SPARSE SIGNAL RECONSTRUCTION FROM NOISY COMPRESSIVE MEASUREMENTS USING CROSS VALIDATION

Petros Boufounos, Marco F. Duarte, Richard G. Baraniuk

Rice University, Electrical and Computer Engineering, Houston, TX 77005

ABSTRACT

Compressive sensing is a new data acquisition technique that aims to measure sparse and compressible signals at close to their intrinsic information rate rather than their Nyquist rate. Recent results in compressive sensing show that a sparse or compressible signal can be reconstructed from very few incoherent measurements. Although the sampling and reconstruction process is robust to measurement noise, all current reconstruction methods assume some knowledge of the noise power or the acquired signal to noise ratio. This knowledge is necessary to set algorithmic parameters and stopping conditions. If these parameters are set incorrectly, then the reconstruction algorithms either do not fully reconstruct the acquired signal (underfitting) or try to explain a significant portion of the noise by distorting the reconstructed signal (overfitting). This paper explores this behavior and examines the use of cross validation to determine the stopping conditions for the optimization algorithms. We demonstrate that by designating a small set of measurements as a validation set it is possible to optimize these algorithms and reduce the reconstruction error. Furthermore we explore the trade-off between using the additional measurements for cross validation instead of reconstruction.

Index Terms— Data acquisition, sampling methods, data models, signal reconstruction, parameter estimation.

1. INTRODUCTION

Compressive sensing (CS) is a new data acquisition technique that aims to measure sparse and compressible signals at close to their intrinsic information rate rather than their Nyquist rate [1, 2]. The fundamental premise is that certain classes of signals, such as natural images, have a concise 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 the Fourier and wavelet bases, respectively.

The traditional mode of data acquisition is to first uniformly sample the signal (at or above its Nyquist rate). Since for wideband signals this often results in too many samples, sampling is often followed by a second compression step. In transform-domain compression (transform coding), the raw data samples are transformed to a sparse representation in a sparsity basis; the large coefficients are kept while the small coefficients are discarded, thereby reducing the amount of data required to be stored, processed, or transmitted.

Recent results in CS demonstrate that a condensed version of a sparse or compressible signal can be directly acquired in one step using a low-rate acquisition process that projects it onto a small set of vectors that is incoherent with the sparsity basis. The signal is subsequently recovered using an optimization (linear program) or greedy algorithm that determines the sparsest representation consistent with the measurements. The quality of the reconstruction depends on the sparsity of the original signal, the choice of the reconstruction algorithm, and the degree of incoherence. One of the most attractive features of compressive sensing is that random vectors are incoherent with any sparsity-inducing basis with high probability.

Since noise is often present in real data acquisition systems, a range of different algorithms have been developed that enable exact reconstruction of sparse signals from noisy compressive measurements [1–8]. The reconstruction quality for compressible signals is comparable to that of the signal's optimal sparse approximation (obtained by keeping only the largest coefficients in the sparsity basis).

There are a number of caveats for noisy CS signal reconstruction, however. First, current reconstruction algorithms only provide worst-case performance guarantees. Second, current reconstruction algorithms generally assume that the noise is either bounded or Gaussian with known variance. Third, most current reconstruction algorithms are iterative and use information about the noise magnitude to establish a stopping criterion. Finally, in practice, the sparsity or compressibility of the signal is often unknown; this can lead to either early or late termination of an iterative reconstruction algorithm. In the former case, the signal has not been completely reconstructed (underfitting), while in the latter case some of the noise is treated as signal (overfitting). In both cases, the reconstruction quality is inferior.

In this paper, we take the viewpoint that noisy CS signal reconstruction is essentially a model order selection and parameter estimation problem, which makes *cross validation* (CV) immediately applicable. Cross validation is a statistical technique that separates a data set into a training/estimation set and a test/cross validation set. The test set is used to estimate noise levels and reduce the model order complexity so that it does not overfit the data.

The key property of CS that enables the application of CV techniques stems from the incoherence between the measurement vectors and the sparsity basis. In a nutshell, incoherence makes CS "democratic" in that it spreads the signal information evenly amongst all the measurements, giving each equal weight in the reconstruction [1, 2]. Furthermore, any sufficiently large set of measurements is as suitable as any other for reconstruction. While additional measurements could be used to improve the reconstruction quality, we will show below that holding these measurements out for CV can significantly improve the CS reconstruction of noisy signals.

This paper is organized as follows. Section 2 introduces the requisite background on CS and CV and establishes our notation. Section 3 develops our CS/CV framework for noisy signals. Section 4

This research was supported by ONR grants N00014-06-1-0768 and N00014-06-1-0829; AFOSR grant FA9550-04-0148; DARPA grants N66001-06-1-2011 and N00014-06-1-0610; NSF grants CCF-0431150, CNS-0435425, and CNS-0520280; and the Texas Instruments Leadership University Program. E-mail: {petrosb,duarte,richb}@rice.edu. Web: dsp.rice.edu/cs

provides experimental results for several reconstruction algorithms. We close in Section 5 with a discussion and conclusions.

2. BACKGROUND ON COMPRESSIVE SENSING

2.1. Compressive Measurements

We consider signals $\mathbf{x} \in \mathbb{R}^N$ that are sparse or compressible in some sparsity basis. Without loss of generality, we assume that $\mathbf{x}$ is sparse in the canonical basis of $\mathbb{R}^N$. The signal is sparse if only K of its coefficients are non-zero, with $K \ll N$. The signal is compressible if $\mathbf{x} \in \ell_p$ for $p < 2$; in this case, its sorted coefficients decay rapidly and are well-approximated as K-sparse.

We measure the signal by taking inner products with a set of M vectors $\{\phi_i, i = 1, \ldots, M\}$ that are incoherent with the sparsity basis. By incoherent we mean that none of the vectors $\{\phi_i, i = 1, \ldots, M\}$ have a sparse or compressible representation in terms of the sparsity basis; see [2] for a precise definition. The measurement process is a linear map $\mathbf{\Phi} : \mathbb{R}^N \to \mathbb{R}^M$. The measurements are then corrupted by additive zero-mean white noise $\mathbf{n}$ of variance σ^2 per measurement — i.e., with an expected noise variance of $\sigma^2_{\mathbf{n}} = M\sigma^2$. The resulting measurement vector $\mathbf{y}$ is expressed as

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} + \mathbf{n}, \qquad E\{\mathbf{n}\mathbf{n}^T\} = \sigma^2 \mathbf{I}_M, \tag{1}$$

where $\mathbf{I}_M$ is the $M \times M$ identity matrix.

2.2. Reconstruction from Noise-free Measurements

If the measurement process satisfies the Restricted Isometry Property (RIP) conditions described in [2], then a sparse/compressible signal can be recovered exactly/approximately using sparse reconstruction algorithms that determine the sparsest signal $\widehat{\mathbf{x}}$ that explains the measurements $\mathbf{y}$ [1, 2, 4]. Specific reconstruction algorithms include linear programming (Basis Pursuit) [9] and Orthogonal Matching Pursuit (OMP) [4]; numerical experiments demonstrate good performance using Matching Pursuit (MP) [10] for reconstruction even though there are no theoretical guarantees. MP is often preferred to OMP due to its significantly reduced computational complexity.

2.3. Reconstruction from Noisy Measurements

In the presence of measurement noise, variations of the aforementioned algorithms have been shown to reliably approximate the original signal, assuming certain noise or signal parameters are known. All the algorithms used in compressive sensing solve one of the following formulations.

Basis Pursuit with Inequality Constraints relaxes the requirement that the reconstructed signal exactly explain the measurements. Instead, the constraint is expressed in terms of the maximum distance of the measurements from the re-measured reconstructed signal. The reconstruction solves the program

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1} \text{ s.t. } \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_{\ell_2} \le \epsilon, \tag{2}$$

some small $\epsilon > 0$. In [3] it is shown that if the noise is power-limited to ϵ and enough measurements are taken, then the reconstructed signal $\widehat{\mathbf{x}}$ is guaranteed to be within $C\epsilon$ of the original signal $\mathbf{x}$:

$$\|\mathbf{x} - \widehat{\mathbf{x}}\|_{\ell_2} \le C\epsilon,$$

where the constant C depends only on the measurement parameters, and not on the level of noise. Unfortunately, noise in practice is not necessarily power limited, and, even when it is, the power limit is usually unknown.

The Dantzig Selector is an alternative convex program useful when the noise is unbounded [5]. Specifically, for the measurement assumptions in (1) and if enough measurements are taken, the convex program

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1} \text{ s.t. } \|\mathbf{\Phi}^*(\mathbf{y} - \mathbf{\Phi}\mathbf{x})\|_{\ell_\infty} \le \sqrt{2\log N}\sigma$$

reconstructs a signal that satisfies

$$\|\mathbf{x} - \widehat{\mathbf{x}}\|_{\ell_2} \le \sigma \cdot C\sqrt{2\log(N)K}.$$

Similar results apply for compressible signals. Unfortunately, this optimimization also requires a priori knowledge of the error variance and the signal sparsity.

The Lasso and Basis Pursuit De-Noising are two alternative formulations of the same objective. In particular, Basis Pursuit De-Noising relaxes the hard constraint on the reconstruction error magnitude with a soft weight λ in the following program:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1} + \lambda\|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_{\ell_2}. \tag{3}$$

With appropriate parameter correspondence, this formulation is equivalent to the Lasso [11]:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{\Phi}\mathbf{x}\|_{\ell_2} \text{ s.t. } \|\mathbf{x}\|_{\ell_1} \le q. \tag{4}$$

Furthermore it is demonstrated in [11] that as λ ranges from zero to infinity, the solution path of (3) is the same at the solution path of (4) as q ranges from infinity to zero. An efficient algorithm that traces this path is mentioned and experimentally analyzed in [12]. It follows that determining the proper value of λ, even if all the solutions are available, is akin to determining the power limit ϵ of the noise.

These three reconstruction formulations are based on the same principle: that ℓ_1 minimization, under certain conditions on the measurement vectors and sparsity basis, recovers the support (locations) of the non-zero coefficients of the sparse representation. These algorithms are often followed by a subsequent step, known as debiasing, in which a standard least squares problem is solved on the support. Debiasing has been shown to lower the reconstruction error; see Sec. 4.

Matching Pursuit (MP) is a greedy algorithm that iteratively incorporates in the reconstructed signal the component from the measurement set that explains the largest portion of the residual from the previous iteration [13]. At each iteration i, the algorithm computes:

$$\begin{aligned}
c_k^{(i)} &= \langle \mathbf{r}^{(i-1)}, \phi_k \rangle, \\
\widehat{k} &= \arg\max_k |c_k^{(i)}|, \\
\mathbf{x}^{(i)} &= \mathbf{x}^{(i-1)} + c_{\widehat{k}}^{(i)} \delta_{\widehat{k}}, \\
\mathbf{r}^{(i)} &= \mathbf{r}^{(i-1)} - c_{\widehat{k}}^{(i)} \phi_{\widehat{k}},
\end{aligned}$$

where $\mathbf{r}^{(i)}$ is the residual after iteration i, with $\mathbf{r}^{(i)} = \mathbf{y} - \mathbf{\Phi}\mathbf{x}^{(i)}$. The algorithm is initialized using $\mathbf{x}^{(0)} = \mathbf{0}$ and $\mathbf{r}^{(0)} = \mathbf{y}$ and terminated when the residual has magnitude below a set threshold.

Orthogonal Matching Pursuit (OMP) additionally orthogonalizes the residual against all measuremnent vectors selected in previous iterations. Although this step increases the complexity of the algorithm, it improves its performance and provides better reconstruction guarantees compared to plain old MP. The orthogonalization step is similar to the debiasing modification of the above optimization based algorithms, with the exception that it is performed at every iteration of the algorithm rather than at the end.

Both MP and OMP can be shown to converge to a solution that fully explains the data and the noise. However, only OMP is guaranteed to converge to a sparse solution. It is theoretically possible that MP converges to a dense solution that explains the measurements and the noise, but does not approximate well the original signal. Experimental results, on the other hand, demonstrate good performance. They further demonstrate that proper termination of the algorithms is a practical way to reject the measurement noise in the reconstruction. However, the conditions for proper termination are heuristic. Even in the case the stopping condition from the Basis Pursuit De-Noising algorithm is heuristically used, namely $\|\mathbf{r}^{(i)}\| \leq \epsilon$, the noise level ϵ is still a necessary input to the algorithm, as is the case with all the methods described in this section. In order to determine the correct parameter values, we propose to apply CV.

2.4. Cross Validation

Cross Validation is a statistical technique to determine the appropriate model order complexity and thus avoid overfitting a model to a set of sample data. CV first divides the data to two distinct sets: a training/estimation set, and a test/cross validation set. The model fitting algorithm operates on the training set, and then its performance is predicted on the testing set. As the algorithm estimates the global parameters of the model and increases the model complexity and the accuracy of the estimated parameters, the prediction performance on the CV set increases. However, when the algorithm begin overfitting the training set, its performance on the testing set decreases. Thus, a further increase in the model complexity is not beneficial and the algorithm should terminate.

Cross validation can be immediately applied to any iterative algorithm, as described in Sec. 3.2. From the family of matching pursuits we focus on the OMP algorithm due to its superior performance in our experiments. Similarly, several algorithms exist that solve each of the ℓ_1-based formulations mentioned in Sec. 2.3. From those algorithms, we focus the **Homotopy continuation** algorithm introduced in [14] as a solution to the Lasso formulation. This is the same as the **Least Angle Regression** (LARS) algorithm with the Lasso modification described in [12]. The key property of Homotopy continuation is that as it iterates it visits all of the solutions to the minimization in Eq. (4), and consequently (3), as λ increases from zero to infinity. Thus we are able to introduce CV at each iteration in order to determine the appropriate parameter values from all of the solutions visited.

There exist algorithms that solve the Lasso formulation more efficiently than the Homotopy continuation algorithm if the parameter values are predetermined and known in advance; for examples see [6–8]. To successfully use these algorithms with cross validation, however, it is necessary to execute them a number of times, each for a different parameter value, thus defeating the speedup of the algorithms.

3. CROSS VALIDATION FOR COMPRESSIVE SENSING

We now demonstrate how CV can be profitably applied to CS reconstruction of noisy data using any iterative signal estimation algorithm.

3.1. Cross Validation Measurement Model

Consistent with the measurement model of (1), we propose to take additional measurements using:

$$\mathbf{y}_{\rm cv} = \mathbf{\Phi}_{\rm cv}\mathbf{x} + \mathbf{n}_{\rm cv},$$
$$E\{\mathbf{n}_{\rm cv}\mathbf{n}_{\rm cv}^T\} = \sigma^2\mathbf{I}_{M_{\rm cv}},$$

in which $M_{\rm cv}$ denotes the number of additional measurements, $\mathbf{\Phi}_{\rm cv}$ denotes the CV measurement matrix and $\mathbf{n}_{\rm cv}$ denotes the noise added on the CV measurements. Since the same data acquisition system is assumed to be used to obtain both the reconstruction and the cross-validation measurements, the CV noise has the same per-measurement variance as the estimation noise in (1). It is, thus, assumed that the cross validation measurement matrix is generated in the same way as the acquisition measurement matrix and properly normalized such that the signal to noise ratio for the signal acquisition measurements and the cross validation measurements is the same. For example, if the reconstruction measurement matrix $\mathbf{\Phi}$ is normalized to have unit column norm, then the CV measurement matrix $\mathbf{\Phi}_{\rm cv}$ is normalized to have column norm equal to $\sqrt{M_{\rm cv}/M}$. On the other hand, if $\mathbf{\Phi}$ contains i.i.d. entries drawn form a distribution with a certain variance, then $\mathbf{\Phi}_{\rm cv}$ also contains i.i.d. entries drawn from the same distribution, with the same variance.

3.2. CV-based Modification to CS Reconstruction Algorithms

Consider any iterative CS reconstruction algorithm for noisy data such as MP, OMP, or Homotopy continuation/LARS. Each iteration produces a signal estimate $\widehat{\mathbf{x}}^{(i)}$. To be able to use cross-validation we modify the CS reconstruction algorithm by wrapping each iteration in a CV step as follows:

1. **Initialize:**
   Set $\epsilon = \|\mathbf{y}_{\rm cv}\|_2$, $\widetilde{\mathbf{x}} = 0$, $i = 1$.
   Initialize the estimation algorithm.
2. **Estimate:**
   Execute one estimation iteration to compute $\widehat{\mathbf{x}}^{(i)}$.
3. **Cross validate:**
   If $\|\mathbf{y}_{\rm cv} - \mathbf{\Phi}_{\rm cv}\widehat{\mathbf{x}}^{(i)}\|_2 < \epsilon$ then set $\epsilon = \|\mathbf{y}_{\rm cv} - \mathbf{\Phi}_{\rm cv}\widehat{\mathbf{x}}^{(i)}\|_2$, $\widetilde{\mathbf{x}} = \widehat{\mathbf{x}}^{(i)}$, and terminate.
4. **Iterate:**
   Increase i by 1 and iterate from 2.

We refer to this modification as the CS-CV modification to the CS algorithm (e.g., OMP-CV). CS-CV can be terminated after a sufficient number of iterations have been executed, the number of which depends on the original CS reconstruction algorithm. At the termination of the modified algorithm, $\widetilde{\mathbf{x}}$ contains the estimated signal with the minimum cross validation error; ϵ is the norm of this error.

If enough CV measurements are used, then the CV error after each iteration $\|\mathbf{y}_{\rm cv} - \mathbf{\Phi}_{\rm cv}\widehat{\mathbf{x}}^{(i)}\|_2$ follows a relatively smooth path as follows: as the algorithm improves the estimate of the signal, the CV error decreases; as the algorithm starts overfitting the noise, the CV error increases. The number of iterations should be sufficient for the original CS reconstruction algorithm to reconstruct the signal and start oververfitting the noise. For example, if OMP-CV is used on a K-sparse signal, it should iterate at least K times. In practice, CS-CV can be terminated if a number of iterations have produced no improvement in the CV error. For example, a typical run of OMP is shown in Fig. 1. At each iteration, the norm of the residual on $\mathbf{y}$, the measurements used in the reconstruction of the signal, (solid line) decreases. However, the reconstruction error (dash-dot line)

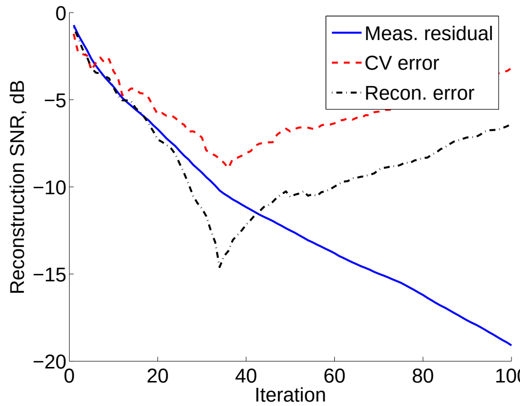

Fig. 1. *Evolution of the error at each iteration of a typical run of CV-based OMP reconstruction. The errors are normalized with respect to their initial values.*

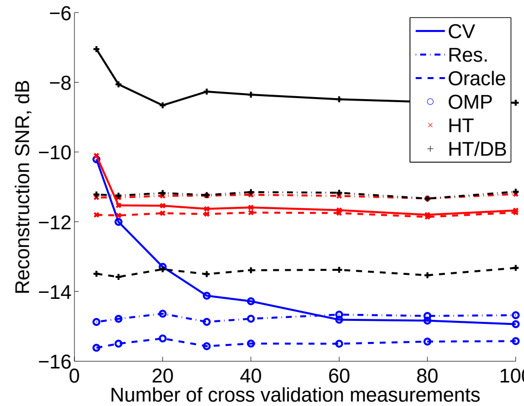

Fig. 2. *Comparison of standard reconstruction algorithms from noisy compressive sensing measurements with several stopping criteria. We fix the parameters $N = 1000$, $K = 50$, $M = 400$, and $\sigma_{\mathbf{n}} = 2$. CV performs better than the residual criterion (Res.) when sufficient CV measurements are used; the oracle performance is shown for reference. Additionally, OMP outperform both Homotopy continuation (HT) and Homotopy continuation with debiasing (HT/DB).*

increases after the 36th iteration. CV ably detects this change, as demonstrated with the residual on $\mathbf{y}_{\rm cv}$ (dashed line). In practice, the algorithm can be safely terminated after the 40th iteration.

4. EXPERIMENTAL RESULTS

This section presents experiments that demonstrate that CV can be combined with the aforementioned standard algorithms for signal reconstruction from noisy compressive measurements. We test how the problem parameters affect the performance of the reconstruction algorithms and compare the results with the performance of the algorithms assuming the noise parameters are known in advance.

4.1. Experiment Setup

For all the experiments we use exactly sparse signals of length $N = 1000$, with the support of the signal randomly selected from a uniform distribution. The non-zero coefficients are drawn from a standard Gaussian distribution. The signals are measured using a measurement matrix $\mathbf{\Phi}$ and a CV matrix $\mathbf{\Phi}_{\rm cv}$ that have i.i.d. entries drawn from a Rademacher distribution of $\{\pm 1/\sqrt{M}\}$ with equal probability. The noise added to the measurements is properly normalized to maintain the acquired signal-to-noise ratio (SNR) between the reconstruction and the CV measurements constant.

For each algorithm we evaluate the performance using the new CV stopping criterion and the inequality stopping criterion of Eq. (2); we refer to the latter as the residual stopping criterion; for reference, we also evaluate the performance of an oracle that selects the solution for the iteration that yields the lowest distortion reconstruction. We evaluate the performance of the algorithm as a function of the number of CV measurements and use the SNR of the reconstruction as the performance metric. We average 200 repetitions of each experiment, with different realizations of the sparse supports, random measurement and CV matrices, and additive noise.

4.2. Number of Measurements

When the total number of measurements $M + M_{\rm cv}$ is fixed, there is a tradeoff between the number of compressive measurements M and the number of CV measurements $M_{\rm cv}$. On the one hand, any measurements taken beyond the minimum number required to resolve and reconstruct an estimate [15] can be used to reduce the reconstruction error, which suggests increasing M and decreasing $M_{\rm cv}$. On the other hand, increasing $M_{\rm cv}$ will improve the CV estimate and thus ensure that the CV optimum is closer to the oracle optimum.

The first experiment demonstrates how the number of cross validation measurements affects the performance of CV. In this experiment signals of sparsity $K = 50$ are sampled and reconstructed using $M = 400$ measurements and three different reconstruction algorithms: OMP, Homotopy continuation, and Homotopy continuation with debiasing. The measurement noise is a white Gaussian process with $\sigma_{\mathbf{n}} = 2$. Figure 2 plots the results for a varying number of CV measurements $M_{\rm cv}$ – additional to the M sampling measurements – with the same measurement noise characteristics. We see that some small number of CV measurements causes a large improvement in performance – outperforming the residual criterion – with additional measurements not providing significant improvement. We also see that OMP outperforms the Homotopy-based algorithms.

The second experiment examines the tradeoff between allocating the measurements to reconstruction or CV. Assuming a fixed number of total measurements $M + M_{\rm cv} = 400$ and the same signal sparsity and noise parameters as in the previous experiment, we evaluate the performance of the three algorithms as $M_{\rm cv}$ varies from 5 to 100. For comparison, we also plot the performance of these algorithms assuming the same number of M measurements is used for reconstruction for two different cases: assuming an oracle guides the reconstruction and stops the algorithm optimally, and assuming the exact variance of the noise added is known and used to stop the algorithms, as described in Sec. 2.3. Note that the simulation uses the sample variance of the noise realization, not just the value of the variance parameter σ used to generate the noise; in practice this value is unknown at reconstruction. The results in Fig. 3 demonstrate that, although CV performance improves as $M_{\rm cv}$ increases, the reconstruction performance decreases as the reconstruction measurements M decrease, even if the optimal stopping iteration is known.

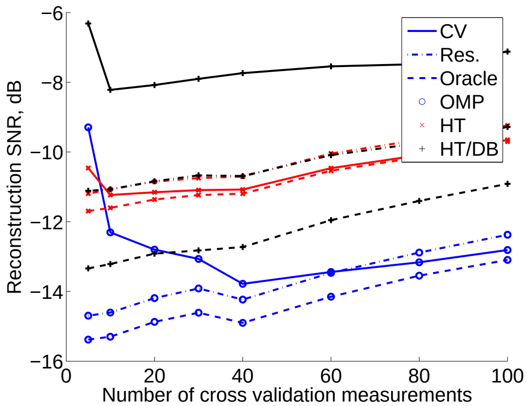

Fig. 3. *Tradeoff of measurements for CV and reconstruction for different algorithms. We fix the parameters $N = 1000$, $K = 50$ and $\sigma_{\mathbf{n}} = 2$. The total number of measurements is help constant at $M + M_{\rm cv} = 400$. Increasing the number of CV measurements improves the ability to detect the optimal stopping iteration. Increasing the number of reconstruction measurement increases robustness to the additive measurement noise.*

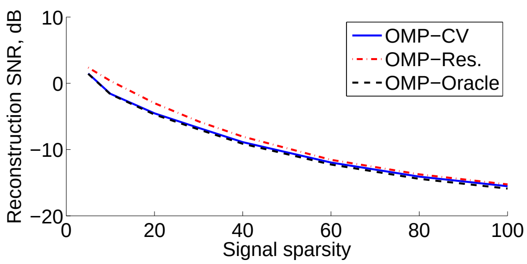

Fig. 4. *Performance of the CV performance criterion as a function of the signal sparsity. CV approaches the oracle performance as K decreases; the residual criterion does not.*

4.3. Parameter sensitivity

We examine the performance of CV-based CS algorithms to the other sundry parameters change. For presentation clarity, we focus on OMP for these experiments.

In the first experiment the sparsity K of the signal varies from 5 to 100. We use $M = 5K$ measurements for reconstruction and $M_{\rm cv} = 100$ measurements for CV. The noise variance is $\sigma_{\mathbf{n}} = 2$. Figure 4 shows that as the signal sparsity decreases, the CV performance approaches that of the oracle, while the difference in performance between the residual criterion and the oracle is roughly constant.

In the second experiment we vary the expected magnitude of the added noise $\sigma_{\mathbf{n}}$ from 0.2 to 200 and leave $K = 50$, $M = 360$ and $M_{\rm cv} = 40$. Figure 5 shows that the residual error stopping criterion performs better than CV only at very large noise variance, in which case all three stopping conditions perform inadequately in reconstructing the signal. In the range of practically acceptable levels of noise, CV consistently outperforms the residual error criterion.

5. CONCLUSIONS

This paper has proposed and explored CV for model selection in CS reconstruction from noisy measurements. Using a small subset of the noisy measurements to perform validation of the reconstruction, it is possible to obtain performance similar to that of standard reconstruction algorithms without requiring knowledge of the noise parameters

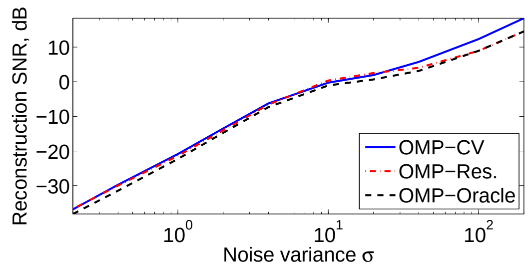

Fig. 5. *Performance of the CV performance criterion as a function of the noise magnitude. For the cases when the oracle obtains an acceptable SNR, CV outperforms the residual error criterion.*

or statistics. There is, however, a tradeoff between the use of additional measurements in the reconstruction algorithm to reduce the reconstruction error vs. the use of these measurements for CV to improve the detection of the optimal stopping iteration. In future work, we will explore CV-based schemes that rotate compressive measurements in and out of the training and test data sets, eventually using all the data to both estimate the parameters and cross validate the data.

6. REFERENCES

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

[2] E. J. Candès, "Compressive sampling," in *Proc. Int. Congress of Mathematicians*, Madrid, Spain, Aug. 2006, vol. 3, pp. 1433–1452.

[3] E. J. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.

[4] J. Tropp and A. C. Gilbert, "Signal recovery from partial information via orthogonal matching pursuit," Apr. 2005, Preprint.

[5] E. J. Candès and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," *Ann. Statistics*, 2006, To appear.

[6] J. Haupt and R. Nowak, "Signal reconstruction from noisy random projections," *IEEE Trans. Info. Theory*, vol. 52, no. 9, pp. 4036–4048, Sept. 2006.

[7] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projections for sparse reconstruction: Application to compressed sensing and other inverse problems," 2007, Preprint.

[8] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "A method for large-scale ℓ_1-regularized least squares problems with applications in signal processing and statistics," 2007.

[9] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. on Sci. Comp.*, vol. 20, no. 1, pp. 33–61, 1998.

[10] M. F. Duarte, M. B. Wakin, and R. G. Baraniuk, "Fast reconstruction of piecewise smooth signals from random projections," in *Proc. SPARS05*, Rennes, France, Nov. 2005.

[11] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Royal Statist. Soc.*, vol. 58, 1996.

[12] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Annals of Statistics*, vol. 32, no. 2, pp. 407–499, Apr. 2004.

[13] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, no. 12, 1993.

[14] M. R. Osborne, B. Presnell, and B. A. Turlach, "A new approach to variable selection in least squares problems," *IMA J. Numerical Analysis*, vol. 20, no. 3, pp. 389–403, 2000.

[15] D. L. Donoho and J. Tanner, "Thresholds for the recovery of sparse solutions via ℓ_1 minimization," in *Proc. Conf. Info. Sci. Sys.*, Princeton, NJ, March 2006.