small number of random projections of a signal can contain most of its salient information [1, 2, 3]. For example, the groundbreaking work in [1] has shown that k random Fourier projections contain enough information to reconstruct piecewise smooth signals at a distortion level nearly equivalent to that attainable from k optimally selected projections. More generally, assume that a signal $f \in R^n$ is well approximated in some orthonormal basis in the following sense. Let $f^{(m)}$ denote the best m-term approximation of f in terms of this basis and suppose that the *average squared error* behaves like

$$\frac{\|f - f^{(m)}\|^2}{n} \equiv \frac{1}{n}\sum_{i=1}^{n}(f_i - f_i^{(m)})^2 = O(m^{-2\alpha})$$

for some $\alpha \geq 1$. The parameter α governs the degree to which f is compressible with respect to the basis. In a noiseless setting, it can be shown that an approximation of such a signal can be recovered from k random projections with an average squared error that is $O\left((k/\log n)^{-2\alpha}\right)$, nearly as good as the best k-term approximation error; see [1, 2, 3]. On the other hand, we show in this paper that one can accurately reconstruct such a signal from k random projections (corrupted with noise) to yield an average squared reconstruction error that is $O\left((k/\log n)^{\frac{-2\alpha}{2\alpha+1}}\right)$. For sparse signals a stronger result is obtained - the expected average squared reconstruction error is $O\left((k/\log n)^{-1}\right)$.

As a motivation for the use of random projections, consider the following simple example. Suppose the signal f^* is a vector of length n with one nonzero entry of amplitude $\sqrt{n}$ such that $\|f^*\|^2/n = 1$. First, consider random spatial point sampling where observations are noise-free (i.e., each sample is of the form $y_j = f^*(t_j)$, where t_j is selected uniformly at random from the set $\{1, \ldots, n\}$). The squared reconstruction error is 0 if the spike is located and 1 otherwise, and the probability of not finding the spike in k trials is $(1 - 1/n)^k$, giving an average squared error of $(1 - 1/n)^k \cdot 1 + \left[1 - (1 - 1/n)^k\right] \cdot 0 = (1 - 1/n)^k$. If n is large, we can approximate this by $(1 - 1/n)^k \approx e^{-k/n}$, which is approximately equal to 1 if $k \ll n$. On the other hand, we present a sampling and reconstruction scheme in this paper based on k random projections (corrupted with noise), which yields an average squared reconstruction error that is $O(k/\log n)^{-1}$. This shows that even given the advantage of being noiseless, the reconstruction error from conventional point sampling can be far greater than that resulting from random projections.

We will investigate an application of these results to wireless sensing. Consider an *ensemble* of n wireless sensors that can be jointly synchronized for phase-coherent communications with a remote destination. Suppose that each sensor multiplies its measurement by a random number (known only to itself and the destination) and then all n sensors transmit their (randomly scaled) measurements to the remote destination. Furthermore, suppose these data are transmitted via uncoded amplitude modulation in a phase-coherent fashion, so that the received signal (summation of

This research was supported in part by the NSF under grants CCR-0113385 and CCR-0310889, and by the ONR under grant N00014-00-1-0966.

the n transmitted signals) constitutes a random projection of the form described above. Repeating this process k times provides a very efficient transmission of k random projections of the sensor readings to the destination. Moreover, the received signal power is a factor n times the total transmit power (ignoring a constant depending on the distance of the destination from the sensors) due to the phase-coherent nature of the transmissions. A remarkable consequence of this fact is that the total transmit power required in order to reconstruct the sensor values at the destination to within an accuracy of $O\left((k/\log^2 n)^{\frac{-2\alpha}{2\alpha+1}}\right)$ decays rapidly to zero as the density of sensors increases.

This paper is organized as follows. In Section 2 we state the basic problem and main theoretical results of the paper. In Section 3 we derive reconstruction error rates for signals that are compressible in terms of a certain orthonormal basis. In Section 4 we describe the optimization problem in detail and list some possible solution strategies. In Section 5 we investigate the performance of our reconstruction method. In Section 6 we explain the benefits of this scheme in the context of remote wireless sensing, and we make concluding remarks in Section 7.

2. MAIN RESULTS

Consider a vector $f^* = [f_1^* \; f_2^* \dots f_n^*]^T \in R^n$ and assume that $\sum_{i=1}^n (f_i^*)^2 \equiv \|f^*\|^2 \leq nB^2$ for a known constant $B > 0$. Suppose that we are able to make k measurements of f^* in the form of noisy random projections. Specifically, let $\Phi = \{\phi_{i,j}\}$ be an $n \times k$ array of bounded, i.i.d. zero-mean random variables of variance $E[\phi_{i,j}^2] = 1/n$. Our observations take the form

$$y_j = \sum_{i=1}^n \phi_{i,j} f_i^* + w_j, \;\; j = 1, \dots, k \qquad (1)$$

where $w = \{w_j\}$ are i.i.d. zero-mean random variables, independent of $\{\phi_{i,j}\}$, with variance σ^2. Our goal is to recover an estimate of f^* from these observations.

Define the risk of a candidate reconstruction f to be

$$\begin{aligned} R(f) &= E\left[\frac{1}{k}\sum_{i=1}^k \left(y_j - \sum_{i=1}^n \phi_{i,j} f_i\right)^2\right] \\ &= \frac{\|f^* - f\|^2}{n} + \sigma^2. \end{aligned}$$

The expectation here (and throughout the rest of the paper) is with respect to Φ and w, the norm is the Euclidean distance, and we have used the fact that $\{\phi_{i,j}\}$ and $\{w_j\}$ are independent random variables and $E[\phi_{i,j}^2] = 1/n$. Assume that both Φ and $\{y_j\}$ are available. Then we can compute the empirical risk

$$\widehat{R}(f) = \frac{1}{k}\sum_{j=1}^k \left(y_j - \sum_{i=1}^n \phi_{i,j} f_i\right)^2$$

which is an unbiased estimate of $R(f)$. Our goal is to use the empirical risk to obtain an estimator $\widehat{f}$ of f^* and to bound the resulting error $E[\|\widehat{f} - f^*\|^2]$.

Our estimator is based on a complexity-regularized empirical risk minimization and we use the Craig-Bernstein concentration inequality [4] to control the estimation error of the reconstruction process. The Craig-Bernstein inequality states that for random variables U_j the event $\mathcal{E}$, defined as

$$\mathcal{E} \equiv \sum_{j=1}^k \frac{(U_j - E[U_j])}{k} \geq \frac{t}{k\epsilon} + \frac{\epsilon \, k \, var\left(\sum_{j=1}^k \frac{U_j}{k}\right)}{2(1-\zeta)}$$

satisfies

$$P(\mathcal{E}) \leq e^{-t}$$

for $0 < \epsilon h \leq \zeta < 1$ and $t > 0$, provided the random variables U_j each satisfy the moment condition

$$E\left[|U_j - E[U_j]|^k\right] \;\leq\; \frac{k! \, var(U_j)\, h^{k-2}}{2}$$

for some $h > 0$ and all integers $k \geq 2$. For our purposes, $U_j = \left[(y_j - \phi_j^T f^*)^2 - (y_j - \phi_j^T f)^2\right]$, the excess risk for a given projection. The moment condition obviously depends on the nature of Φ and w. In this paper we focus on (normalized) Rademacher projections, in which case each $\phi_{i,j}$ is $\pm 1/\sqrt{n}$ with equal probability, and assume that w is a sequence of zero-mean Gaussian noises. Generalizations to other random projections and noise models may be possible; this would only require one to verify the moment conditions required by the Craig-Bernstein inequality.

Suppose that we have a countable collection $\mathcal{F}$ of candidate reconstruction functions, such that each $f \in \mathcal{F}$ satisfies $\|f\|^2 \leq nB^2$. Further, assume that assigned to each $f \in \mathcal{F}$ is a non-negative number $c(f)$ such that $\sum_{f\in\mathcal{F}} 2^{-c(f)} \leq 1$. Select a reconstruction according to the complexity-regularized empirical risk minimization

$$\widehat{f}_k \;=\; \arg\min_{f\in\mathcal{F}} \left\{\widehat{R}(f) + \frac{c(f)\log 2}{k\epsilon}\right\}$$

where $\epsilon > 0$ is a constant that depends on the function bound B and the noise variance σ^2. Then we have the following oracle inequality.

Theorem 1 *There exists a constant $C_1 > 0$ such that*

$$E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] \leq C_1 \min_{f\in\mathcal{F}} \left\{\frac{\|f - f^*\|^2}{n} + \frac{c(f)\log 2 + 4}{k\epsilon}\right\}. \qquad (2)$$

Our derivation of the oracle bound (2) is similar in nature to the procedure used in [5], but a main contribution of our work is the verification of the moment condition for certain unbounded random variables. Detailed derivation of the oracle bound, including a theoretical framework under which unbounded noise can be handled in the context of the moment condition, can be found in [6].

Suppose f^* is compressible with respect to a certain orthonormal basis (e.g., wavelet). We can obtain explicit bounds on the error in terms of the number of random projections k and the degree to which f^* is compressible. Let $f^{(m)}$ denote the best m-term approximation of f^* in the basis. That is, if f^* has a representation $f^* = \sum_{i=1}^n \theta_i \psi_i$ in the basis $\{\psi_i\}$, then $f^{(m)} = \sum_{i=1}^m \theta_{(i)} \psi_{(i)}$, where coefficients and basis functions are re-ordered such that $|\theta_{(1)}| \geq |\theta_{(2)}| \geq \cdots \geq |\theta_{(n)}|$. Assume that the *average squared error* $\|f^* - f^{(m)}\|^2/n \equiv \frac{1}{n}\sum_{i=1}^n (f_i^* - f_i^{(m)})^2$ behaves like

$$\frac{\|f^* - f^{(m)}\|^2}{n} = O(m^{-2\alpha})$$

for some $\alpha \geq 1$. Now by choosing $\mathcal{F}$ to be a suitably quantized collection of functions (represented in terms of the basis $\{\psi_i\}$), we have the following error bound.

Theorem 2 *If*

$$c(f) = 2\log(n) \times \{\text{\# non-zero coefficients of } f\}$$

then there exists a constant $C_2 > 0$ such that

$$E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] \leq C_2 \left(\frac{k}{\log n}\right)^{\frac{-2\alpha}{2\alpha+1}}. \qquad (3)$$

Note that the exponent $-2\alpha/(2\alpha+1)$ is the "usual" exponent governing the rate of convergence in nonparametric function estimation. A stronger result is obtained if the signal is sparse, as stated in the following Corollary.

Corollary 1 *Suppose that f^* has at most m nonzero coefficients. Then there exists a constant $C_2' > 0$ such that*

$$E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] \leq C_2' \left(\frac{k}{m\log n}\right)^{-1}. \qquad (4)$$

Similar results hold if the signal is additionally contaminated with noise prior to the random projection process.

Corollary 2 *Suppose our observation model takes the form*

$$y_j = \sum_{i=1}^n \phi_{i,j}\,(f_i^* + \eta_i) + w_j, \quad j = 1, \ldots, k$$

where $\{\eta_i\}$ are i.i.d. zero-mean Gaussian random variables with variance σ_s^2 that are independent of $\{\phi_{i,j}\}$ and $\{w_j\}$. Then Theorems 1 and 2 and Corollary 1 hold with slightly different constants C_1, C_2, C_2', and ϵ.

This result follows from the fact that the projected noise term, $\sum_{i=1}^n \phi_{i,j}\eta_i$, is equivalent in distribution to a Gaussian random variable given our choice of $\{\phi_{i,j}\}$ and $\{\eta_i\}$. Further, this projected noise term is independent of $\{\phi_{i,j}\}$ so it can be absorbed into the noise terms $\{w_j\}$. This results in a setup identical in form to the original observation model given in (1) where the only change is in the variance of the noise term. For a complete derivation of this result, see [6].

3. ERROR BOUNDS FOR COMPRESSIBLE SIGNALS

Suppose that f^* is compressible in a certain orthonormal basis $\{\psi_i\}_{i=1}^n$. Specifically, let $f^{(m)}$ denote the best m-term approximation of f^* in terms of $\{\psi_i\}$, and assume that the error of the approximation behaves like

$$\frac{\|f - f^{(m)}\|^2}{n} \leq C_3\, m^{-2\alpha}$$

for some $\alpha \geq 1$ and a constant $C_3 > 0$.

Let us use the basis $\{\psi_i\}$ for the reconstruction process. Any vector $f \in R^n$ can be expressed in terms of the basis $\{\psi_i\}$ as $f = \sum_{i=1}^n \theta_i\psi_i$, where $\theta = \{\theta_i\}$ are the coefficients of f in this basis. Let T denote the transform that maps coefficients to functions, so that $f = T\theta$. Let Θ denote the collection of all coefficient vectors satisfying $\|T\theta\|^2 \leq nB^2$ and whose components are uniformly quantized in magnitude to n levels. Let $\mathcal{F}$ denote the set of all functions of the form $f = T\theta$, $\theta \in \Theta$. Furthermore, if $f = T\theta$, then let $c(f) \equiv c(\theta) = 2\log(n)\sum_{i=1}^n I_{\theta_i \neq 0} = 2\log n\|\theta\|_0$. It is easily verified that $\sum_{f\in\mathcal{F}} 2^{-c(f)} \leq 1$ by noting that each $\theta \in \Theta$ can be uniquely encoded via a prefix code consisting of $2\log n$ bits per non-zero coefficient ($\log n$ bits for the locations and $\log n$ bits for the quantized values) in which case the codelengths $c(f)$ must satisfy the Kraft inequality [7].

The oracle inequality

$$E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] \leq C \min_{f\in\mathcal{F}}\left\{\frac{\|f - f^*\|^2}{n} + \frac{c(f)\log 2 + 4}{k\epsilon}\right\}$$

can be written in terms of the new class of candidate reconstructions as

$$E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] \leq C \min_{\theta\in\Theta}\left\{\frac{\|\theta - \theta^*\|^2}{n} + \frac{c(\theta)\log 2 + 4}{k\epsilon}\right\}$$

where $f^* = T\theta^*$. For each integer $m \geq 1$, let $\theta^{(m)}$ denote the coefficients corresponding to the best m-term approximation of f^* and let $\theta_q^{(m)}$ denote the nearest element in Θ. The maximum possible dynamic range for the coefficient magnitudes, $\pm\sqrt{n}B$, is quantized to n levels, giving $\|\theta_q^{(m)} - \theta^{(m)}\|^2 \leq C_4$, for a constant $C_4 > 0$. Thus, we have

$$\begin{aligned}
\|\theta_q^{(m)} - \theta^*\|^2 &= \|\theta_q^{(m)} - \theta^{(m)} - \theta^{(m)} - \theta^*\|^2 \\
&\leq \|\theta_q^{(m)} - \theta^{(m)}\|^2 + \\
&\quad 2\|\theta_q^{(m)} - \theta^{(m)}\| \cdot \|\theta^{(m)} - \theta^*\| + \\
&\quad \|\theta^{(m)} - \theta^*\|^2 \\
&\leq C_4 + 2m^{-\alpha}\sqrt{nC_3C_4} + C_3 n m^{-2\alpha}.
\end{aligned}$$

Now inserting $\theta_q^{(m)}$ in place of θ in the oracle bound, noting that $c(\theta_q^{(m)}) = 2m\log n$, and minimizing the upper bound with respect to the choice of m, we obtain

$$E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] \leq C_2\left(\frac{k}{\log n}\right)^{\frac{-2\alpha}{2\alpha+1}}$$

for a constant $C_2 > 0$.

Suppose now that f^* has only m nonzero coefficients. In this case, we have

$$\begin{aligned}
\|\theta_q^{(m)} - \theta^*\|^2 &= \|\theta_q^{(m)} - \theta^{(m)} - \theta^{(m)} - \theta^*\|^2 \\
&\leq \|\theta_q^{(m)} - \theta^{(m)}\|^2 + \\
&\quad 2\|\theta_q^{(m)} - \theta^{(m)}\| \cdot \|\theta^{(m)} - \theta^*\| + \\
&\quad \|\theta^{(m)} - \theta^*\|^2 \\
&\leq C_4
\end{aligned}$$

since $\|\theta^{(m)} - \theta^*\| = 0$. Now the penalty term dominates in the oracle bound and we obtain

$$E\left[\frac{\|\widehat{f}_k - f^*\|^2}{n}\right] \leq C_2'\left(\frac{k}{m\log n}\right)^{-1}$$

for a constant $C_2' > 0$.

4. OPTIMIZATION SCHEME

Let us assume that we wish to reconstruct our signal in terms of the basis $\{\psi_i\}$. The reconstruction

$$\widehat{f}_k = \arg\min_{f\in\mathcal{F}}\left\{\widehat{R}(f) + \frac{c(f)\log(2)}{k\epsilon}\right\}$$

is equivalent to $\widehat{f}_k = T\widehat{\theta}_k$ where

$$\widehat{\theta}_k = \arg\min_{\theta\in\Theta}\left\{\widehat{R}(T\theta) + \frac{c(\theta)\log(2)}{k\epsilon}\right\}.$$

Thus, the optimization problem can then be written as

$$\widehat{\theta}_k = \arg\min_{\theta\in\Theta}\left\{\|y - PT\theta\|^2 + \frac{2\log(2)\log(n)}{\epsilon}\|\theta\|_0\right\}$$

where $P = \Phi^T$, the transpose of the $n \times k$ projection matrix Φ, y is a column vector of the k observations, and $\|\theta\|_0 = \sum_{i=1}^n I_{\{\theta_i \neq 0\}}$.

The optimization problem above can be tackled using a number of existing methods, including the Bound Optimization technique proposed in in [8] and applied in in this context in [6], or the Coordinate Descent method derived in [9]. Both methods cited above result in a simple iterative scheme where the optimization step involves a linear update followed by a coordinate-wise (diagonal) thresholding operation.

5. RESULTS

We tested the performance of our method on two examples. In each case, the optimization problem was solved using the Coordinate Descent method in [9]. If we define $\tau = \log(2)\log(n)/\epsilon$, $B = (PT)^T(PT)$, $b = (PT)^T y$, and $v = B\theta$, then the update rule for each component of $\widehat{\theta}_k$ can be written as:

$$\widehat{\theta}_{k,i}^{new} = \begin{cases} \widehat{\theta}_{k,i}^{old} + \frac{b_i - v_i}{B_{i,i}} & \text{if } |\widehat{\theta}_{k,i}^{old} + \frac{b_i - v_i}{B_{i,i}}| \geq \sqrt{\frac{2\tau}{B_{i,i}}} \\ 0 & \text{otherwise} \end{cases}$$

for $i = 1\ldots n$. Each such optimization step must be accompanied by an update of the vector v, computed using the new value of $\widehat{\theta}$. The algorithm terminates when the entries uniformly satisfy $|\widehat{\theta}_{k,i}^{new} - \widehat{\theta}_{k,i}^{old}| \leq \delta$, for a small positive tolerance δ.

For the first example, we consider a one-dimensional signal being measured by equally-spaced sensors. The signal to be determined is the Stanford WaveLab Blocks signal of length $2^{10} = 1024$, and each sensor measurement is corrupted by zero-mean Gaussian noise.

Figure 1 shows the original noiseless signal, the noisy signal as measured by the sensors, and the reconstruction obtained for one realization of random basis elements and noise. Most of the dominant signal features have been captured in and extracted from the limited number of observations. Also notice that the reconstruction shows no noise artifacts like those seen in the actual sensor measurements. The per-pixel squared reconstruction error for this example is 0.24. Averaged over 50 trials, the per-pixel squared reconstruction error is 0.34 with standard deviation 0.08. The threshold was set at $\tau = 1.17$.

For the second example, we consider the problem of sensing a distributed field (e.g., star map, presence or absence of a chemical agent) using a collection of n sensors

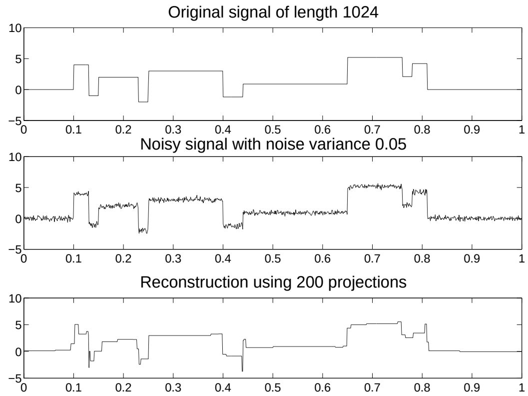

Fig. 1. Simulation results for the Blocks signal of length 1024. Blocks signal (top), noisy sensor readings (middle), reconstructed signal using 200 random projections of noisy sensor readings (bottom). The sensor noise variance is $\sigma^2 = 0.05$.

distributed uniformly over a two-dimensional region of interest. Such signals might be sparse in the spatial measurement (pixel) basis.

Figure 2 shows the original noisy field along side the reconstruction obtained from a limited number of random projections. Besides being significantly less noisy, the reconstruction again exhibits all of the significant location information found in the original signal. The per-pixel squared reconstruction error for this example is 0.003. Averaged over 50 trials, the per-pixel squared reconstruction error is 0.004 with standard deviation 0.005. The threshold was set at $\tau = 0.41$.

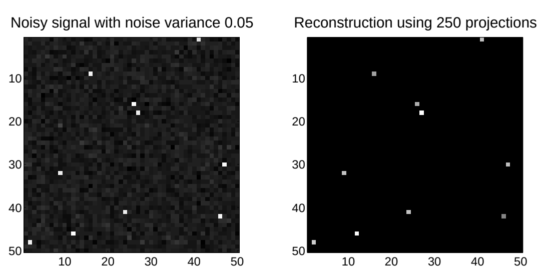

Fig. 2. Simulation results for a 50×50 signal that is sparse in the pixel basis. Noisy sensor readings (left), signal reconstruction using 250 random projections of noisy sensor readings (right). Spikes have amplitude 5 and the noise variance is $\sigma^2 = 0.05$.

6. WIRELESS SENSING

Our theory and method above can be applied to wireless sensing as follows. Consider the problem of sensing a distributed field (e.g., temperature, light, chemical) using a collection of n wireless sensors distributed uniformly over a region of interest. Such systems are often referred to as *sensor networks*. The goal is to obtain an accurate, high-resolution reconstruction of the field at a remote destination. One approach to this problem is to require each sensor to digitally transmit its measurement to the destination, where field reconstruction is then performed. Alternatively, the sensors might collaboratively process their measurements to reconstruct the field themselves and then transmit the result to the destination (i.e., the nodes collaborate to compress their data prior to transmission). Both approaches pose significant demands on communication resources and infrastructure, and it has recently been suggested that non-collaborative analog communication schemes offer a more resource-efficient alternative [10, 11, 12].

Assume that the sensor data is to be transmitted to the destination across an additive white Gaussian noise channel. Suppose the destination broadcasts (perhaps digitally) a random seed to the sensors. Each node modifies this seed in a unique way known to only itself and the destination (e.g., this seed could be multiplied by the node's address or geographic position). Each node generates a pseudorandom Rademacher sequence, which can also be constructed at the destination. Then the nodes *phase-coherently* transmit the random projections to the destination. This is accomplished by requiring each node to simply multiply its reading by an element of its random sequence in each projection/communication step and transmit the result to the destination via amplitude modulation. If the transmissions from all n sensors can be synchronized so that they all arrive in phase at the destination, then the averaging inherent in the multiple access channel computes the desired inner product. After receiving k projections, the destination can employ the reconstruction algorithm above using a basis of choice (e.g., wavelet). The communications procedure is completely non-adaptive and potentially very simple to implement. The collective functioning of the wireless sensors in this process is more akin to an ensemble of phase-coherent emitters than it is to conventional networking operations. Therefore, we prefer the term *sensor ensemble* instead of sensor network in this context.

A remarkable aspect of the sensor ensemble approach is that the power required to achieve a target distortion level can be very minimal. Let σ_s^2 and σ_c^2 denote the noise variance due to sensing and communication, respectively. Thus, each projection received at the destination is corrupted by a noise of total power $\sigma_s^2 + \sigma_c^2$. The sensing noise variance is assumed to be a constant and the additional variance due to the communication channel is assumed to scale like the inverse of the total received power

$$\sigma_c^2 \propto \frac{1}{n^2 P}$$

where P is the transmit power per sensor. Note that although the total transmit power is nP, the received power is a factor of n larger as a result of the *power amplification* effect of the phase-coherent transmissions [12]. In order to achieve rates of distortion decay that we claim, it is sufficient that the variance due to the communication channel behaves like a constant. Therefore, we require only that $P \propto n^{-2}$. This results in a rather surprising conclusion. Optimal reconstruction is possible at the destination with total transmit power nP tending to zero as the density of sensors increases. If conventional spatial point samples were taken instead (e.g., if a single sensor is selected at random in each step and transmits its measurement to the destination), then the power required per sample would be a constant, since only one sensor would be involved in such a transmission. Thus, it appears that random projection sampling could be much more desirable in wireless sensing applications.

7. CONCLUSIONS

We have shown that compressible signals can be accurately recovered from random projections contaminated with noise. The squared error bounds for compressible signals are $O((k/\log n)^{\frac{-2\alpha}{2\alpha+1}})$, which is within a logarithmic factor of the usual nonparametric estimation rate, and $O((k/\log n)^{-1})$ for sparse signals. We demonstrated the effectiveness of random projection sampling with several examples. An interesting line of related and independent work, that we became aware of while finishing this paper, considers signal reconstruction from random projections corrupted by an unknown but bounded perturbation and employs a theoretical approach quite different from ours [13, 14]. One of the most promising potential applications of our theory and method is to wireless sensing, wherein one realizes a large transmission power gain by random projection sampling as opposed to conventional spatial point sampling.

8. REFERENCES

[1] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *Preprint*, June, 2004.

[2] E. Candes and T. Tao, "Near optimal signal recovery from random projections and universal encoding strategies," *Preprint*, October, 2004.

[3] D. L. Donoho, "Compressed sensing," *Preprint*, September, 2004.

[4] C. Craig, "On the Tchebychef inequality of Bernstein," *The Annals of Mathematical Statistics*, vol. 4, no. 2, pp. 94–102, May 1933.

[5] A. R. Barron, "Complexity regularization with application to artificial neural networks," in *Nonparametric Functional Estimation and Related Topics*. Kluwer Academic Publishers, 1988, pp. 561–576.

[6] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *Submitted to IEEE Transactions on Information Theory*, March, 2005.

[7] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[8] M. Figueiredo and R. Nowak, "A bound optimization approach to wavelet-based image deconvolution," *preprint, submitted to IEEE ICIP*, January, 2005.

[9] ——, "A coordinate descent algorithm for basis pursuit," *In Preparation*, 2005.

[10] M. Gastpar and M. Vetterli, "Source-channel communication in sensor networks," in *Proc. IPSN'03*, Apr. 2003, pp. 162–177.

[11] ——, "To code, or not to code: Lossy source-channel communication revisited," *IEEE Trans. Info. Th.*, vol. 49, no. 5, pp. 1147–1158, 2003.

[12] W. Bajwa, A. Sayeed, and R. Nowak, "Matched source-channel communication for field estimation in wireless sensor networks," in *Proc. IPSN 2005, Los Angeles*, 2005.

[13] E. Candes and T. Tao, "Decoding by linear programming," *Preprint*, December, 2004.

[14] E. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Preprint*, February, 2005.