# LIE OPERATORS FOR COMPRESSIVE SENSING

*Chinmay Hegde*[†], *Aswin Sankaranarayanan*[‡], *Richard Baraniuk*[*]

[†]Massachusetts Institute of Technology, Cambridge MA, USA
[‡]Carnegie Mellon University, Pittsburgh PA, USA
[*]Rice University, Houston TX, USA

## ABSTRACT

We consider the efficient acquisition, parameter estimation, and recovery of signal ensembles that lie on a low-dimensional manifold in a high-dimensional ambient signal space. Our particular focus is on randomized, compressive acquisition of signals from the manifold generated by the transformation of a base signal by operators from a Lie group. Such manifolds factor prominently in a number of applications, including radar and sonar array processing, camera arrays, and video processing. Leveraging the fact that Lie group manifolds admit a convenient analytical characterization, we develop new theory and algorithms for: (1) estimating the Lie operator parameters from compressive measurements, and (2) recovering the base signal from compressive measurements. We validate our approach with several of numerical simulations, including the reconstruction of an affine-transformed video sequence from compressive measurements.

## 1. INTRODUCTION

In recent years, the notion of *sparsity* as a model for data has come to the fore. Data vectors of length-$N$ are said to be $K$-sparse if only $K$ of their $N$ entries are nonzero; geometrically, the set of all sparse vectors can be identified with a particular *union of $K$-dimensional subspaces* in $\mathbb{R}^N$. Sparsity lies at the heart of Compressive Sensing (CS), an emergent alternative to the classical Shannon/Nyquist framework for signal acquisition. CS theory stipulates that instead of recording each of the $N$ entries (or samples) of a data vector $x \in \mathbb{R}^N$, the data acquisition system records the values obtained by performing $M \ll N$ inner products (or *measurements*) with $x$, such that $y = \Phi x$ where $\Phi \in \mathbb{R}^{M \times N}$. If $x$ is $K$-sparse in a known orthonormal basis $\Psi$, then it can be *exactly recovered* from the $M$ measurements $y$ provided that the entries of the vectors $\phi_i$ are drawn independently from certain probability distributions and $M = O(K \log N/K)$; further, this recovery can be performed efficiently in polynomial time. [1, 2].

Thus, CS can be viewed as a data acquisition framework that exploits a low-dimensional geometric model (sparsity) to motivate low-complexity sampling and reconstruction using random measurements. However, in several applications, the geometry of the data can be better modeled by a nonlinear low-dimensional *manifold*. Manifolds have been used as effective models in a range of machine learning applications, including supervised and semi-supervised classification and regression on different data types including images, audio, and text. The problem of specializing CS to manifold-modeled data has been examined in the literature both from a theoretical as well as an algorithmic standpoint [3–7].

In this paper, we focus our attention on geometric models in which the nonlinear manifold can be specified by matrix transformation operators, or *Lie operators*. Manifolds defined by Lie operators are widely encountered in practice. For instance, such models can be used for families of 1D signals formed by shifting, scaling and dilating a fixed base signal $x_0$ (applicable to radar, sonar and antenna arrays); families of 2D images subjected to transformations such as translation, rotation, scaling, and illumination change (applicable to multi-view camera networks); and a temporally varying sequence of images (applicable to video acquisition) [8,9]. Specifically, we study and solve two problems pertaining to CS sampling and recovery in this context:

- **Parameter estimation**. Given a base signal $x_0$, a manifold $\mathcal{L}$ parameterized by a Lie group, and compressive measurements $y$ of a signal $x = T(\mathbf{z})x_0 \in \mathcal{L}$ we develop an algorithm to accurately estimate the unknown manifold parameter vector $\mathbf{z}$. Our algorithm is algebraic in nature and involves solving a system of multivariate polynomials. We study the case of a one-dimensional (1D) manifold of shifted signals, and show how it can be extended to higher dimensional manifolds.
- **Base signal estimation**. Given parameter values $\mathbf{z}_j,\ j = 1, \ldots, J$, and corresponding compressive measurements of the corresponding signals $x_j$, we develop an algorithm to estimate the unknown base signal $x_0$. We also calculate the number of measurements required to stably recover the base signal.

We validate our approach with numerical simulations on both synthetic and real data studying the performance of our algorithms and their stability with respect to noise.

## 2. BACKGROUND

### 2.1. Lie Operators

Consider a data vector $x_0$ belonging to a space $\mathcal{S}$. Consider a (nonlinear) transformation of $x_0$ to another point $x \in \mathcal{S}$, so that $x = Tx_0$. Consider a family of *transformation operators* $\mathcal{T}$ such that $T \in \mathcal{T}$. In this paper, we focus on sets $\mathcal{T}$ which are endowed with an additional *group structure*, i.e., $\mathcal{T}$ satisfies the standard group axioms: (i) for $T_1, T_2 \in \mathcal{T}$, the composite operator $T_1 \circ T_2$ also belongs to $\mathcal{T}$; (ii) $\mathcal{T}$ contains the identity transformation $I$; (iii) the transformations in $\mathcal{T}$ are associative, i.e., $(T_1 \circ T_2) \circ T_3 = T_1 \circ (T_2 \circ T_3)$; (iii) for every $T \in \mathcal{T}$, there exists $T' \in \mathcal{T}$ such that $T \circ T' = T' \circ T = I$.

Suppose further that the space of transformations $\mathcal{T}$ is differentiable with respect to a suitably-defined distance measure $d_{\mathcal{T}}(\cdot, \cdot)$. Then, $\mathcal{T}$ assumes the structure of a differentiable manifold and is termed a *Lie group* [10]. We restrict our attention to finite dimensional Lie groups, i.e., we assume that $\mathcal{T}$ can be constructed by a finite set of transformation operators $\{T_1, \ldots, T_L\} \subset \mathcal{T}$ (elements of this finite set are also known as *generators*). Consider the set

$$\mathcal{L} = \{x \in \mathcal{S} \mid x = Tx_0,\ T \in \mathcal{T}\}. \tag{1}$$

We can identify $\mathcal{L}$ as being an $L$-dimensional manifold embedded in $\mathcal{S}$. For the rest of the paper, we will concern ourselves with finite-dimensional data vectors ($\mathcal{S} = \mathbb{R}^N$), so that the Lie transformation operators $T$ reduce to *matrix operators* of size $N \times N$.

Suppose that the Lie group $\mathcal{T}$ can be parameterized by a $L$-dimensional vector $\mathbf{z} = (\mathbf{z}_1, \ldots, \mathbf{z}_L)$; thus, $T = T(\mathbf{z})$ for each $T \in \mathcal{T}$. Then, the group structure endowed on $\mathcal{T}$ enables the following convenient formula: $x = \left(\prod_k e^{z_k G_k}\right) x_0$, where $G_k$ represents the matrix logarithm of the $k^{\text{th}}$ generator $T_k$ [8]. Note that the transformation operators will not be commutative in general. This restriction, in turn, is captured in the non-commutativity of the terms in the matrix exponential.

### 2.2. Compressive Sensing (CS)

A data vector $x \in \mathbb{R}^N$ is said to be $K$-sparse in the orthonormal basis $\Psi \in \mathbb{R}^{N \times N}$ if the basis representation $u = \Psi^T x$ has only $K$ nonzeros. Classical data acquisition has been largely based on the Shannon/Nyquist paradigm, where the signal $x$ is first sampled (or recorded) at a rate specified by its Fourier bandwidth, and subsequently compressed using a nonlinear encoding scheme. Compressive Sensing (CS) obviates the need for this "two-step" procedure in the following manner. Suppose we create a matrix $\Phi \in \mathbb{R}^{M \times N},\ M < N$, such that the elements of $\Phi$ are drawn randomly and independently from certain probability distributions (such as a Gaussian or a Bernoulli distribution). Then, a CS acquisiton system records inner products (or measurements) $y_i = \langle \phi_i, x \rangle,\ i = 1, \ldots, M$ (where $\phi_i$ denotes the $i^{\text{th}}$ row of $\Phi$), without ever having access to the full original signal $x$.

The key premise in CS is that if $x$ were $K$-sparse in a given basis $\Psi$, then with merely a number of measurements $M = O(K \log \frac{N}{K})$, the full signal $x = \Psi u$ can be *stably recovered in polynomial time* [1, 2]; in fact, this is achieved by solving the $\ell_1$-optimization problem:

$$\min \|u\|_1, \text{ s.t. } y = \Phi \Psi u. \tag{2}$$

This core framework has been extended in various ways. In particular, parallel CS frameworks have been built for general union-of-subspaces models [11], structured sparsity models [12], and multi-signal models [13].

### 2.3. CS for Manifold Models

A limited amount of progress has been made in extending the CS acquisition framework to nonlinear manifold models. It is known that for data belonging to a $L$-dimensional submanifold $\mathcal{M}$ of $\mathbb{R}^N$, merely $M = O(L \log N)$ random projections are sufficient to preserve both Euclidean distances as well as geodesics between points in $\mathcal{M}$ [3,4]. The type of results rely on the values of certain geometric parameters (such as manifold curvature) that are hard to either analytically compute or estimate from training data.

The simplest type of algorithms for signal recovery typically rely on a variation of gradient descent [4]. Consequently, the risk of encountering local minima is severe, and guarantees on algorithm performance and convergence are hard to achieve. More sophisticated algorithms for manifold-based recovery assume the availability of an *orthogonal projection operator* onto the manifold [6, 7]; for general nonlinear manifolds, these can be rather hard to construct. A non-parametric Bayesian approach to manifold-based CS recovery has also been proposed in [14].

Each of the above-listed methods applies to generic manifold models. In contrast, we focus on the special case when the underlying manifolds possess richer additional structure (specifically, a Lie group structure). In the rest of the paper, we develop a CS framework for this specific case that is useful in a range of applications.

## 3. LIE OPERATORS FOR CS

Suppose a signal $x \in \mathcal{L} \subset \mathbb{R}^N$ is such that $x = T(\mathbf{z}) x_0$, where $\mathcal{L}$ is a submanifold of $\mathbb{R}^N$ and $T \in \mathcal{T}$ is a Lie operator. As in the usual CS setting, suppose we are only given access to $M \ll N$ compressive measurements of $x$:

$$y = \Phi x = \Phi T(\mathbf{z}) x_0. \tag{3}$$

A fundamental CS task is to recover $x$ from $y$. In the most general case, $\Phi$ has a nullspace of dimension $N - M$ and thus there is no unique solution for $x$. Nonetheless, the problem can be solved with the aid of some additional information. For example, the recovery scheme might possess knowledge of the base signal $x_0$, the Lie operator $T(\cdot)$, the parameter vector $\mathbf{z}$, or some combination of these quantities. We study two subproblems within this broad setup.

### 3.1. Parameter Estimation

Suppose that only the base signal $x_0$ and the generators of the Lie group $\mathcal{T}$ are known. Then, the CS recovery problem reduces to solving for the unknown parameter vector $\mathbf{z} = (z_1, \ldots, z_L)$ in (3). In the case of general manifolds, this is a hard problem, and optimization-based tools such as Newton's method (or gradient descent) will often fail to converge to the correct solution.

Alternatively, consider the $m^{\text{th}}$ measurement $y_m$. Then,

$$y_m = \langle \phi_m, \prod_{k=1}^{L} e^{z_k G_k} x_0 \rangle = \langle \phi_m, \prod_{k=1}^{L} V_k e^{z_k \Lambda_k} V_k^{-1} x_0 \rangle,$$

where $V_k \Lambda_k V_k^{-1}$ corresponds the eigendecomposition of $G_k$. Note that $\Lambda_k$ is a diagonal eigenvalue matrix and thus $e^{D_k}$ is well-defined. Represent the $N$ diagonal entries of $\Lambda_k$ by $\lambda_{kj},\ j = 1, 2, \ldots, N$. Then, we observe that for any fixed $\lambda_{kj}$, $y_m$ is *linear* in $e^{\lambda_{kj} z_k}$. Consequently, $y_m$ is a *multi-linear* function of the exponential terms $e^{\lambda_{kj} z_k}$:

$$y_m = \sum_i c_i \prod_{j,k} e^{\lambda_{kj} z_k}. \tag{4}$$

The coefficients $c_i$ can be computed in closed form by expanding the matrix equations in each eigenvalue decomposition and collecting-ing the corresponding terms. We can further simplify (4) as follows: we perform a Taylor series expansion of each exponential term, and truncate each expansion to $d$ terms. Thus, we obtain a multivariate polynomial in $\mathbf{z}$ of total degree $Ld$:

$$y_m = \sum_\alpha c_\alpha z^\alpha, \ \ m = 1, \ldots, M, \tag{5}$$

where the sum is over non-negative integer $L$-tuples $(\alpha_1, \ldots, \alpha_L)$ such that $\sum_{i=1}^{L} \alpha_i \leq Ld$. Thus, for every measurement $y_m$, we obtain a new multivariate polynomial in the unknown parameter vector $\mathbf{z}$.

We have reduced the general nonlinear system of equations (3) to a system of $M$ polynomials in $L$ variables. Solving such a system of equations can be carried out, for example, by computing a Groebner basis for the *ideal* generated by the $M$ polynomials using Buchberger's algorithm, followed by back substitution [15]. We omit the details of the complex machinery used to solve the above polynomial system, and instead remark that thanks to recent advances in computational algebraic geometry, accurate software solutions for such problems exist in packages such as MAPLE.

The set of common solutions satisfying the system of $M$ polynomial equations is called an *affine variety*. It is known that the dimension of the affine variety of a system of $M$ polynomial equations (5) in $L$ variables is at most $L-M$; if $L = M$ and the coefficients of the polynomial system are random, then the set of solutions is of dimension 0 or finite [16]. Therefore, $M = L+1$ equations are in fact sufficient to *uniquely* determine the parameter vector $\mathbf{z}$. As opposed to general descent-based methods on the manifold, this approach is guaranteed to produce accurate parameter estimates. However, it could still be computationally expensive (akin to the $\ell_0$ minimization problem in standard CS). The worst case complexity bounds for Buchberger's algorithm are doubly exponential in the number of variables $L$. However, for small values of $L$ ($\leq 3$) this approach is still viable.

To provide an illustration of our approach, we consider the simple case where the manifold of interest is 1-dimensional, i.e., $L = 1$. An example of such a manifold is the set of shifted versions of a base 1D signal. In such cases, the Lie group $\mathcal{T}$ is specified by a single matrix transformation operator $T$ and a corresponding matrix logarithm $G$. Given random compressive measurements of a vector $x \in \mathcal{T}$, we have that

$$\begin{aligned} y_m &= \langle \phi_m, e^{zG} x_0 \rangle = \langle \phi_m, V e^{z\Lambda} V^{-1} x_0 \rangle \\ &= \langle \phi_m, \sum_{j=1}^{N} (v_j \widehat{v_j}^T x_0) e^{z\lambda_j} = \sum_{j=1}^{N} (\phi_m^T v_j)(\widehat{v_j}^T x_0) e^{\lambda_j z}, \end{aligned}$$

where $G = V \Lambda V^{-1}$ and $v_j,\ \widehat{v_j}$ denote the $j^{\text{th}}$ column of $V,\ V^{-1}$, respectively. Thus, the $m^{\text{th}}$ measurement equals a weighted linear sum of exponential functions in $z$. This sum of exponentials in turn can be approximated to arbitrary precision by a polynomial (of sufficiently high degree $d$) in $z$. The value of the parameter can subsequently be estimated by solving for the common real root of the system of $M$ polynomial equations.

### 3.2. Base Signal Estimation

Suppose now that the parameter vector $\mathbf{z}$ and the generators of the Lie group $\mathcal{T}$ are known, but that the base signal $x_0$ is unknown. Suppose also that the base signal is sparse in a known orthonormal basis $\Psi$. Further suppose that we observe compressive measurements $y = \Phi T(\mathbf{z}) x_0$. Then, CS recovery can be achieved by solving the $\ell_1$-optimization problem

$$\min \|u_0\|_1, \text{ s.t. } y = \Phi T(\mathbf{z}) \Psi u_0,$$

where $u_0 = \Psi^T x_0$.

This formulation is particularly useful in multi-signal acquisition scenarios as follows. Consider an ensemble of signals $x_j = T(\mathbf{z}_j) x_0,\ j = 1, \ldots, J$ lying on the manifold $\mathcal{L}$. Suppose we obtain $M$ measurements $y_j = \Phi_j x_j$ of each signal $x_j$. Adopting standard approaches [5, 17], we may stack the measurements and factor out a composite "measurement" matrix $\widetilde{\Phi}$ to obtain:

$$\widetilde{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_J \end{bmatrix} = \widetilde{\Phi} \begin{bmatrix} T(\mathbf{z}_1) \\ \vdots \\ T(\mathbf{z}_J) \end{bmatrix} x_0 = \widetilde{\Phi} \widetilde{T} \Psi u_0. \quad (6)$$

This can be again solved using $\ell_1$ methods, with the important distinction that the $NJ \times N$ matrix of Lie operators $\widetilde{T}$ (and correspondingly, the transformed sparsifying dictionary $\widetilde{T}\Psi$) is *no longer orthonormal*. Therefore, in general (6) is much harder to solve than standard CS recovery via $\ell_1$-optimization (2), and conventional bounds on the number of compressive measurements, as well as guarantees on algorithmic performance, no longer apply.

Nevertheless, recent theoretical results [18, 19] indicate that stable CS recovery in the case of redundant dictionaries using $\ell_1$ optimization is feasible under a certain restriction relating the maximum value of the dictionary *coherence*.[1] Thus, we arrive at the following proposition regarding base signal recovery.

**Proposition 1.** *Suppose that $x_0$ is $K$-sparse in a given orthonormal basis $\Psi$ and that we observe $M$ random Gaussian measurements $y_j = \Phi_j T(\mathbf{z}_j) x_0$ for each $j = 1, \ldots, J$. Let $\mu$ be the coherence of the matrix $\widetilde{T}$. If $K < 1 + 16\mu^{-1}$, then with high probability $x_0$ can be recovered by solving an $\ell_1$-optimization, provided $M = O\left(c(\mu) \frac{K}{J} \log \frac{JN}{K}\right)$.*

This proposition can be derived in a straightforward fashion by combining Theorem II.2 and Corollary II.4 in [18]. We make two remarks on Proposition 1. First, $M$ is inversely proportional to the number of signals $J$. Therefore, for sufficiently high $J$, we only need *a constant number of measurements for each signal*, irrespective of the sparsity (or complexity) of the base signal $x_0$. In scenarios (such as high-speed video acquisition) where the size of the signal ensemble is large and the budget for the number of measurements per signal is short, our proposed formulation can be particularly useful. Second, certain combinations of $T$ and $\mathbf{z}$ may result in a poorly conditioned $\widetilde{T}$, thus effectively negating the possibility of ever recovering $x_0$ irrespective of the number of measurements required. This intuition can be useful during acquisition system design, since a careful analysis of the operator matrix $\widetilde{T}$ may reveal "good" choices for the parameters $\mathbf{z}$ in order to favor better reconstruction of the overall ensemble of signals.

## 4. NUMERICAL EXPERIMENTS

### 4.1. Parameter Estimation

We test our parameter recovery algorithm via polynomial root-finding on a simple 1D manifold $\mathcal{L}$ of shifts of a randomly generated base signal $x_0$ of length $N = 128$. The base signal is circularly shifted by an amount $z = 6$ and compressively acquired using merely $M = 25$ random measurements. The shift parameter $z$ is then estimated using two methods: our proposed polynomial root finding method (with degree parameter $d = 20$), and gradient descent on the manifold $\mathcal{L}$ with an initial starting guess of $\widehat{z} = 0$. The signal is reconstructed using the estimated value of the shift.

Instead of adopting the computationally intensive Groebner basis approach, we computed the real roots of each polynomial equation (of degree $d$) and constructed a simple histogram of the set of all such estimated roots; the peak of the histogram was chosen as the estimate of the parameter $z$. The complexity of this algorithm is given by $O(Md^2 \log^2(d))$ [20]. The results of the estimation procedure are displayed in Fig. 1(a). Our method yields an accurate estimate of the shift, in contrast with the gradient descent, which gets stuck in a local minimum close to $z = 0$.

In order to quantitatively analyze our proposed algebraic approach for extracting parameter estimation from compressive measurements, we conduct a Monte-Carlo study consisting of $E = 100$

[1] The coherence $\mu$ is defined as the maximum absolute value of the pairwise inner products among all (normalized) columns of the dictionary and takes on values between 0 and 1.

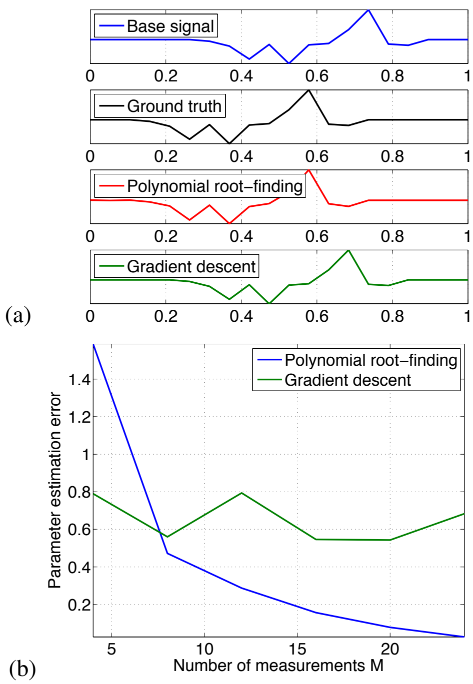



**Fig. 1**. *(a) Sample signal reconstruction results using different methods. (b) Monte Carlo performance comparison of parameter estimation using different methods.*

trials. For each trial, we generate a 1D base signal of length $N = 64$ with random coefficients and subject it to a (possibly non-integer) random shift $z$ chosen uniformly between 0 and $z_{\max} = 10$. We acquire $M$ compressive measurements, and assuming that the base signal (and correspondingly, the 1D manifold $\mathcal{L}$) is known, estimate the shift parameter $z$ from the measurements using our proposed polynomial root-finding approach, as well as gradient descent on the manifold with an initial shift guess of $z = 0$. The variation in the relative parameter estimation error with increasing number of measurements $M$ for the two methods is plotted in Fig. 1(b). Due to the highly nonconvex structure of $\mathcal{L}$, gradient descent often gets stuck in local minima, and its performance is poor even for high $M$. However, our proposed approach degrades gracefully in performance with increasing $M$ and gives fairly accurate results for $M > 20$.

### 4.2. Base Signal Estimation

We test our base signal estimation algorithm using signals belonging to a 1D manifold of shifted Gaussian pulses. We conduct a Monte Carlo run of $E = 100$ trials to study the quality of base signal reconstruction with the number of acquired measurements per signal. For each trial, we generated $J = 5$ signals that were random shifts of a truncated Gaussian pulse ($N = 512$, $K = 15$) and acquired $M$ measurements per signal. Then, we used (6) to estimate the base signal jointly from the aggregate measurements, and compared the reconstruction error with the average error obtained by performing independent estimation from each set of $M$ measurements. The gains of employing our approach can be visualized in Fig. 2, particularly for very low measurement rates $M/K$.

Finally, we test our approach on real images from the DARPA VIVID database. These images comprise a video sequence of $J =$

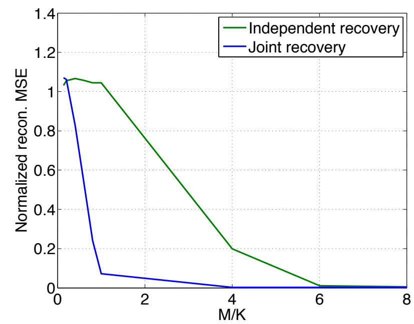
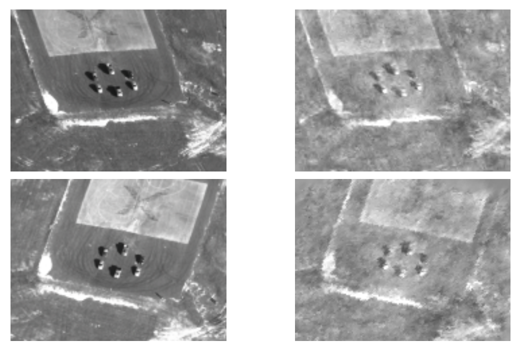



**Fig. 2**. *Monte Carlo performance comparison of base signal estimation using different methods. Our joint recovery formulation (Eq. 6) results in superior estimates than independent signal recovery, particularly for small values of measurement rates.*



Ground truth $\quad \frac{M}{N} = 0.15\%$, SNR = 15.97dB

**Fig. 3**. *CS recovery of UAV images using Lie operators.* $N = 160 \times 120 = 19200$, $J = 600$. *Accurate base signal reconstruction can be performed with measurement rates as low as* $M/N = 0.3\%$.

600 frames of size $N = 120 \times 160$ acquired by a camera mounted on an unmanned aerial vehicle (UAV). Due to the altitude of the imaging apparatus, the scene depth can approximately be modeled as constant across all pixels in all frames, and hence the each image can be approximately modeled as an affine transformation of a base image $x_0$ ($L = 6$). We assume that the affine parameters $\mathbf{z}$ are known for each image frame; these are precomputed using a simple affine registration algorithm. (In practice, these can be assumed to be available using an auxiliary source of information, such as a GPS or an accelerometer mounted on the UAV.) We convert the images to grayscale and compute $M$ measurements per frame and estimate the base signal using our proposed formulation (6), assuming that the base image is sparse in the 2D-wavelet basis specified by the Daubechies-4 filter. Once the base image is estimated, we can reconstruct each of the images via (1). Sample recovered images as well as average reconstruction SNR values are displayed in Fig. 3. We observe that excellent recovery of the base image can be obtained using just $M = 120$ measurements per frame, corresponding to a compression ratio $N/M$ of 160.

An interesting question is whether *both* parameter estimation and base signal estimation can be performed simultaneously from compressive measurements (via, say, alternating methods, similar to those developed in [21, 22]). We defer this to future work.

## 5. REFERENCES


[1] E. Candès, “Compressive sampling,” in *Proc. Int. Congress of Math.*, Madrid, Spain, Aug. 2006.

[2] D. Donoho, “Compressed sensing,” *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[3] R. Baraniuk and M. Wakin, “Random projections of smooth manifolds,” *Found. Comput. Math.*, vol. 9, no. 1, pp. 51–77, 2009.

[4] A. Efthekari and M. Wakin, “New analysis of manifold embeddings and signal recovery from compressive measurements,” Preprint, 2013.

[5] J. Park and M. Wakin, “A multiscale algorithm for reconstructing videos from streaming compressive measurements,” *J. Electronic Imaging*, vol. 22, no. 2, 2013.

[6] P. Shah and V. Chandrasekharan, “Iterative projections for signal identification on manifolds,” in *Proc. Allerton Conf. on Comm., Contr., and Comp.*, Monticello, IL, Sept. 2011.

[7] C. Hegde and R. Baraniuk, “Signal recovery on incoherent manifolds,” *IEEE Trans. Inform. Theory*, vol. 58, no. 12, pp. 7204–7214, Dec. 2012.

[8] X. Miao and R. Rao, “Learning the Lie groups of visual invariance,” *Neural Computation*, 2007.

[9] B. Culpepper and B. Olshausen, “Learning transport operators for image manifolds,” in *Proc. Adv. in Neural Processing Systems (NIPS)*, Dec. 2009.

[10] W. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press, London, England, 2003.

[11] Y. Eldar and M. Mishali, “Robust recovery of signals from a structured union of subspaces,” *IEEE Trans. Inform. Theory*, vol. 55, no. 11, Nov. 2009.

[12] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, “Model-based compressive sensing,” *IEEE Trans. Inform. Theory*, vol. 56, no. 4, pp. 1982–2001, 2010.

[13] S. Cotter, B. Rao, K. Engan, and K. Kreutz-Delgado, “Sparse solutions to linear inverse problems with multiple measurement vectors,” *IEEE Trans. Signal Processing*, vol. 53, no. 7, pp. 2477–2488, 2005.

[14] M. Chen, J. Silva, J. Paisley, C. Wang, D. Dunson, and L. Carin, “Compressive Sensing on Manifolds Using a Nonparametric Mixture of Factor Analyzers: Algorithm and Performance Bounds,” *IEEE Trans. Signal Processing*, 2010.

[15] J. Little and D. O’Shea, *Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra*, Springer, 2007.

[16] B. Sturmfels, *Solving systems of polynomial equations*, American mathematical society, 2002.

[17] A. Sankaranarayanan, P. Turaga, R. Chellappa, and R. Baraniuk, “Compressive acquisition of linear dynamical systems,” *SIAM J. Imaging Sci.*, vol. 6, no. 4, pp. 2109–2133, 2013.

[18] H. Rauhut, K. Schnass, and P. Vandergheynst, “Compressed sensing and redundant dictionaries,” *IEEE Trans. Inform. Theory*, vol. 54, no. 5, pp. 2210–2219, 2008.

[19] E. Candès, Y. Eldar, D. Needell, and P. Randall, “Compressed sensing with coherent and redundant dictionaries,” *Appl. Comput. Harmon. Anal.*, vol. 31, no. 1, pp. 59–73, 2010.

[20] M. Kim and S. Sutherland, “Polynomial root-finding algorithms and branched covers,” *Arxiv preprint math/9201280*, 1991.

[21] H. Braun, P. Turaga, C. Tepedelenlioglu, and A. Spanias, “Optical flow for compressive sensing video reconstruction,” in *IEEE Intl. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2013.

[22] A. C. Sankaranarayanan, C. Studer, and R. G. Baraniuk, “CS-MUVI: Video compressive sensing for spatial-multiplexing cameras,” in *IEEE Intl. Conf. Computational Photography (ICCP)*, 2012.