Compressive Domain Interference Cancellation

Mark A. Davenport,* Petros T. Boufounos,† and Richard G. Baraniuk*

*Rice University, Electrical and Computer Engineering, Houston, TX 77005
†Mitsubishi Electric Research Laboratories, Cambridge, MA 02139

Abstract—In this paper we consider the scenario where a compressive sensing system acquires a signal of interest corrupted by an interfering signal. Under mild sparsity and orthogonality conditions on the signal and interference, we demonstrate that it is possible to efficiently filter out the interference from the compressive measurements in a manner that preserves our ability to recover the signal of interest. Specifically, we develop a filtering method that nulls out the interference while maintaining the restricted isometry property (RIP) on the set of potential signals of interest. The construction operates completely in the compressive domain and has computational complexity that is polynomial in the number of measurements.

I. Introduction

A. Motivation

Recent results in compressive sensing (CS) have shown that an N-dimensional signal can be efficiently captured and recovered using $M \ll N$ randomized linear measurements provided that the signal can be sparsely expressed in a known basis or frame [1, 2]. This has motivated a number of practical hardware designs that enable efficient acquisition of sparse signals at the expense of a slight increase in the computation required to recover the signal [3–5].

The compressive measurements acquired by these systems capture the entire signal space, which often includes undesired interference. Since CS-based signal acquisition and processing has been shown to be more susceptible to noise and interference than classical methods [6], it therefore seems prudent to eliminate as much noise and interference as possible prior to any processing.

In this paper we develop an efficient compressive domain filtering algorithm that eliminates signal interference while preserving the geometry of the set of possible signals of interest. Specifically, we show that if the interfering signal lives in a known subspace that is orthogonal to the signal of interest, then we can project the compressive measurements into an orthogonal subspace and thus eliminate the interference. We further demonstrate that the projection operator maintains the *restricted isometry property* (RIP) for the set of signals of interest. Thus, the projected measurements retain
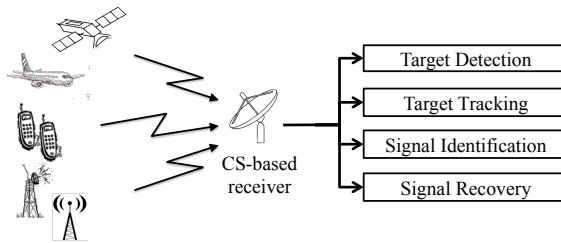

Fig. 1. Example application: broadband signal monitoring.

sufficient information to enable the direct recovery of this signal of interest, or alternatively to enable the use of efficient compressive-domain algorithms for further processing.

B. Applications

In many applications we aim to acquire or monitor a small class of signals in a very large signal space. Such cases are amenable to CS-based acquisition when the signals of interest are sparse or compressible. For example, consider the wide-band signal monitoring and processing station shown in Fig. 1 that receives signals from a variety of sources, including various television, radio, and cell-phone transmissions, radar signals, and satellite communication signals.

In many cases, the monitoring station is only interested in a single type of signal, and the other signals act as interference. A naïve way to proceed would be to recover all of the signal components using a standard CS algorithm (such as ℓ_1-minimization or a greedy algorithm), separate the components due to each of the sources, and then process each recovered component separately. However, this approach does not exploit the particular properties of each component. For example, the problems of detecting the satellite signal, tracking an airplane trajectory, or classifying a cell-phone signal could be amenable to a variety of compressive-domain processing techniques [7–9].

Unfortunately, it is typically more difficult to apply compressive-domain processing techniques when several signals are acquired simultaneously, since the signals

that are not of interest become sources of noise and interference. For example, the presence of a strong television signal might interfere with our ability to detect a weak signal of interest, especially in the compressive domain. It is therefore desirable to filter out the interference before processing in order to provide a degree of robustness.

The compressive domain interference cancellation approach we develop below enables us to efficiently filter out interference from the signals of interest when certain mild sparsity and orthogonality conditions are met. Our filtering approach maintains the geometry of the set of signals of interest, while it eliminates the interference.

Our experiments demonstrate that this approach can improve the performance of subsequent compressive processing/recovery stages. Specifically we demonstrate that rejecting the interference in the compressive domain is beneficial in subsequent signal recovery compared to the naïve approach of first recovering the signal and then rejecting the interference components. The gains are measurable both in the recovery performance and in the computational efficiency of the recovery algorithm.

C. Contrast to classical interference cancellation

The approach we propose in this paper is very different in scope from classical interference cancellation—a widely explored and mature topic with a variety of adaptive and non-adaptive approaches [10]. There are two primary approaches in the literature. Analog cancellation assumes that the interfering signal has enough known parameters that it can be identified and removed in the analog domain before signal acquisition. It is implied that the interference content is not interesting to the acquisition system or its user and can be safely discarded. Digital cancellation assumes that the interference is bandlimited and follows a similar approach after sampling the signal at the Nyquist rate dictated by the combined bandwidth of the signal of interest and the interference.

In contrast, our approach emphasizes the value of simple, general-purpose hardware that captures and processes all the signal of interest together with the interference in a small set of compressive samples. Our approach enables storing and processing of the interfering signal using other system components, a setting in which classical interference cancellation techniques are difficult to apply.

D. Contributions

To summarize, the main contributions of this paper are the following: *(i)* we prove that we can use a simple projection operator to filter out sparse interference with known support in some basis from a compressively acquired signal; *(ii)* we prove that our interference cancellation maintains the RIP for the set of possible signals of interest; *(iii)* we demonstrate that the projection operator can be efficiently computed and applied to the measurements; and *(iv)* we show experimentally that interference cancellation followed by recovery performs significantly better than recovery followed by cancellation both in terms of recovery error and computational efficiency.

II. COMPRESSIVE SENSING

In the standard CS framework, we acquire a signal $x \in \mathbb{R}^N$ via the linear measurements

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

where Φ is an $M \times N$ measurement matrix modeling the sampling system and $y \in \mathbb{R}^M$ is the vector of samples acquired. For simplicity, we deal with real-valued rather than quantized measurements y. Classical sampling theory dictates that to ensure that there is no loss of information the number of samples M should be greater than the signal dimension N. CS, on the other hand, enables us to acquire significantly fewer samples than N, as long as the signal x is sparse or compressible in some basis [1, 2]. Specifically, if we are interested in signals that are K-sparse when represented in the *sparsity basis* Ψ, i.e., $x = \Psi\alpha$ with $\|\alpha\|_0 \leq K$[1], then one can acquire only $M = O(K \log(N/K))$ measurements and still recover the signal x.

A fundamental question in CS concerns the properties of Φ that guarantee satisfactory performance of the sensing system. In [11], Candès and Tao introduced the *restricted isometry property* (RIP) of a matrix Φ and established its important role in CS. Slightly adapted from [11], we say that a matrix Φ satisfies the RIP of order K if there exists constants $a, b \in \mathbb{R}$, $0 < a \leq b < \infty$, such that

$$a\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq b\|x\|_2^2, \tag{2}$$

holds for all x such that $\|x\|_0 \leq K$. In other words, Φ acts as an approximate isometry on the set of vectors that are K-sparse. When we wish to acquire signals that are sparse with respect to the sparsity basis Ψ, we would like to have Φ act as an approximate isometry on vectors that are K-sparse with respect to Ψ. An important result is that for any given Ψ, if we draw a random matrix Φ whose entries ϕ_{ij} are independent realizations from a sub-Gaussian distribution, then $\Phi\Psi$ will satisfy the RIP of order K provided that $M = O(K \log(N/K))$ [12]. Thus, it is common practice to assume for the sake

---

[1]$\| \cdot \|_0$ denotes the ℓ_0 quasi-norm, which simply counts the number of non-zero entries of a vector.

of simplicity that Ψ is the canonical (identity) basis, a convention that we follow for the rest of this paper.

The RIP is a necessary condition if we wish to be able to recover all sparse signals x from the measurements y. Specifically, if $\|x\|_0 = K$, then Φ must satisfy the RIP of order $2K$ with $a > 0$ in order to ensure that any algorithm can recover x from the measurements y. Furthermore, the RIP also suffices to ensure that a variety of practical algorithms can successfully recover any compressible signal from noisy measurements. The following theorem, a slight modification of Theorem 1.2 from [13], makes this precise by bounding the recovery error of x with respect to the sampling noise and with respect the best approximation of x with a K-sparse signal, known as best K-term approximation of x and denoted using x_K.

Theorem 1. *Suppose that Φ satisfies the RIP of order $2K$ with isometry constants satisfying $b/a < 1 + \sqrt{2}$. Given measurements of the form $y = \Phi x + e$, where $\|e\|_2 \le \epsilon$, the solution to*

$$\widehat{x} = \operatorname{argmin} \|x\|_1 \text{ subject to } \|\Phi x - y\|_2 \le \epsilon \quad (3)$$

obeys

$$\|\widehat{x} - x\|_2 \le C_0 \epsilon + C_1 \frac{\|x - x_K\|_1}{\sqrt{K}},$$

where

$$C_0 = \frac{4\sqrt{2}b}{(\sqrt{2}+1)a - b}, \quad C_1 = \frac{(\sqrt{2}-1)a + b}{(\sqrt{2}+1)a - b}.$$

While ℓ_1-minimization techniques are a powerful method for CS signal recovery, there also exist a variety of greedy algorithms that are commonly used in practice and for which performance guarantees similar to that of Theorem 1 can be established. In particular, a greedy algorithm called CoSaMP was recently shown to satisfy similar performance guarantees under slightly stronger assumptions on the RIP constants [14].

III. Compressive Interference Cancellation

Suppose that our signal $x \in \mathbb{R}^N$ consists of two components:

$$x = x_S + x_I,$$

where x_S represents the *signal of interest* and x_I represents *interference* that we would like to reject. We acquire measurements of both components simultaneously

$$y = \Phi(x_S + x_I). \quad (4)$$

Our goal is to remove the contribution of x_I to the measurements y while preserving the information about x_S. While it is not strictly necessary, we assume for

simplicity that the sparsity basis Ψ is the same both for x_S and x_I. As noted above, without loss of generality we may assume this sparsity basis is the canonical (identity) basis.

We cancel the interference by first constructing a linear operator P that operates on the measurements y. The design of P is based solely on the measurement matrix Φ and knowledge of the support of x_I. Specifically, we assume that x_I belongs to a known K_I-dimensional subspace of $\mathbb{R}^N$ having basis $\{e_j\}_{j \in J}$, where e_j denotes the vector of all zeros with a 1 in the j-th position and J denotes a set of indices. Note that if Φ_J denotes the matrix consisting of the columns of Φ indexed by the set J, then Φx_I lies in $\mathcal{R}(\Phi_J)$, the range of Φ_J. The operator P we construct should map $\mathcal{R}(\Phi_J)$ to zero; i.e., we want the nullspace of P to be equal to $\mathcal{R}(\Phi_J)$. P will obviously depend on the set J, but for simplicity we omit this dependence in our notation.

There are a variety of methods for constructing a P with the desired nullspace properties. Furthermore, each construction might be computed with several numerical methods, which affect the performance and stability of the construction. For example, if $\widetilde{\Phi}_J$ is any orthonormal basis for $\mathcal{R}(\Phi_J)$, then

$$P = I - \widetilde{\Phi}_J \widetilde{\Phi}_J^*$$

is an orthogonal projection whose nullspace is equal to $\mathcal{R}(\Phi_J)$. One could obtain $\widetilde{\Phi}_J$ via Gram-Schmidt orthogonalization of the columns of Φ_J or via the singular value decomposition of Φ_J [15]. If Φ satisfies the RIP, then Φ_J is well-conditioned, and so these methods will provide a stable method for computing $\widetilde{\Phi}_J$. Extensive investigation of the numerical properties of these constructions is beyond the scope of this paper.

We will focus on the case where Φ admits a fast transform-based implementation; we will construct P to leverage Φ's structure and ease the computational cost of applying P. For example, Φ may consist of random rows of a Discrete Fourier Transform or a permuted Hadamard Transform matrix. In this case, rather than constructing the matrix $\widetilde{\Phi}_J$, we use

$$P = I - \Phi_J \Phi_J^\dagger,$$

where $\Phi_J^\dagger$ is the pseudoinverse $\Phi_J^\dagger = (\Phi_J^* \Phi_J)^{-1} \Phi_J^*$. Note that since $\Phi x_I \in \mathcal{R}(\Phi_J)$ there exists an $\alpha \in \mathbb{R}^{K_I}$ such that

$$\begin{aligned} P\Phi x_I &= P\Phi_J \alpha \\ &= (I - \Phi_J(\Phi_J^* \Phi_J)^{-1} \Phi_J^*)\Phi_J \alpha \\ &= \Phi_J \alpha - \Phi_J \alpha = 0. \end{aligned} \quad (5)$$

3

Thus for any x_I supported on the set J, $P\Phi x_I = 0$, i.e., P eliminates the interference x_I from the samples y. However, unlike the prior construction, if we have a fast transform-based implementation of Φ and Φ^*, then we can use the conjugate gradient method or Richardson iteration to efficiently compute Py [14].

From (4) and (5), $Py = P\Phi x_S + P\Phi x_I = P\Phi x_S$. We now need to ensure that $P\Phi x_S$ contains sufficient information about x_S. In particular, we wish to show that the matrix $P\Phi$ satisfies a relaxed version of the RIP. Towards this end, we first establish a useful bound.

Lemma 1. *Suppose that x and z are orthogonal and that $\|x \pm z\|_0 \leq K$. If Φ satisfies the RIP of order K, then*

$$\frac{|\langle \Phi x, \Phi z \rangle|}{\|\Phi x\|_2 \|\Phi z\|_2} \leq \frac{b-a}{2a}.$$

Proof: First assume that $\|x\|_2 = \|z\|_2 = 1$. Since x and z are orthogonal, $\|x \pm z\|_2^2 = \|x\|_2^2 + \|z\|_2^2 = 2$, and hence from the RIP we have that

$$2a \leq \|\Phi x \pm \Phi z\|_2^2 \leq 2b.$$

From the parallelogram identity we obtain

$$
\begin{aligned}
|\langle \Phi x, \Phi z \rangle| &\leq \frac{1}{4} \left| \|\Phi x + \Phi z\|_2^2 - \|\Phi x - \Phi z\|_2^2 \right| \\
&\leq \frac{1}{4}(2b - 2a) = \frac{b-a}{2}.
\end{aligned}
$$

From the bilinearity of the inner product and the RIP, for x, z with arbitrary norm, the lemma follows:

$$
\begin{aligned}
|\langle \Phi x, \Phi z \rangle| &\leq \frac{b-a}{2} \|x\|_2 \|z\|_2 \\
&\leq \frac{b-a}{2a} \|\Phi x\|_2 \|\Phi z\|_2.
\end{aligned}
$$

∎

Thus, any sparse signal that is orthogonal to x_I will remain nearly orthogonal to Φx_I in the compressive domain. Using this result we can show that if Φ has the RIP, then $P\Phi$ satisfies the RIP restricted to the set of signals that are orthogonal to the interference subspace.

Theorem 2. *Given an index set J with cardinality $\#J \leq K_I$, let Φ_{J^c} denote the matrix consisting of the columns of Φ indexed by the set $J^c = \{1, 2, \ldots, N\} \setminus J$. If Φ satisfies the RIP of order $K = 2K_S + K_I$, then $P\Phi_{J^c}$ satisfies*

$$\widetilde{a}\|\widetilde{x}\|_2^2 \leq \|P\Phi_{J^c}\widetilde{x}\|_2^2 \leq b\|\widetilde{x}\|_2^2,$$

for all $\widetilde{x} \in \mathbb{R}^{N-K_I}$ such that $\|\widetilde{x}\|_0 \leq 2K_S$, where

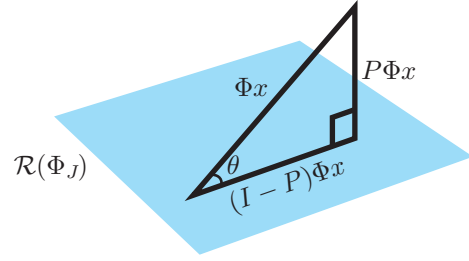$$\widetilde{a} = a - \frac{(b-a)^2}{4a}.$$

Fig. 2. *Decomposition of Φx into $P\Phi x$ and $(I-P)\Phi x$.*

Proof: We first let $\widetilde{x}$ be any vector such that $\|\widetilde{x}\|_0 \leq 2K_S$. Now define x so that $x_n = \widetilde{x}_n$ for $n \in J^c$ and $x_n = 0$ for $n \in J$, i.e., x is the extension of $\widetilde{x}$ into $\mathbb{R}^N$. Then $\Phi_{J^c}\widetilde{x} = \Phi x$. We can decompose Φx as $\Phi x = P\Phi x + (I-P)\Phi x$. Since P is an orthogonal projection we can write

$$\|\Phi x\|_2^2 = \|P\Phi x\|_2^2 + \|(I-P)\Phi x\|_2^2. \tag{6}$$

This is illustrated in Fig. 2. Our goal is to show that $\|\Phi x\|_2 \approx \|P\Phi x\|_2$, or equivalently that $\|(I-P)\Phi x\|_2$ is small. Towards this end, note that if θ is the angle between Φx and $(I-P)\Phi x$, then

$$\cos\theta = \frac{\|(I-P)\Phi x\|_2}{\|\Phi x\|_2} = \frac{\langle (I-P)\Phi x, \Phi x \rangle}{\|(I-P)\Phi x\|_2 \|\Phi x\|_2}. \tag{7}$$

Note that $(I-P)$ is a projection onto $\mathcal{R}(\Phi_J)$. Thus there exists an α such that $(I-P)\Phi x = \Phi_J \alpha$. Furthermore, by assumption, x is orthogonal to $\mathcal{R}(\Phi_J)$. Hence we may apply Lemma 1 to obtain

$$\frac{|\langle (I-P)\Phi x, \Phi x \rangle|}{\|(I-P)\Phi x\|_2 \|\Phi x\|_2} \leq \frac{b-a}{2a}.$$

Combining this with (7), we obtain

$$\|(I-P)\Phi x\|_2 \leq \frac{b-a}{2a}\|\Phi x\|_2.$$

Since we trivially have that $\|(I-P)\Phi x\|_2 \geq 0$, we can combine this with (6) to obtain

$$\left(1 - \left(\frac{b-a}{2a}\right)^2\right)\|\Phi x\|_2^2 \leq \|P\Phi x\|_2^2 \leq \|\Phi x\|_2^2.$$

Since $\|x\|_0 \leq 2K_S$, we have that

$$\left(1 - \left(\frac{b-a}{2a}\right)^2\right)a\|x\|_2^2 \leq \|P\Phi x\|_2^2 \leq b\|x\|_2^2.$$

Recalling that $\Phi_{J^c}\widetilde{x} = \Phi x$, and since $\|x\|_2 = \|\widetilde{x}\|_2$, the theorem follows. ∎

One can easily verify that if $b/a \leq 1.91$, then $b/\widetilde{a} \leq 1 + \sqrt{2}$. Thus, from Theorem 1 we conclude that under a slightly more restrictive bound on the required

RIP constants, we can directly recover a sparse signal of interest x_S that is orthogonal to the interfering x_I without actually recovering x_I. This has a number of practical benefits. For instance, if we were instead to attempt to first recover x and then cancel x_I, then we would require RIP of order $2(K_S + K_I)$ to ensure that this recover-then-cancel approach will be successful. In contrast, our approach requires RIP of order only $2K_S + K_I$. In certain cases (when K_I is significantly larger than K_S), this results in a substantial decrease in the required number of measurements. Furthermore, since most recovery algorithms have computational complexity that is at least linear in the number of coefficients of the recovered signal, this can also result in substantial computational savings.

IV. Experiments

In this section we evaluate the performance of our proposed cancellation method in the context of attempting to recover the signal of interest x_S while canceling out the interfering signal x_I. Rather than ℓ_1-minimization we use the iterative CoSaMP greedy algorithm since it more naturally naturally lends itself towards a simple modification described below. More specifically, we evaluate three interference cancellation approaches:

1) **Cancel-then-recover:** This is the approach advocated in this paper. We cancel out the contribution of x_I to the measurements y and directly recover x_S using the CoSaMP algorithm.

2) **Modified recovery:** Since we know the support of x_I, rather than cancelling out the contribution from x_I to the measurements, we modify a greedy algorithm such as CoSaMP to exploit the fact that part of the support of x is known in advance. This modification is made simply by forcing CoSaMP to always keep the elements of J in the active set at each iteration. After recovering $\widehat{x}$, we then set $\widehat{x}_n = 0$ for $n \in J$ to filter out the interference.

3) **Recover-then-cancel:** In this approach, we simply ignore that we know the support of x_I and try to recover the signal x using the standard CoSaMP algorithm, and then set the $\widehat{x}_n = 0$ for $n \in J$ as before.

In our experiments, we set $N = 1000$, $M = 200$, and $K_S = 10$. We then considered a range of values of K_I from 1 to 100. For each value of K_I, we generated 2000 test signals where the coefficients were selected according to a Gaussian distribution, and then contaminated with an N-dimensional Gaussian noise vector. As a reference for comparison, we also considered an oracle decoder that is given the support of both x_I and x_S and
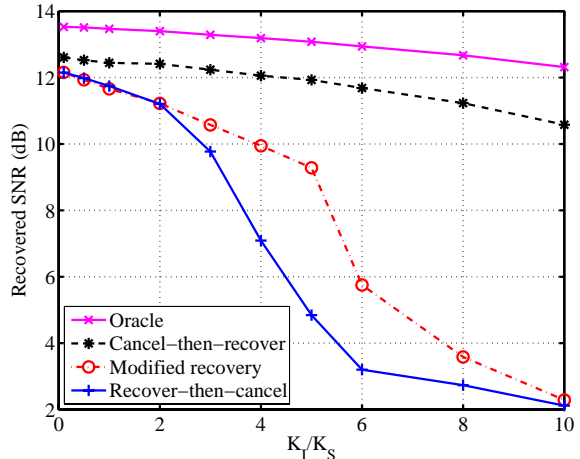

Fig. 3. *SNR of x_S recovered using the three different cancellation approaches for different ratios of K_I to K_S compared to the performance of an oracle.*

solves the least-squares problem restricted to the known support set.

We considered a range of signal-to-noise ratios (SNRs) and signal-to-interference ratios (SIRs). Fig. 3 shows the results for the case where x_S and x_I are normalized to have equal energy (an SIR of 0dB) and where the variance of the noise is selected so that the SNR is 15dB. Our results were consistent for a wide range of SNR and SIR values, and we omit the plots due to space limitations.

Our results show that the cancel-then-recover approach performed significantly better than both of the other methods as K_I grows larger than K_S. While both of the other methods begin to suffer as K_I grows large, the cancel-then-recover approach continues to perform almost as well as the oracle decoder for the entire range of K_I. We also note that the while the modified recovery method did perform slightly better than the recover-then-cancel approach, the improvement is relatively minor.

We observe similar results in Fig. 4 for the recovery time (which includes the cost of computing the matrix P in the cancel-then-recover approach), with the cancel-then-cancel approach performing significantly faster than the other approaches as K_I grows larger than K_S.

V. Discussion

In this paper, we have demonstrated that with a small penalty on the RIP constants we can efficiently null out any known sparse interference via simple and efficient filtering in the compressive domain.

Several of the compressive-domain processing algorithms described in the Introduction rely on the compressive sensing system preserving the geometry of the
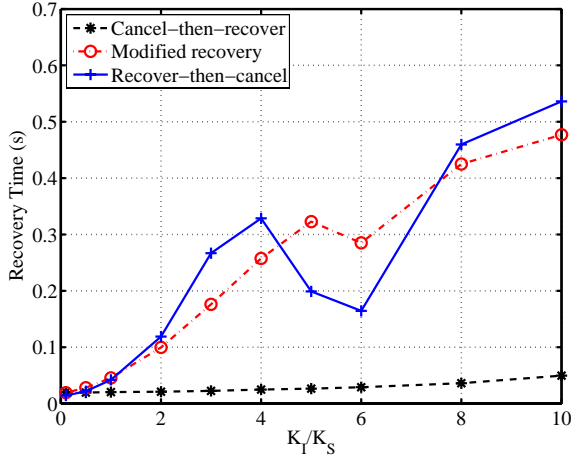
Fig. 4. *Recovery time for the three different cancellation approaches for different ratios of K_I to K_S.*

signal space [7–9]. Our interference cancellation method maintains the RIP property of the sampling system which guarantees that distances between signals in the space are not affected by the sampling process. Thus we ensure the stable embedding of signals of interest in the compressive domain and provide robustness to measurement noise and sampling non-idealities.

The interference cancellation computation is in general more efficient than signal recovery, even when the cost of computing the projection/cancellation operator P is taken into account. Furthermore, in many practical applications, there may exist extremely efficient methods for computing Py. Alternatively, the projection operator P can be precomputed and the cost amortized over several applications of the same sampling operator Φ.

The work presented in this paper paves the way for compressive-domain filtering and signal processing. For example, a future step in this direction is to generalize our filters beyond simple projection schemes to more arbitrary compressive domain filters. Furthermore, there exist even more efficient and structured approaches to these problems for practical structured CS systems such as the random sampler [4] or the random demodulator [5]. More generally, one could study the impact of such filtering approaches on a wide variety of compressive domain processing algorithms.

REFERENCES

[1] E. J. Candès, "Compressive sampling," in *Proc. International Congress of Mathematics*, Madrid, Spain, 2006, vol. 3, pp. 1433–1452.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Info. Theory*, vol. 52, no. 4, pp. 1289–1306, Sep. 2006.

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

[4] J. N. Laska, S. Kirolos, Y. Massoud, R. G. Baraniuk, A. C. Gilbert, M. Iwen, and M. J. Strauss, "Random sampling for analog-to-information conversion of wideband signals," in *Proc. IEEE Dallas Circuits and Systems Workshop (DCAS)*, Dallas, TX, Oct. 2006.

[5] J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk, "Beyond nyquist: Efficient sampling of sparse bandlimited signals," 2009, Preprint.

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

[7] M. A. Davenport, M. B. Wakin, and R. G. Baraniuk, "Detection and estimation with compressive measurements," Tech. Rep. TREE 0610, Rice University ECE Department, Houston, TX, 2006.

[8] M. A. Davenport, M. F. Duarte, M. B. Wakin, J. N. Laska, D. Takhar, K. F. Kelly, and R. G. Baraniuk, "The smashed filter for compressive classification and target recognition," in *Proc. IS&T/SPIE Symposium on Electronic Imaging: Computational Imaging*, San Jose, CA, Jan. 2007, vol. 6498.

[9] C. Hegde, M.B. Wakin, and R.G. Baraniuk, "Random projections for manifold learning," in *Neural Information Processing Systems (NIPS)*, 2007.

[10] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part IV: Optimum Array Processing*, John Wiley & Sons, 2002.

[11] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, pp. 4203–4215, Dec. 2005.

[12] R. G. Baraniuk, M. Davenport, R. A. 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.

[13] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," in *Compte Rendus de l'Academie des Sciences, Paris, Series I*, 2008, vol. 346, pp. 589–592.

[14] D. Needell and J. A. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Accepted to Appl. Comp. Harmonic Anal.*, 2008.

[15] L. N. Trefethen and III D. Bau, *Numerical Linear Algebra*, SIAM Review, 1997.