

Interference-Driven Adaptation in Sparse Approximation

Bob L. Sturm and John J. Shynk
Department of Electrical and Computer Engineering
University of California
Santa Barbara, CA 93106-9560

Abstract—Sparse approximation attempts to find an efficient signal representation by adaptively building a signal vector space from elements of a usually redundant and overcomplete dictionary of atoms. Often, however, the representations produced by iterative descent methods, such as orthogonal matching pursuit (OMP), will contain atoms that are poorly chosen and are later confused to be features of the signal. Poorly selected atoms bring about the selection of other atoms that serve to correct for previous choices using destructive interference. This behavior diminishes the efficiency of a representation. In this paper, we propose and study a modification of the atom selection in OMP that takes into account the aforementioned effects. We find that a pursuit adapting to the interference between atoms can create a more efficient representation than that created by OMP. The representations created are more a representation of the signal and its features and less a reflection of the decomposition process.

I. INTRODUCTION

Recent works on signal processing and data analysis continue to demonstrate the benefits of sparse approximation (see, e.g., [1]–[4]). Applications include compressive sampling [5], coding and compression of audio [6] and image data [7], denoising and data recovery [8], and content discrimination and classification [9]–[11]. Sparse approximation often uses a highly redundant set (*dictionary*) of functions (*atoms*) to find representations of signals that are more efficient and robust than can be obtained by linear transformations. The freedom to choose the dictionary provides considerable flexibility, and much research has been devoted to the benefits of properly designing or choosing a dictionary (see, e.g., [1]–[4], [7], [8], [12]–[14]).

The characteristics of a pursuit algorithm are highly dependent upon the dictionary used and the decomposition process. A pursuit using a dictionary well-correlated with the structures in a particular signal can generate a very efficient representation; however, an iterative pursuit that “greedily” selects atoms can pay a penalty in efficiency, and create representations with obscured relationships to the features, or content, of a signal [4], [15]–[17]. If a low-order representation contains atoms that “correct” some aspects of other atoms, then the representation is less a reflection of the signal, and more a symptom that the pursuit and dictionary have not efficiently described the signal. When this occurs, the representation has a reduced utility for applications that depend on efficiency and low distortion, such as audio coding, or on clear relationships

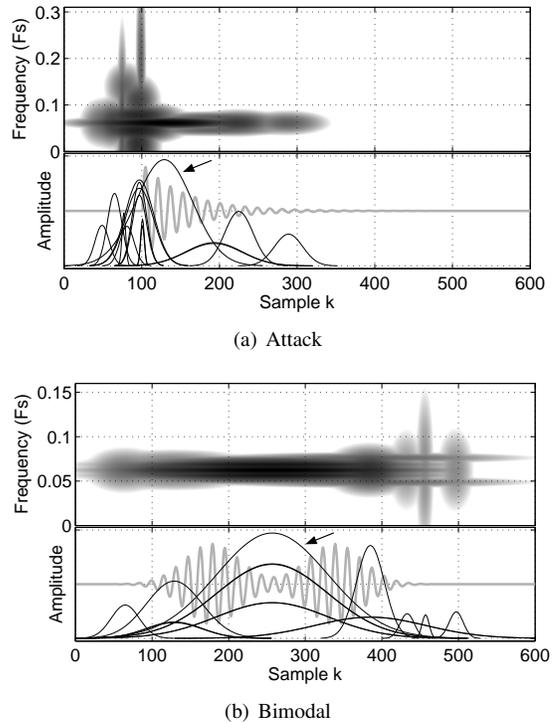


Fig. 1. Energy distribution in time-frequency domain (top), and the time-domain signal (gray) overlaid with exaggerated atom envelopes (black) of 15-order representations found using OMP.

between representation features and signal content, such as tasks of discrimination and classification.

Consider the two representations shown in Fig. 1, created by orthogonal matching pursuit (OMP) [2] using a redundant dictionary of Gabor atoms (modulated Gaussian functions) and Dirac functions (spikes). Observe that these representations contain atoms that function entirely in a corrective sense. In Fig. 1(a), atoms placed prior to the signal onset are a direct result of the left-hand tail of the large-scale atom (arrow). Similarly, in Fig. 1(b), the first atom selected (arrow) combines the two modes into one, which requires other atoms to create the “trough.”

Past approaches to the problem of erroneous features in sparse approximation have attempted to mitigate their production by altering the decomposition process, or by specifying dictionaries that are more similar to specific signal content. High-resolution pursuit [4] uses a dictionary in which large-

scale atoms consist of linear combinations of small-scale atoms (e.g., B-splines). Since each atom is selected based upon minimizing the worst fit of its smaller-scale constituent atoms, this approach places restrictions on the contents of a dictionary. One might use atoms that are more similar to specific structures in a signal, such as asymmetric damped sinusoids for speech [12], or modulated Gamma envelopes for transients [6]. One might also consider using a union of several dictionaries to provide more choices [18]; however, this can still lead to problems of correction, and make the pursuit more costly to perform. Ideally, we want a pursuit to efficiently represent only those signal structures that “make sense” using a dictionary, and omit the rest (e.g., those features that can not be represented without much correction).

In prior work [15]–[17], we have studied this phenomenon of correction for matching pursuit (MP), and have designed a measure of it called *interference*. In this paper, we motivate using the interference between atoms in a representation to inform the selection of atoms in a pursuit to avoid or postpone the correction of atoms. We demonstrate that taking interference into account, and thus adapting to the interference in the pursuit, can result in more efficient representations.

II. ORTHOGONAL MATCHING PURSUIT AND INTERFERENCE

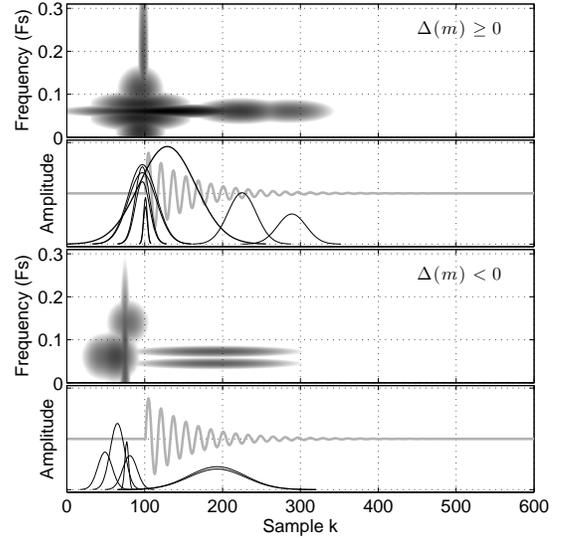
Consider a signal $\mathbf{x} \in \mathbb{C}^K$ and a dictionary $\mathcal{D}_N \triangleq \{\mathbf{d}_i \in \mathbb{C}^K : \|\mathbf{d}_i\|_2 = 1\}_{i=1}^N$, described in matrix form as $\mathbf{D} = [\mathbf{d}_1|\mathbf{d}_2|\dots|\mathbf{d}_N]_{K \times N}$ where $\text{rank}(\mathbf{D}) = K$ and usually $N \gg K$. We express the n th-order representation of \mathbf{x} as $X^n = \{\mathcal{H}_n, \mathbf{a}(n), \mathbf{r}(n)\}$, where \mathcal{H}_n is a set of n atoms selected from \mathcal{D}_N , written in matrix form as $\mathbf{H}(n) = [\mathbf{h}_0|\mathbf{h}_1|\dots|\mathbf{h}_{n-1}]_{K \times (n-1)}$. The column vector $\mathbf{a}(n) = [a_0, a_1, \dots, a_{n-1}]^T$ weights the selected atoms, and $\mathbf{r}(n)$ is the residual defined below. The n th-order approximation of \mathbf{x} is $\hat{\mathbf{x}}(n) \triangleq \mathbf{H}(n)\mathbf{a}(n)$. OMP updates the representation elements using the following rules:

$$X^n \rightarrow X^{n+1} : \begin{cases} \mathcal{H}_{n+1} = \{\mathcal{H}_n, \mathbf{h}_n\} \\ \mathbf{a}(n+1) = [\mathbf{G}(n+1)]^{-1} \mathbf{H}^H(n+1) \mathbf{x} \\ \mathbf{r}(n+1) = \mathbf{x} - \mathbf{H}(n+1) \mathbf{a}(n+1) \end{cases} \quad (1)$$

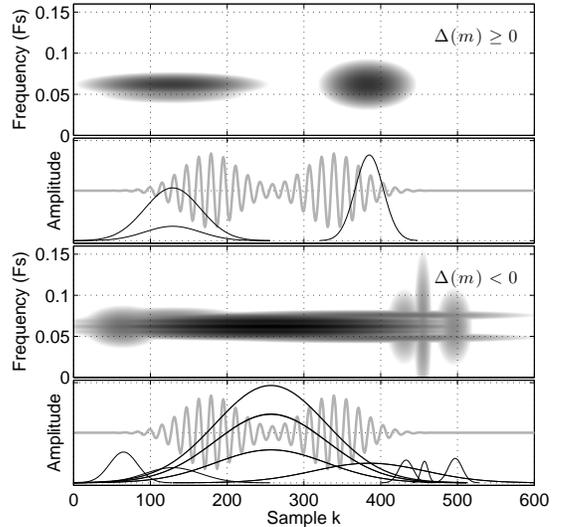
where $\mathbf{G}(n) \triangleq \mathbf{H}^H(n) \mathbf{H}(n)$ is the Gramian of the representation basis \mathcal{H}_n (the superscript H denotes complex conjugate vector transpose), and the OMP atom selection is as follows:

$$\mathbf{h}_n = \arg \max_{\mathbf{d} \in \mathcal{D}_N} |\mathbf{d}^H \mathbf{r}(n)|. \quad (2)$$

By using an overcomplete dictionary, it is generally the case that $\mathbf{G}(n)$ will not be an identity matrix, which implies that the atoms of a representation “interact” in a linear combination to form the approximation $\hat{\mathbf{x}}(n)$. In previous work, we have proposed an measure of this interaction [15]: the *interference* associated with the m th atom $\mathbf{h}_m \in \mathcal{H}_n$ of the n th-order representation $X^n = \{\mathcal{H}_n, \mathbf{a}(n), \mathbf{r}(n)\}$ of \mathbf{x} is defined as the



(a) Attack



(b) Bimodal

Fig. 2. Separation of atoms in the representations shown in Fig. 1 based on the sign of interference (3).

following energy:

$$\begin{aligned} \Delta(m) &\triangleq \|\hat{\mathbf{x}}(n)\|_2^2 - (\|\hat{\mathbf{x}}(n) - a_m \mathbf{h}_m\|_2^2 + \|a_m \mathbf{h}_m\|_2^2) \quad (3) \\ &= 2 \text{Real} \{a_m^* \mathbf{h}_m^H (\hat{\mathbf{x}}(n) - a_m \mathbf{h}_m)\} \quad (4) \end{aligned}$$

for $m = 0, 1, \dots, n-1$.

Using this definition of interference, we can classify how an atom of a representation interacts on average with other atoms in terms of energy. If $\Delta(m) < 0$, then the atom \mathbf{h}_m *destructively interferes* with the others; if $\Delta(m) > 0$, then it *constructively interferes* with the others; and if $\Delta(m) = 0$, then it is orthogonal to the $(n-1)$ th-order approximation. The separation of atoms of the representations in Fig. 1 based on the sign of the interference is shown in Fig. 2. Observe that for Attack, many of the destructively interfering atoms occur prior to the signal onset. These correct for the first atom selected,

which introduces energy into the residual signal. For Bimodal, all but three of the atoms of its representation are destructively interfering. This suggests that most of the terms in its 15-order representation serve more to correct the approximation rather than to represent the signal.

III. INTERFERENCE-DRIVEN ADAPTATION IN ORTHOGONAL MATCHING PURSUIT

The previous examples suggest that an efficient representation has few terms that require correction, and that an efficient pursuit will defer corrections until “coherent” aspects of the signal have been addressed by the pursuit, i.e., those features that can be represented without much correction. Let us inspect the interference associated with a new atom if it is selected by OMP in the updated representation X^{n+1} . From (1) and the iterative weight update rule in [2], observe that

$$\Delta(\mathbf{d}) \triangleq 2\text{Real} \left\{ a_{\mathbf{d}}^* \mathbf{d}^H \left([\mathbf{H}(n)] \mathbf{d} \begin{bmatrix} \mathbf{a}^{(n)} - a_{\mathbf{d}} \mathbf{b} \\ a_{\mathbf{d}} \end{bmatrix} - a_{\mathbf{d}} \mathbf{d} \right) \right\} \quad (5)$$

where

$$\mathbf{b} \triangleq [\mathbf{G}(n)]^{-1} \mathbf{H}^H(n) \mathbf{d} \quad (6)$$

$$a_{\mathbf{d}} \triangleq \frac{\mathbf{d}^H \mathbf{r}(n)}{\|\mathbf{d} - \mathbf{d}_{\mathcal{H}_n}\|_2^2} = \frac{\mathbf{d}_{\mathcal{H}_n^\perp}^H \mathbf{r}(n)}{\|\mathbf{d}_{\mathcal{H}_n^\perp}\|_2^2} \quad (7)$$

$$\mathbf{d}_{\mathcal{H}_n} \triangleq \mathbf{H}(n) \mathbf{b} \implies \mathbf{d}_{\mathcal{H}_n^\perp} \triangleq \mathbf{d} - \mathbf{d}_{\mathcal{H}_n}. \quad (8)$$

This expression simplifies to

$$\Delta(\mathbf{d}) = 2\text{Real} \{ a_{\mathbf{d}}^* \mathbf{d}^H \hat{\mathbf{x}}(n) \} - 2|a_{\mathbf{d}}|^2 \|\mathbf{d}_{\mathcal{H}_n}\|_2^2. \quad (9)$$

The interference of \mathbf{d} thus depends on its projection onto the current approximation $\hat{\mathbf{x}}(n)$, and the representation basis \mathcal{H}_n . For real atoms, this can be expressed as

$$\Delta(\mathbf{d}) = 2 \frac{\mathbf{d}^T \mathbf{r}(n)}{\|\mathbf{d}_{\mathcal{H}_n^\perp}\|_2^2} \left[\mathbf{d}^T \hat{\mathbf{x}}(n) - \frac{\|\mathbf{d}_{\mathcal{H}_n}\|_2^2}{\|\mathbf{d}_{\mathcal{H}_n^\perp}\|_2^2} \mathbf{d}^T \mathbf{r}(n) \right]. \quad (10)$$

In the pursuit of a representation, let us assume that atoms which are good candidates to be selected will likely satisfy $\|\mathbf{d}_{\mathcal{H}_n}\|_2^2 < \|\mathbf{d}_{\mathcal{H}_n^\perp}\|_2^2 \approx 1$. Let us also assume for these atoms that as the pursuit progresses, $|\mathbf{d}^T \mathbf{r}(n)|$ decreases. With these in mind, and since the sign of $\Delta(\mathbf{d})$ is more of interest than its exact value, we propose to simplify (10) to

$$\Delta'(\mathbf{d}) \triangleq 2 [\mathbf{d}^T \mathbf{r}(n)] \mathbf{d}^T \hat{\mathbf{x}}(n). \quad (11)$$

Essentially (11) indicates whether or not \mathbf{d} will be constructively or destructively interacting with the representation of \mathbf{x} , and approximately by what amount.

Finally, we may incorporate (11) into the atom selection criteria of OMP in (2) as a penalty in the following way:

$$\mathbf{h}_n = \arg \max_{\mathbf{d} \in \mathcal{D}_N} \left[|\mathbf{d}^T \mathbf{r}(n)| + \gamma(n) [\mathbf{d}^T \mathbf{r}(n)] \mathbf{d}^T \hat{\mathbf{x}}(n) \right] \quad (12)$$

where $-\infty < \gamma(n) < \infty$ is an iteration-dependent scalar that weights the importance of using interference in the atom selection rule. We see that when $\gamma(n) = 0$, the pursuit update in (1) is just OMP. When $\gamma(n) > 0$, the selection

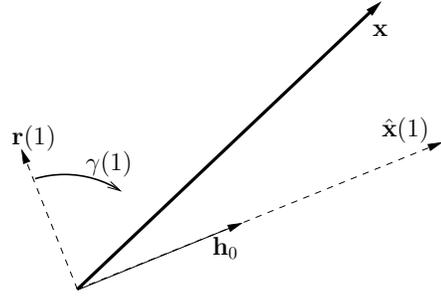


Fig. 3. Geometric interpretation of interference-driven adaptation in a greedy iterative descent pursuit.

rule in (12) will encourage atoms that constructively interfere with the others ($\Delta(\mathbf{d}) > 0$), and will discourage atoms that destructively interfere ($\Delta(\mathbf{d}) < 0$). The opposite case occurs when $\gamma(n) < 0$.

Figure 3 illustrates how the modified pursuit works. After the first iteration, OMP selects an atom that has the maximum projection (magnitude) onto the residual $\mathbf{r}(1)$. If a good match is found, it will be nearly orthogonal to the original signal \mathbf{x} . When $\gamma(1) > 0$, interference influences the atom selection by effectively pulling the residual toward the original signal \mathbf{x} and its first-order approximation $\hat{\mathbf{x}}(1)$. The next atom selected should thus be less orthogonal to the original signal. By selecting an atom based on its projection onto the residual and the approximation of the signal, this pursuit essentially takes into account the information about how it has already represented the signal.

IV. COMPUTER SIMULATIONS

In the following experiments, we have decomposed the signals shown in Fig. 1, as well as a sinusoid, using (1) with (12) and the same overcomplete dictionary of Gabor atoms and Dirac spikes used in Fig. 1. The interference weight used is constant:

$$\gamma(n) = c \quad \forall n \geq 0. \quad (13)$$

Figure 4 shows the signal-to-residual ratio (SRR) as a function of the pursuit iteration for several positive values of c . The benefits of incorporating interference in the atom selection is clearly manifested by a decrease in distortion using fewer terms, i.e., an increase in the efficiency of the representation. This effect is much more prominent for Attack than for Bimodal and Sine. It is clear with Bimodal that the first atom selected (arrow) spans the support of the entire signal, which necessitates correction around its center. Since $\mathcal{H}_0 = \emptyset$, there is no information present to guide the first atom selection.

The relationship between the distortion as a function of the number of atoms and the value of $\gamma(n)$ is obviously complex. Figure 5 shows the isolines of distortion in terms of the number of atoms required as a function of c (recall that OMP corresponds to the case of $c = 0$). For Attack, we see a dramatic difference in the number of atoms required to achieve a specific distortion. This difference grows as the distortion decreases. The other two signals do not show such

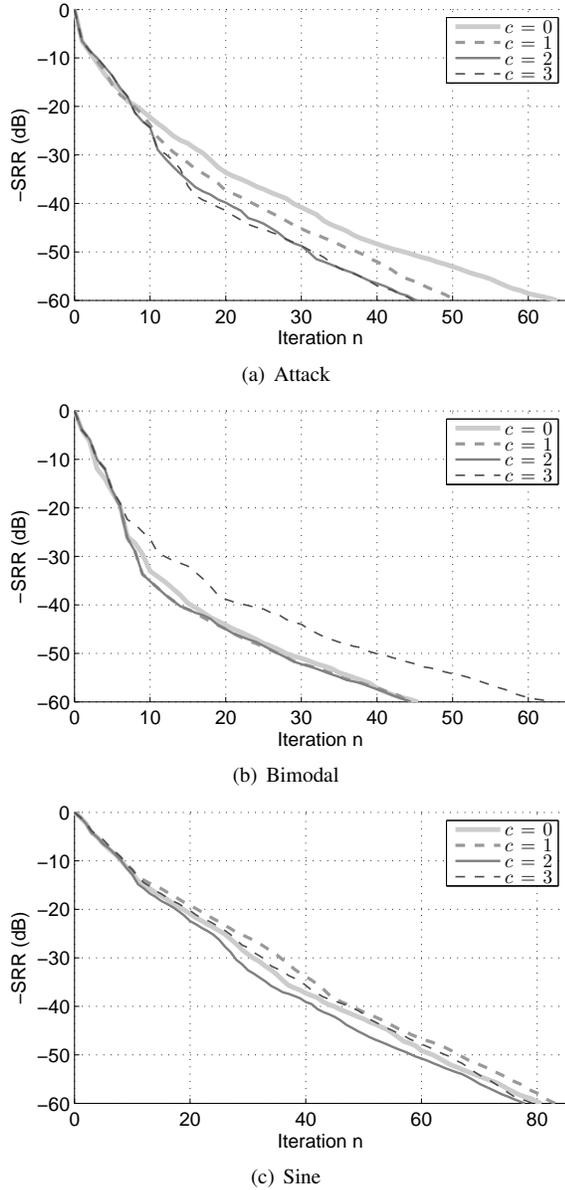


Fig. 4. Residual energy decay as a function of iteration using (12) with (13).

a large difference in terms of the number of atoms found at a given SRR compared with basic OMP. It is interesting to see that for Bimodal and $c < 0$, there exists a slightly more sparse solution for an SRR of 60 dB, than when $c = 0$.

The representations created by the modified pursuit can vary widely depending on the value of c , even if there is little change in the distortion. This can be seen in Fig. 6, which shows the time-domain envelopes of the atoms selected by the pursuit for several values of c . For Attack, we find that as c increases, the resulting atom distribution appears to adapt better than OMP to the sharp onset and subsequent decay of the signal. The error introduced by the first atom preceding the onset is lessened over time because its energy is reduced with the reprojection in (1) (it appears to be irrelevant to a constructive representation of \mathbf{x}). We observe this same effect

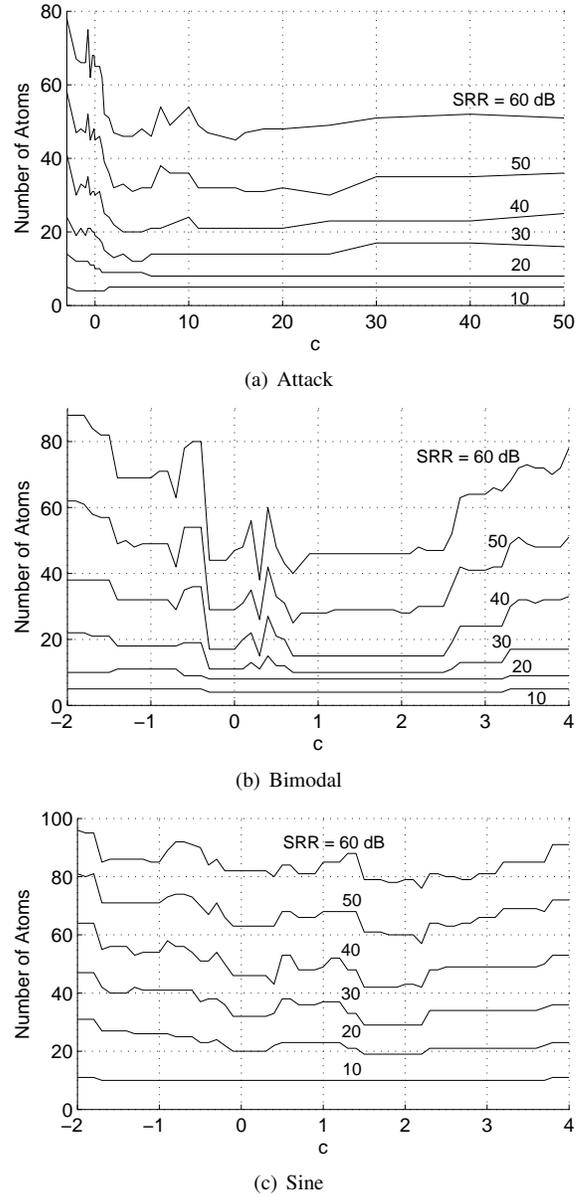


Fig. 5. Atoms found for SRR using (12) with (13) as a function of c .

for Bimodal, where the representation becomes more of a reflection of the signal structures. For the constant envelope Sine, the ideal behavior of overlapped atoms with uniform amplitudes is observed for $c = 2$.

V. CONCLUSION

We have presented a modification to the atom selection criteria in OMP that takes into account the fact that atoms selected from redundant and overcomplete dictionaries can result in less meaningful terms in the signal representation. By incorporating interference into the criteria used to select atoms, we effectively assign a cost to the selection based on their interaction with the current approximation. We have empirically shown that the resulting representations can be more efficient and representative of the signals than those

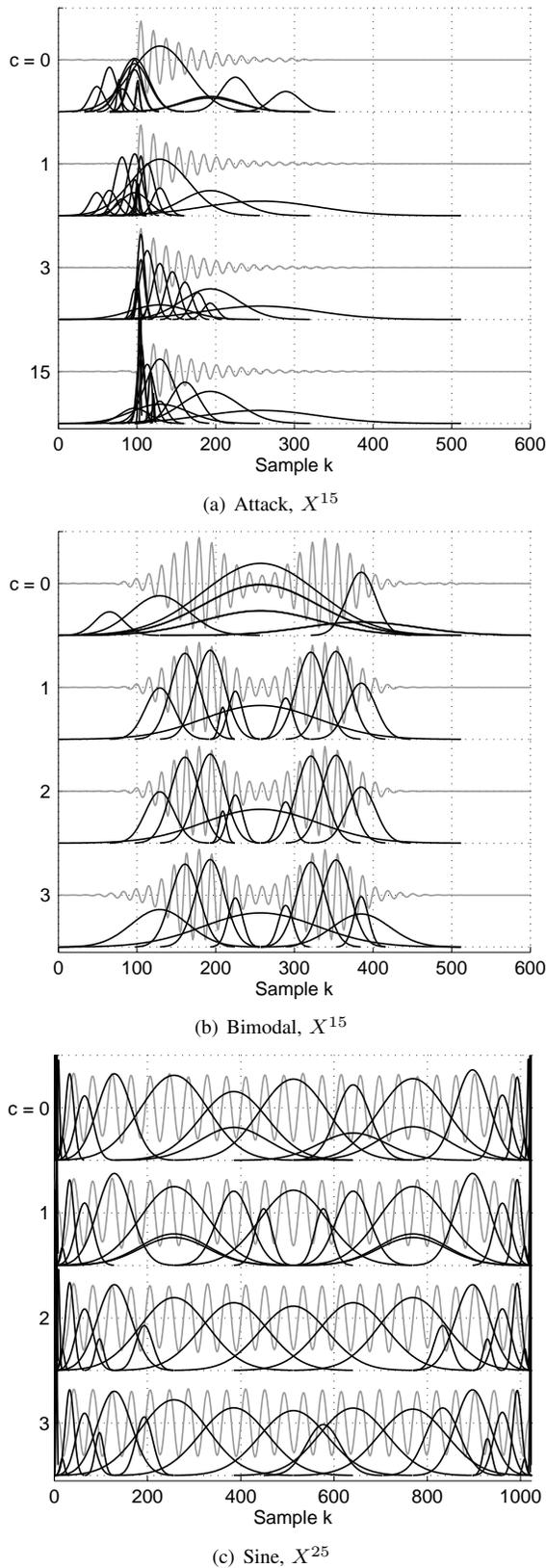


Fig. 6. Time-domain atom envelopes of representations created using (12) with (13) as a function of c . Atom envelopes are exaggerated using the square-root of the energy.

produced by OMP without interference-driven adaptation. Future work will look at increasing efficiency more by rejecting atoms from the basis based on their interaction with the signal representation, which we expect to aid cases like Bimodal.

ACKNOWLEDGMENTS

The authors would like to thank Mads Christensen for helpful discussions. This work was supported in part by the National Science Foundation under Grant CCF 0729229.

REFERENCES

- [1] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.
- [2] Y. Pati, R. Rezaifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 1993, vol. 1, pp. 40–44.
- [3] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, Aug. 1998.
- [4] S. Jaggi, W. C. Karl, S. Mallat, and A. S. Willsky, "High resolution pursuit for feature extraction," *Applied and Computational Harmonic Analysis*, vol. 5, no. 4, pp. 428–449, Oct. 1998.
- [5] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [6] M. G. Christensen and S. van de Par, "Efficient parametric coding of transients," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 4, pp. 1340–1351, July 2006.
- [7] R. M. Figueras i Ventura, P. Vandergheynst, and P. Frossard, "Low-rate and flexible image coding with redundant representations," *IEEE Trans. Image Process.*, vol. 15, no. 3, pp. 726–739, Mar. 2006.
- [8] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.
- [9] K. Umamathy, S. Krishnan, V. Parsa, and D. G. Jamieson, "Discrimination of pathological voices using a time-frequency approach," *IEEE Trans. Biomedical Eng.*, vol. 52, no. 3, pp. 421–430, Mar. 2005.
- [10] U. Malinowska, P. J. Durka, K. J. Blinowska, W. Szelensberger, and A. Wakarow, "Micro- and macrostructure of sleep EEG: A universal adaptive time-frequency parameterization," *IEEE Eng. Med. Biol. Mag.*, vol. 25, no. 4, pp. 26–31, July/Aug. 2006.
- [11] S. Chu, S. Narayanan, and C.-C. J. Kuo, "Environmental sound recognition using MP-based features," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, Las Vegas, NV, Mar. 2008, pp. 1–4.
- [12] M. Goodwin and M. Vetterli, "Matching pursuit and atomic signal models based on recursive filter banks," *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 1890–1902, July 1999.
- [13] M. Aharon, M. Elad, and A.M. Bruckstein, "K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov 2006.
- [14] P. Leveau, E. Vincent, G. Richard, and L. Daudet, "Instrument-specific harmonic atoms for mid-level music representation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 1, pp. 116–128, Jan. 2008.
- [15] B. L. Sturm, J. J. Shynk, L. Daudet, and C. Roads, "Dark energy in sparse atomic estimations," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 3, pp. 671–676, Mar. 2008.
- [16] B. L. Sturm, J. J. Shynk, and L. Daudet, "A short-term measure of dark energy in sparse atomic estimations," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2007, pp. 1126–1129.
- [17] B. L. Sturm, J. J. Shynk, and L. Daudet, "Measuring interference in sparse atomic estimations," in *Proc. Conf. Info. Sciences Syst.*, Princeton, NJ, Mar. 2008, pp. 961–966.
- [18] L. Daudet, "Sparse and structured decompositions of signals with the molecular matching pursuit," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 5, pp. 1808–1816, Sept. 2006.