

REGULARIZATION OF GEOPHYSICAL INVERSION USING DICTIONARY LEARNING

Michael Bianco and Peter Gerstoft

Scripps Institution of Oceanography, UCSD
La Jolla, CA 92093-0238

ABSTRACT

Densely sampled dynamic geophysical data are often modeled using principal components analysis (PCA, a.k.a. empirical orthogonal function or EOF analysis) to provide constraints for their inversion with remote sensing techniques. We show that overcomplete sparsifying dictionaries, generated using dictionary learning, provide a more informative basis for geophysical signal representation. Relative to EOFs, all the vectors in learned dictionaries represent significant variance in the geophysical signals. Since many geophysical inverse problems are ill-posed, this behavior makes learned dictionaries ideal for both minimizing the solution dimension and improving the resolution of parameter estimates. The K-SVD algorithm is applied to ocean sound speed profile (SSP) data. It is shown that learned dictionaries improves SSP inversion resolution.

Index Terms— Dictionary learning, machine learning, inverse problems, geophysics, oceanography

1. INTRODUCTION

Dictionary learning algorithms may provide optimal regularization bases for geophysical inversion. Inversion for geophysical phenomena is typically ill-posed and requires significant regularization to obtain physically plausible solutions and moderate the size of the parameter search [1]. If many representative measurements are available, the dimension of the model typically is reduced using principal component analysis (PCA) [2]. PCA, or empirical orthogonal function (EOF, in the geosciences) analysis, provides a set of orthogonal shape functions which describe the features of largest variance in the data [2, 3]. However, this requirement of orthogonality may limit the regularization effectiveness.

Many signals, including natural images [4, 5], audio [6], and seismic profiles [7] are well approximated using sparsifying dictionaries. Given a signal, a dictionary is defined here as a set of ℓ_2 -normalized vectors which describe the signal using few coefficients. The sparse processor is then an ℓ_2 -norm cost function with an ℓ_0 -norm penalty on the number of non-zero coefficients [8]. Signal sparsity is exploited for a number

of purposes including signal compression and denoising [9]. Here, signal sparsity is exploited for inverse problem regularization of dynamic geophysical phenomena.

Dictionaries of vectors that approximate a given class of signals using few coefficients can be designed with dictionary learning. A popular dictionary learning approach, the K-SVD algorithm [9], based on vector quantization (VQ) [6], finds a dictionary of vectors that optimally partition the data from the training set such that the few dictionary vectors describe each data example. Relative to EOFs, the learned dictionary vectors are not constrained to be orthogonal and potentially provide more optimal signal compression because the vectors are on average, nearer to the signal examples [10]. Dictionary learning has been applied in the geophysics community to improve denoising results in seismics [11] and ocean acoustics [12, 13].

In this paper, dictionaries describing 1D ocean sound speed profile (SSP) data are generated using the K-SVD algorithm and the reconstruction performance is evaluated against EOF methods. In Section 2, the EOF methods and sparse processing are briefly introduced. In Section 3, dictionary learning and the K-SVD algorithm are described. In Section 4, SSP reconstruction results are given using the two methods. It is shown that each vector in the learned dictionaries explains more SSP variability than the leading order EOFs trained on the same data. Further, it is demonstrated that SSPs can be reconstructed up to acceptable error using as few as one non-zero coefficient. This compression can improve the resolution of ocean SSP estimates with negligible computational burden. More details of the approach and further experimental results are available in a forthcoming paper [14].

2. EOF AND SPARSE METHODS

EOF analysis seeks to reduce representation complexity of continuously sampled space-time fields by finding spatial patterns which explain much of the variance of the process. These spatial patterns or EOFs correspond to the principal components, from principal component analysis (PCA), of the temporally varying field [3]. Here, the field is a collection of ocean SSP anomaly vectors $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_M] \in \mathbb{R}^{K \times M}$ which are sampled over K discrete points in depth and M

This work is supported by the Office of Naval Research, Grant No. N00014-11-1-0439.

instants in time. They are defined as

$$\mathbf{y}_m = \mathbf{c}_m - \bar{\mathbf{c}} \quad (1)$$

where $\mathbf{c}_m \in \mathbb{R}^K$ is an observed K -point SSP and $\bar{\mathbf{c}} \in \mathbb{R}^K$ is the mean SSP over M observations.

The singular value decomposition (SVD) [15] finds the EOFs as the eigenvectors of $\mathbf{Y}\mathbf{Y}^T$ by

$$\mathbf{Y}\mathbf{Y}^T = \mathbf{P}\mathbf{\Lambda}^2\mathbf{P}^T, \quad (2)$$

where $\mathbf{P} = [\mathbf{p}_1, \dots, \mathbf{p}_L] \in \mathbb{R}^{K \times L}$ are L EOFs (eigenvectors, $L = \min(M, K)$) and $\mathbf{\Lambda}^2 = \text{diag}([\lambda_1^2, \dots, \lambda_L^2]) \in \mathbb{R}^{L \times L}$ are the total variances of the data along the principal directions defined by the EOFs \mathbf{p}_l , with $\lambda_1^2 \geq \dots \geq \lambda_L^2$.

Since the leading-order EOFs often explain much of the variance in \mathbf{Y} , the representation of \mathbf{Y} can be compressed by retaining only the leading order EOFs $P < K$ for reconstruction of \mathbf{y}_m . Each of the SSP anomaly vectors \mathbf{y}_m are approximated as $\mathbf{y}_m = \sum_{p=1}^P \mu_p \mathbf{p}_p$, where the EOF coefficients μ_p are solved as $\mu_p = \mathbf{p}_p^T \mathbf{y}_m$. For ocean SSPs, usually no more than 5 EOF coefficients have been used to reconstruct ocean SSPs [16, 17].

A signal \mathbf{y}_m , whose model is sparse in the dictionary $\mathbf{Q} = [\mathbf{q}_1, \dots, \mathbf{q}_N] \in \mathbb{R}^{K \times N}$, is reconstructed to an acceptable error using few vectors \mathbf{q}_n [8]. The inversion for these sparse coefficients is phrased as an ℓ_2 -norm minimization problem with an ℓ_0 -norm penalization on the number of non-zero coefficients

$$\hat{\mathbf{x}}_m = \arg \min_{\mathbf{x}_m \in \mathbb{R}^N} \|\mathbf{y}_m - \mathbf{Q}\mathbf{x}_m\|_2 \quad \text{subject to} \quad \|\mathbf{x}_m\|_0 \leq T, \quad (3)$$

where $\mathbf{x}_m \in \mathbb{R}^N$ is the vector of coefficients for \mathbf{Q} , $\hat{\mathbf{x}}_m$ is the sparse estimate of \mathbf{x}_m , and T is the number of non-zero coefficients in the solution ($T \ll K$). The sparse reconstruction $\hat{\mathbf{y}}_m \in \mathbb{R}^K$ of the signal \mathbf{y}_m is then $\hat{\mathbf{y}}_m = \mathbf{Q}\hat{\mathbf{x}}_m$. The ℓ_0 -norm constraint is non-convex and imposes combinatorial search for the exact solution to (3). Here, orthogonal matching pursuit (OMP) [18] is used as the sparse solver.

3. DICTIONARY LEARNING WITH K-SVD

The K-SVD algorithm [9] is inspired by the iterative K-means for VQ codebook design [6]. The N columns of the dictionary \mathbf{Q} , like the entries in VQ codebooks, correspond to partitions in \mathbb{R}^K . However, they are constrained to have unit ℓ_2 -norm and thus separate the magnitude (coefficients \mathbf{x}_n) from the shapes (dictionary entries \mathbf{q}_n) for the sparse processing objective (3). When $T = 1$, the ℓ_2 -norm objective in (3) is minimized by the dictionary entry \mathbf{q}_n that has the greatest inner product with example \mathbf{y}_m [8]. Thus for $T = 1$, $[\mathbf{q}_1, \dots, \mathbf{q}_N]$ define radial partitions of \mathbb{R}^K . This corresponds to a special case of VQ, called gain-shape VQ [6]. For $T = 1$, the sequential updates of the K-SVD provide optimal dictionary updates for gain-shape VQ [6, 9]. Optimal updates to

the gain-shape dictionary will, like K-means updates, either improve or leave unchanged the MSE and convergence to a local minimum is guaranteed. For $T > 1$, convergence of the K-SVD updates to a local minimum depends on the accuracy of the sparse-solver used in the sparse coding stage [9].

The dictionary learning objective is

$$\min_{\mathbf{Q}} \left\{ \min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{Q}\mathbf{X}\|_{\mathcal{F}}^2 \quad \text{subject to} \quad \forall m, \|\mathbf{x}_m\|_0 \leq T \right\}, \quad (4)$$

where $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_M]$ is the matrix of coefficient vectors corresponding to examples $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_M]$, and \mathcal{F} is the Frobenius norm. The K-means algorithm is generalized to the dictionary learning problem as the two steps: 1) sparse coding and 2) dictionary update.

In the K-SVD algorithm, each iteration i sequentially improves both the entries $\mathbf{q}_n \in \mathbf{Q}^i$ and the coefficients in $\mathbf{x}_m \in \mathbf{X}^i$, without change in support. Expressing the coefficients as row vectors $\mathbf{x}_T^n \in \mathbb{R}^N$ and $\mathbf{x}_T^j \in \mathbb{R}^N$, which relate all examples \mathbf{Y} to \mathbf{q}_n and \mathbf{q}_j , respectively, the ℓ_2 -penalty from (4) is rewritten as

$$\|\mathbf{Y} - \mathbf{Q}\mathbf{X}\|_{\mathcal{F}}^2 = \left\| \mathbf{Y} - \sum_{n=1}^N \mathbf{q}_n \mathbf{x}_T^n \right\|_{\mathcal{F}}^2 = \|\mathbf{E}_j - \mathbf{q}_j \mathbf{x}_T^j\|_{\mathcal{F}}^2, \quad (5)$$

where

$$\mathbf{E}_j = \left(\mathbf{Y} - \sum_{n \neq j} \mathbf{q}_n \mathbf{x}_T^n \right). \quad (6)$$

Thus, in (5) the ℓ_2 -penalty is separated into an error term $\mathbf{E}_j = [\mathbf{e}_{j,1}, \dots, \mathbf{e}_{j,M}] \in \mathbb{R}^{K \times M}$, which is the error for all examples \mathbf{Y} if \mathbf{q}_j is excluded from their reconstruction, and the product of the excluded entry \mathbf{q}_j and coefficients $\mathbf{x}_T^j \in \mathbb{R}^N$.

An update to the dictionary entry \mathbf{q}_j and coefficients \mathbf{x}_T^j which minimizes (5) is found by taking the SVD of \mathbf{E}_j , which provides the best rank-1 approximation of \mathbf{E}_j . However, many of the entries in \mathbf{x}_T^j are zero (examples which don't use \mathbf{q}_j). To update \mathbf{q}_j and \mathbf{x}_T^j with SVD, (5) must be restricted to examples \mathbf{y}_m which use \mathbf{q}_j

$$\|\mathbf{E}_j^R - \mathbf{q}_j \mathbf{x}_R^j\|_{\mathcal{F}}^2, \quad (7)$$

where \mathbf{E}_j^R and \mathbf{x}_R^j are entries in \mathbf{E}_j and \mathbf{x}_T^j , respectively, corresponding to examples \mathbf{y}_m which use \mathbf{q}_j , and are defined as

$$\mathbf{E}_j^R = \{\mathbf{e}_{j,l} \mid \forall l, x_l^j \neq 0\}, \quad \mathbf{x}_R^j = \{x_l^j \mid \forall l, x_l^j \neq 0\}. \quad (8)$$

The K-SVD algorithm is given in Table 1.

4. EXAMPLE

We here apply dictionary learning to ocean SSP data from the HF-97 acoustics experiment [19, 20], conducted off the coast of Point Loma, CA. The reconstruction results are compared

Given: $\mathbf{Y} \in \mathbb{R}^{K \times M}$, $\mathbf{Q}^0 \in \mathbb{R}^{K \times N}$, $T \in \mathbb{N}$, and $i = 0$

Repeat until convergence:

1. Sparse coding

for $m = 1 : M$

a: solve (3) for $\hat{\mathbf{x}}_m$ using OMP

b: $\mathbf{X} = [\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_M]$

2. Dictionary update

for $j = 1 : N$

a: compute reconstruction error \mathbf{E}_j from (6)

b: obtain \mathbf{E}_j^R , \mathbf{x}_R^j corresponding to nonzero \mathbf{x}_T^j

c: apply SVD to \mathbf{E}_j^R : $\mathbf{E}_j^R = \mathbf{U}\mathbf{S}\mathbf{V}^T$

d: update $\mathbf{q}_j^i = \mathbf{U}(:, 1)$, $\mathbf{x}_R^j = \mathbf{V}(:, 1)\mathbf{S}(1, 1)$

e: $\mathbf{Q}^{i+1} = \mathbf{Q}^i$

$i = i + 1$

Table 1: The K-SVD Algorithm [9]

with EOF methods. $M = 1000$ (15 point) profiles were used for the training set. The SSPs were interpolated to $K = 30$ points using a shape-preserving cubic spline. EOFs were calculated from (2) and learned dictionaries were generated with the K-SVD algorithm (Table 1). The number of non-zero coefficients solved with OMP for each dictionary was held fixed at exactly T non-zero coefficients. The initial dictionary \mathbf{Q}^0 was populated using randomly selected examples from the training sets \mathbf{Y} .

The HF-97 learned dictionary, with $N = K$ and $T = 1$, is compared to the EOFs ($K = 30$) in Fig. 1. Only the leading order EOFs (Fig. 1(a)) are informative of ocean SSP variability whereas all shape functions in the dictionary (Fig. 1(b)) are informative (Fig. 1(c)–(d)). By relaxing the requirement of orthogonality for the shape functions, the shape functions are adapted to the data distribution and thereby achieve greater compression. The Gram matrix \mathbf{G} , which gives the coherence of matrix columns, is defined for a matrix \mathbf{A} with unit ℓ_2 -norm columns as $\mathbf{G} = |\mathbf{A}^T \mathbf{A}|$. Fig. 1(e) shows the shapes in the EOF dictionary are orthogonal, whereas those of the learned dictionary (Fig. 1(f)) are not.

4.1. Reconstruction of SSP training data

Reconstruction performance of the EOFs and learned dictionaries are evaluated on SSPs within the training set, using a mean error metric. The coefficients for the learned \mathbf{Q} and initial \mathbf{Q}^0 dictionaries $\hat{\mathbf{x}}_m$ are solved from the sparse objective (3) using OMP. The least squares (LS) solution for the T leading-order coefficients $\mathbf{x}_L \in \mathbb{R}^T$ from the EOFs \mathbf{P} were solved by

$$\mathbf{x}_L = \mathbf{P}_L^\dagger \mathbf{y}_m, \quad (9)$$

where \mathbf{P}_L is the T leading order EOFs from \mathbf{P} , and \mathbf{P}_L^\dagger is its pseudoinverse. The best combination of T EOF coefficients was solved from the sparse objective (3) using OMP for $\mathbf{Q} =$

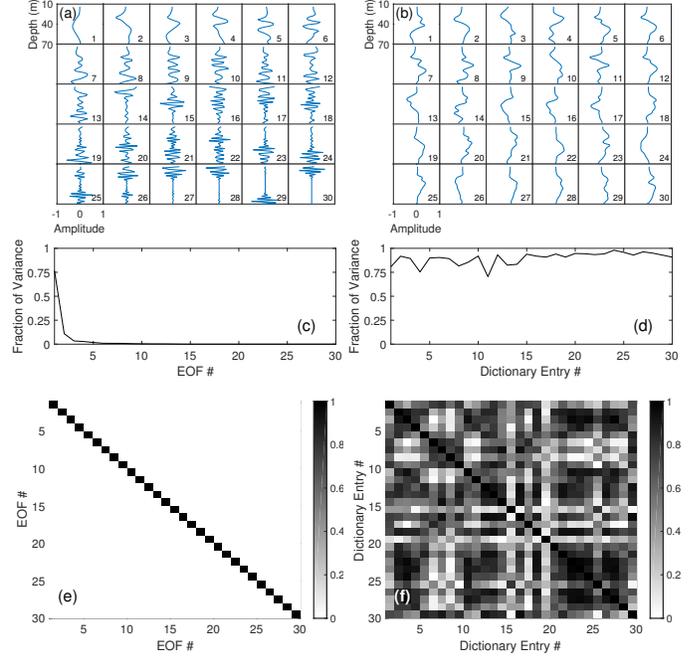


Fig. 1: (a) EOFs and (b) learned dictionary entries ($N = K$ and $T = 1$, sorted by variance $\sigma_{\mathbf{q}_n}^2$). Fraction of (c) total SSP variance explained by EOFs and (d) SSP variance explained for examples using dictionary entries. Coherence of (e) EOFs and (f) dictionary entries.

P. The mean reconstruction error ME for the training set is

$$\text{ME} = \frac{1}{KM} \|\mathbf{Y} - \hat{\mathbf{Y}}\|_1. \quad (10)$$

The reconstruction error using the EOF dictionary is compared to results from dictionaries \mathbf{Q} with $N = 3K$, using T non-zero coefficients. In Fig. 2(a) results are shown for $N = 90$ dictionary entries. Coefficients describing each example \mathbf{y}_m , were solved 1) from the learned dictionary \mathbf{Q} , 2) from \mathbf{Q}^0 , the dictionary consisting of N randomly chosen examples from the training set (to illustrate improvements in reconstruction error made in the K-SVD iterations), 3) the leading order EOFs, and 4) the best combination of EOFs. The mean SSP reconstruction error using the dictionaries trained for each sparsity T is less than EOF reconstruction, for either leading order coefficients or best coefficient combination, for all values of T shown. The best combination of EOF coefficients, chosen approximately using OMP, achieves less error than the LS solution to the leading order EOFs, with added cost of search.

Just one learned dictionary entry achieves the same ME as more than 6 leading order EOFs, or greater than 4 EOFs chosen by search (Fig. 2(a)). To illustrate the representational power of the learned dictionary entries, both true and reconstructed SSPs are shown in Fig. 3. Nine true SSP examples from the training set are reconstructed using one learned dictionary entry. It is shown for each case, that nearly all of the

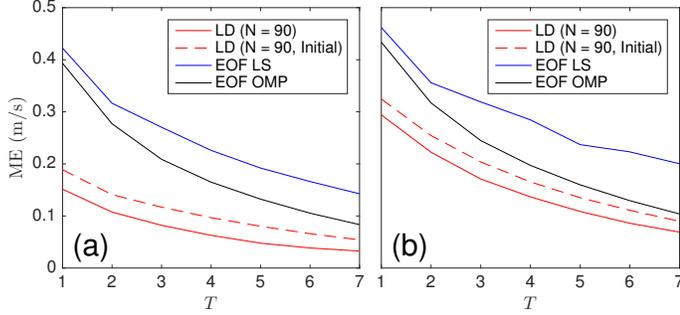


Fig. 2: (a) Mean reconstruction error (ME) versus T non-zero coefficients for SSPs using EOFs (solved using LS and OMP) and learned dictionary (LD, $N = 90$). (b) Mean reconstruction error ME_{CV} for extra-sample data calculated with K -fold cross validation for $J = 10$ folds.

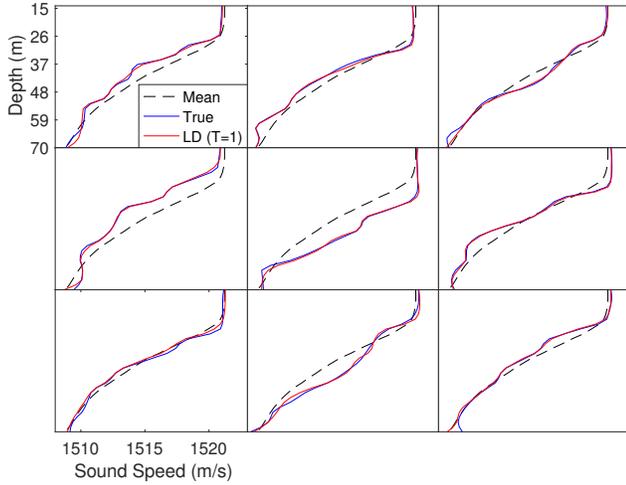


Fig. 3: True SSP reconstruction of 9 example profiles using one entry ($T = 1$) from learned dictionary (LD, $N = 90$).

SSP variability is captured using a single entry in \mathbf{Q} .

4.2. Extra-sample SSP reconstruction

The extra-sample SSP reconstruction performance of learned dictionaries and EOFs is tested using K -fold cross-validation [15]. The SSP data set \mathbf{Y} of M profiles is divided into J subsets with equal numbers of profiles $\mathbf{Y} = [\mathbf{Y}_1, \dots, \mathbf{Y}_J]$, where the fold $\mathbf{Y}_j \in \mathbb{R}^{K \times (M/J)}$. For each of the J folds: 1) \mathbf{Y}_j is the set extra-sample test cases, and the training set \mathbf{Y}_{tn} is

$$\mathbf{Y}_{tn} = \{\mathbf{Y}_l \mid \forall l \neq j\}; \quad (11)$$

2) the dictionary \mathbf{Q}_j and EOFs are derived using \mathbf{Y}_{tn} ; and 3) coefficients estimating test samples \mathbf{Y}_j are solved for \mathbf{Q}_j with sparse processor (3), and for EOFs by solving for leading order terms and by solving with sparse processor. The extra-sample error from cross validation ME_{CV} for each method is

then

$$ME_{CV} = \frac{1}{KM} \sum_{j=1}^J \|\mathbf{Y}_j - \hat{\mathbf{Y}}_j\|_1. \quad (12)$$

ME_{CV} increases over the within-training-set estimates for both the learned and EOF dictionaries, as shown in Fig. 2(b) for $J = 10$ folds. The mean reconstruction error using learned dictionaries, as in the within-training-set estimates, is less than the EOF dictionaries. More than 2 EOFs, choosing best combination by search, or more than 3 leading-order EOFs solved with LS, are required to achieve the extra-sample performance as one learned dictionary entry.

4.3. Solution space for SSP inversion

Acoustic inversion for ocean SSP is a non-linear problem. One approach is coefficient search using genetic algorithms. [2] Discretizing each coefficient into H values, the number of candidate solutions for T fixed coefficients indices is

$$S_{\text{fixed}} = H^T. \quad (13)$$

If the coefficient indices for the solution can vary, as per dictionary learning with learned dictionary $\mathbf{Q} \in \mathbb{R}^{K \times N}$, the number of candidate solutions S_{comb} is

$$S_{\text{comb}} = H^T \frac{N!}{T!(N-T)!}. \quad (14)$$

Given the results in the last paragraph of Section 4.1, and assuming a typical $H = 100$ point discretization of the coefficients and an unknown SSP similar to the training set, the SSP may be constructed up to acceptable resolution using one entry from the learned dictionary with 10^4 possible solutions. To achieve the similar ME, 6 EOFs are required (10^{12} possible solutions), using fixed indices. The best EOF combination requires 4 EOFs (10^{14} possible solutions).

5. CONCLUSION

Given sufficient training data, dictionary learning generates optimal dictionaries for sparse reconstruction of a given signal class. Since these learned dictionaries are not constrained to be orthogonal, the entries fit the distribution of the data such that signal example is approximated using few dictionary entries. Relative to EOFs, each dictionary entry is informative to the signal variability.

The K-SVD dictionary learning algorithm is applied to ocean SSP data from the HF-97 experiment. The learned dictionaries generated describe ocean SSP variability with high resolution using fewer entries than EOFs. As few as one entry from a learned dictionary describes nearly all the variability in each of the observed ocean SSPs. Provided sufficient SSP training data is available, learned dictionaries can improve SSP inversion resolution. This could provide improvements to geoacoustic inversion [2, 21], matched field processing [22], and underwater communications [19].

6. REFERENCES

- [1] R.C. Aster, B. Borchers, and C.H. Thurber, *Parameter estimation and inverse problems*, Elsevier, Oxford, 2nd edition, 2013.
- [2] P. Gerstoft, “Inversion of seismoacoustic data using genetic algorithms and *a posteriori* probability distributions,” *The Journal of the Acoustical Society of America*, vol. 95, no. 2, pp. 770–782, 1994.
- [3] A. Hannachi, I.T. Jolliffe, and D.B. Stephenson, “Empirical orthogonal functions and related techniques in atmospheric science: a review,” *International Journal of Climatology*, vol. 27, pp. 1119–1152, 2007.
- [4] A. Hyvärinen, J. Hurri, and P.O. Hoyer, *Natural Image Statistics: A Probabilistic Approach to Early Computational Vision*, Springer science and business media, 2009.
- [5] C. Christopoulos, A. Skodras, and T. Ebrahimi, “The JPEG2000 still image coding system: an overview,” *IEEE Trans. Cons. Elec.*, vol. 46, pp. 1103–1127, 2000.
- [6] A. Gersho and R.M. Gray, *Vector quantization and signal compression*, Kluwer Academic, Norwell, MA, 1991.
- [7] H.L. Taylor, S.C. Banks, and J.F. McCoy, “Deconvolution with the ℓ_1 -norm,” *Geophysics*, vol. 44, no. 1, pp. 39–52, 1979.
- [8] M. Elad, *Sparse and Redundant Representations*, Springer, New York, 2010.
- [9] M. Aharon, M. Elad, and A. Bruckstein, “K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation,” *IEEE Trans. Sig. Proc.*, vol. 54, pp. 4311–4322, 2006.
- [10] K. Engan, S.O. Aase, and J.H. Husøy, “Multi-frame compression: theory and design,” *Signal Processing*, vol. 80, pp. 2121–2140, 2000.
- [11] S. Beckouche and J. Ma, “Simultaneous dictionary learning and denoising for seismic data,” *Geophysics*, vol. 79, no. 3, pp. 27–31, 2014.
- [12] M. Taroudakis and C. Smaragdakis, “De-noising procedures for inverting underwater acoustic signals in applications of acoustical oceanography,” *EuroNoise 2015 Maastricht*, pp. 1393–1398, 2015.
- [13] T. Wang and W. Xu, “Sparsity-based approach for ocean acoustic tomography using learned dictionaries,” *OCEANS 2016 Shanghai IEEE*, pp. 1–6, 2016.
- [14] M. Bianco and P. Gerstoft, “Dictionary learning of sound speed profiles,” *ArXiv e-prints*, Sept. 2016, <https://arxiv.org/abs/1609.04840>.
- [15] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning: data mining, inference and prediction*, Springer, 2nd edition, 2009.
- [16] C. F. Huang, P. Gerstoft, and W. S. Hodgkiss, “Effect of ocean sound speed uncertainty on matched-field geoacoustic inversion,” *The Journal of the Acoustical Society of America Express Letters*, vol. 123, no. 6, pp. EL162–EL168, 2008.
- [17] P. Gerstoft and D.F. Gingras, “Parameter estimation using multifrequency range-dependent acoustic data in shallow water,” *The Journal of the Acoustical Society of America*, vol. 99, no. 5, pp. 2839–2850, 1996.
- [18] Y.C. Pati, R. Rezaifar, and P. S. Krishnaprasad, “Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition,” *27th. Annual Asilomar Conference on Signals, Systems and Computers, IEEE Proc.*, pp. 40–44, 1993.
- [19] N. Carbone and W.S. Hodgkiss, “Effects of tidally driven temperature fluctuations on shallow-water acoustic communications at 18kHz,” *IEEE Journal of Oceanic Engineering*, vol. 25, no. 1, pp. 84–94, 2000.
- [20] W.S. Hodgkiss, W.A. Kuperman, and D.E. Ensberg, “Channel impulse response fluctuations at 6 kHz in shallow water,” in *Impact of Littoral Environmental Variability of Acoustic Predictions and Sonar Performance*. Springer, 2002.
- [21] M. Bianco and P. Gerstoft, “Compressive acoustic sound speed profile estimation,” *The Journal of the Acoustical Society of America Express Letters*, vol. 139, no. 3, pp. EL90–EL94, 2016.
- [22] A.B. Baggeroer, W.A. Kuperman, and P.N. Mikhalevsky, “An overview of matched field methods in ocean acoustics,” *IEEE Journal of Oceanic Engineering*, vol. 18, no. 4, pp. 401–424, 1993.