8-2009

Sparse Representations in Power Systems Signals

Jack Cooper
Clemson University, jackcooper@gmail.com

Sparse Representations in Power Systems Signals

A Master's Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Mathematical Sciences

by
Jack Cooper
August 2009

Accepted by:
Dr. Taufiquar R Khan, Committee Chair
Dr. Irina V Viktorova
Dr. James Brannan

Abstract

This thesis seeks to detect transient disturbances in power system signals in a sparse framework. To this end, an overcomplete wavelet packet dictionary and damped sinusoid dictionary are considered, and for each dictionary Matching Pursuit is compared with Basis Pursuit. Previous work in developing waveform dictionary theory and sparse representation is reviewed, and simulations are run on a test signal in both noisy and noiseless environments. The solutions are viewed as time-frequency plane tilings to compare the accuracy and sparsity of these algorithms in properly resolving optimal representations of the disturbances. The advantages and disadvantages of each combination of dictionary and algorithm are presented.

Acknowledgments

The author acknowledges the committee members for their help, and also the creaters of the WaveLab and SparseLab software currently being developed at Stanford University. Without these excellent tools, much of the analysis would have been impossible.

Table of Contents

Title Page . . . . **i**

Abstract . . . . **ii**

Acknowledgments . . . . **iii**

List of Tables . . . . **vi**

List of Figures . . . . **vii**

1 Introduction **1**
1.1 The Signal 2
1.2 The Problem 3
1.3 The Dictionary 6
1.4 The Algorithms 8

2 Research Design and Methods **10**
2.1 Test Signals 10
2.2 Dictionary 11
2.3 Motivation for Pursuit Algorithms and Sparsity 14
2.4 Comparing Sparsity 16
2.5 Importance of the Study 16

3 Related Work **17**
3.1 Sparsity Development 17
3.2 Algorithm Development 20
3.3 Dictionary Development 20

4 Results **23**
4.1 Dictionary Coherence 23
4.2 Simulation without Noise 25
4.3 Simulation with Noise 29
4.4 Conclusion of Results 36

5 Conclusions and Discussion **37**
5.1 Answering the Research Questions 37
5.2 Discussion of future work 39

References **40**

List of Tables

4.1 Mutual Coherence Table 24

List of Figures

2.1 Tested Signal without Noise 12
2.2 Tested Signal with Noise 13
2.3 Coefficient vector of wavelet decomposition 15

4.1 Matching Pursuit on Wavelet Packet Dictionary without noise 26
4.2 Basis Pursuit on Wavelet Packet Dictionary without noise 27
4.3 Matching Pursuit on Damped Sinusoid Dictionary without noise . . . 28
4.4 Basis Pursuit on Damped Sinusoid Dictionary without noise 30
4.5 Matching Pursuit on Wavelet Packet Dictionary with noise 31
4.6 Basis Pursuit on Wavelet Packet Dictionary with noise 33
4.7 Matching Pursuit on Damped Sinusoid Dictionary with noise 34
4.8 Basis Pursuit on Damped Sinusoid Dictionary with noise 35

Chapter 1

Introduction

In recent years, the eyes of both industry and government turned toward the technological advancement of power systems to ease problems with growing energy needs, shrinking energy supplies, and irreversible environmental damage. A key component of this focus is power quality monitoring and fault detection. These transient faults need to be understood and characterized to properly identify and subsequently analyze the sources of disruption. Detecting transients within noise is a challenging task to automate, and much research has gone into this pursuit. Algorithms have been proposed to analyze the signal in the temporal domain, various transform domains, and adaptive pursuit algorithms as in this thesis.

In applied mathematics, the research area of sparse representation is also seeing an infusion of interest. A signal can be decomposed by Fourier series, wavelets, or other waveforms, usually with increasing accuracy as the the number of waveforms used increases. Yet, in addition to the accuracy of the representation it has become a priority to pursue a sparse representation. A sparse representation is one in which the signal is decomposed into only a few elements. This clearly results in a higher compression and less data storage, but sparse representation also ensures accurate extraction and sharper classification of the main signal elements.

Two adaptive pursuit algorithms, matching pursuit and basis pursuit, as pertaining to simulated power systems signals are investigated in this thesis. While the fields of power systems analysis and sparsity constraints emerge separately, this thesis seeks their marriage in an effort to produce an optimal method for analyzing power systems signals and faults.

1.1 The Signal

The power system signal is a complicated one because of its dynamic nature, dependent on constantly changing loads. The current model for power system phenomena is developed through [23, 27, 2]. In [23], power system signal is characterized by

1. Low frequency harmonics ranging from the system fundamental frequency (60Hz in the US) to 3kHz. These are caused by power electronic devices, arc furnaces, transformers, rotational machines, and aggregate loads.

2. Swells and Sags of the RMS Voltage from a half cycle of the fundamental up to a minute in duration.

3. Transients observed as impulses superimposed to the fundamental frequency. These can be caused by lightnings, equipment faults, switching operations, and more.

Essentially, power systems are built from transmission lines, sources, and loads. In these circuits the voltages closely follow the solutions of differential equations [17]. Because of this, damped sinusoids are proposed to model the signal. Due to their impulsive nature, this model reflects the transient disturbances well. The proposed model is

$$f(t) = \sum_{q=1}^{Q} A_q e^{-\lambda_q (t-t_0)} \cdot \sin(2\pi f_q (t - t_q) + \phi_q) \cdot u(t - t_q) \tag{1.1}$$

where A_q is the initial amplitude, f_q is the frequency in Hertz, λ_q is the damping factor, and ϕ is the phase angle, and t_q is the starting time of the q^{th} impulse. The function u(t) is zero before t_q and 1 during the impulse. In noise the model becomes

$$f(t) = \sum_{q=1}^{Q} A_q e^{-\lambda_q (t-t_0)} \cdot \sin(2\pi f_q (t - t_q) + \phi_q) \cdot u(t - t_q) + n(t) \tag{1.2}$$

where n(t) is white Gaussian noise. This thesis focuses on the detection of the transients modeled in this way.

1.2 The Problem

A dictionary of size p, D, is defined as a collection of parameterized waveforms $\{\phi_i\}_{i=1}^{p}$ called atoms. Many dictionaries have been developed and tested over the last decade, with a particular emphasis on wavelet dictionaries. With such a general definition of a dictionary, many can be created, but the choice of which waveforms to include is critical for ensuring accurate and sparse representations.

A discretized power system signal can be decomposed into m dictionary waveforms as

$$x = \sum_{i=1}^{m} \alpha_i \phi_i + R^{(m)} \tag{1.3}$$

where $R^{(m)}$ is a residual and α_i represents the coefficient on the dictionary waveform ϕ_i. Assuming the dictionary contains a basis for l^2, an exact decomposition can be made with

$$x = \sum_{\phi \in D} \alpha_i \phi_i. \tag{1.4}$$

What the signal x is decomposed into depends on the choice of dictionary. Ideally, the major features are represented in only a few waveforms from D. In the case of power quality monitoring, the transient disturbances should be isolated as the few major features. If this is done properly, then the transients will be detected and identified.

If the dictionary represents a simple basis, the decomposition will be unique. However, of more interest are overcomplete dictionaries. These dictionaries allow for many different decompositions of the signal. Considering the waveforms of the dictionary as discretized with length n, a dictionary of p waveforms can be represented as an n-by-p matrix, Φ. Then, 1.4 becomes

$$x = \Phi\alpha. \tag{1.5}$$

If the dictionary is a simple basis then Φ is a nonsingular n-by-n matrix and the decomposition can be easily found using the inverse. However, when the dictionary is overcomplete the system becomes n-by-p and the system is underdetermined, meaning there is no unique solution. Which of these solution is best and what method should be used to choose it are the remaining questions.

Minimizing the l^2 norm, a measure of energy, is one way to choose the solution. This problem formulation is

$$\min_{\alpha} \|\alpha\|_2 \text{ subject to } x = \Phi\alpha. \tag{1.6}$$

One of the primary motivations for minimizing the l^2 norm is that the problem formulation has a unique solution of the form

$$\alpha = \Phi^T(\Phi\Phi^T)^{-1}x. \tag{1.7}$$

However, while this minimizes the l^2 norm, it does not result in a sparse solution as the energy is dissipated over the coefficients. A sparse solution is such that $\|\alpha\|_0 << p$ where

$$\|\alpha\|_0 = \#(i : \alpha_i \neq 0). \tag{1.8}$$

The problem formulation in this sparse framework is then

$$\min_{\alpha} \|\alpha\|_0 \text{ subject to } x = \Phi\alpha. \tag{1.9}$$

Unfortunately, solving this formulation of the problem is NP-Hard [1]. Over the last decade, much research has gone into solving this problem, particularly by relaxing the condition slightly to

$$\min_{\alpha} \|\alpha\|_1 \text{ subject to } x = \Phi\alpha. \tag{1.10}$$

When the problem is convexified in this way and then solved, it is known as Basis Pursuit. The solution is found through advanced linear programming techniques. The solution of 1.10 will match the solution of 1.9 under certain conditions developed by Donoho and Elad [7]. This thesis will investigate the effectiveness of this convex relaxation problem on the proposed power system signal.

In the presence of noise, 1.10 must be recast to an approximation

$$\min_{\alpha} \|\alpha\|_1 \text{ subject to } \|x - \Phi\alpha\|_2 \leq \delta, \tag{1.11}$$

which is known as Basis Pursuit Denoising.

The solution to this problem is the same as the solution to the unconstrained optimization problem

$$\min_{\alpha} \lambda \left\|\alpha\right\|_1 + \frac{1}{2} \left\|x - \Phi\alpha\right\|_2^2 \tag{1.12}$$

for a suitable λ value. The higher the λ, the harsher the punishment for the solution to go away from sparsity. The λ value will depend on the δ range for which the solution is to stay within.

Basis Pursuit, solving 1.10, and Basis Pursuit Denoising, solving 1.11 or 1.12, are simply different problem formulations. In fact, Basis Pursuit is a version of Basis Pursuit Denoising with $\delta = 0$. In Basis Pursuit an exact representation is found, but in Basis Pursuit denoising an approximation within a specified tolerance is sought. Most algorithms are optimized to solve the more challenging Basis Pursuit Denoising, but also work for the Basis Pursuit case where $\delta = 0$.

1.3 The Dictionary

This thesis uses two dictionaries, a wavelet packet dictionary and a damped sinusoid dictionary. A wavelet packet dictionary has $n \log_2(n)$ waveforms of length N, indexed by (j, p, k) where $1 \leq j \leq \log_2(N)$, $0 \leq p \leq 2^{-j}N - 1$, and $0 \leq k \leq 2^j - 1$. Each of these atoms has a similar time-frequency localization property to a discrete window function, dilated by 2^j, centered at $2^j(p + \frac{1}{2})$, and modulated by a sinusoidal wave of frequency $2\pi 2^{-j}(k + \frac{1}{2})$ [18, 12, 25].

A variety of mother wavelets can be chosen such as Haar, Daubechies, Coiflets, and Symlets. For power quality application the wavelet should be oscillatory with short support. The most widely chosen such wavelet is the Daubechies 4, db4, wavelet [3, 13].

The second dictionary is a damped sinusoid dictionary. This dictionary is constructed similarly to the wavelet packet dictionary, but using a mother wavelet of a damped sinusoid. Because the features searched for are in fact damped sinusoids, this dictionary should do well in identifying these transients, but its lack of orthogonality and high level of coherence make such a dictionary less popular.

This thesis investigates the aforementioned mutual coherence of these dictionaries. The mutual coherence is defined as

$$\mu(\Phi) = \max_{1 \le k, j \le m, k \ne j} \frac{\left|\phi_k^T \phi_j\right|}{\|\phi_k\|_2 \|\phi_j\|_2}. \tag{1.13}$$

With an appropriately low mutual coherence, the solutions to 1.10 and 1.9 will match. The theorem, proved by Donoho and Elad [7], states

Theorem 1.1 *If a solution α exists obeying*

$$\|\alpha\|_0 < \frac{1}{2}(1 + \frac{1}{\mu(\Phi)})$$

then that solution is both the unique solution of 1.10 and 1.9.

Another theorem from Donoho [1] uses the coherence in relation to the matching pursuit algorithm:

Theorem 1.2 *If a solution α exists obeying*

$$\|\alpha\|_0 < \frac{1}{2}(1 + \frac{1}{\mu(\Phi)})$$

then an Orthogonal Greedy Algorithm with threshold parameter $\epsilon_0 = 0$ is guaranteed to find it exactly.

1.4 The Algorithms

This thesis investigates two different algorithms adaptive pursuit algorithms that pursue sparse solutions. They will be evaluated on the two dictionaries and in noisy and noiseless scenarios.

1.4.1 Matching Pursuit

The matching pursuit algorithm proposed by Mallat and Zhang [19], is a greedy algorithm which addresses solving 1.9 directly by piecing the decomposition together step by step. Starting with an initial approximation $x^{(0)} = 0$ and residual $R^{(0)} = x$ a sparse approximation is built by adding a new atom to the decomposition at each iteration. At the k^{th} stage, the dictionary is searched for the atom that has the largest correlation with the current residual. This is the piece that will reduce the residual the most by adding it to the approximation. The approximation $x^{(k)} = x^{(k-1)} + \alpha_k \phi_k$, where $\alpha_k = \left\langle R^{(k-1)}, \phi_k \right\rangle$, the l^2 inner product. The new residual is $R^{(k)} = x - x^{(k)}$, and the algorithm iterates until a stopping condition is met on the residual.

If the dictionary is orthogonal, the Matching Pursuit algorithm works flawlessly. If this is the case and the signal consists of only m objects, then running Matching Pursuit for m iterations will find this solution. However, as is often the case in overcomplete dictionaries, the dictionary is not always orthogonal. Because the algorithm is a greedy algorithm, a poor choice early on can result in the algorithm spending its time trying to correct itself. This potential drawback does occur in practice [1, 6, 5].

1.4.2 Basis Pursuit

Basis Pursuit is concerned with finding a sparse solution through solving 1.10. The term Basis Pursuit does not refer to a single specific algorithm, but instead refers to any algorithm which solves the Basis Pursuit problem formulation. The methods for accomplishing this task initially suggested by the authors of [6] are of the linear programming variety due to strides made in computational efficiency with interior point methods. The convexity of the problem formulations allows linear program solvers to find the global solution to 1.10. In recent years, many algorithms have been proposed to accurately and quickly solve 1.12, and they are reviewed in [1].

This thesis uses a primal-dual log-barrier linear program technique to solve 1.10. In the presence of noise, 1.12 is transformed to a second order cone program and then solved using a primal-dual log-barrier algorithm. These solvers are readily available in the SparseLab package available for Matlab.

Chapter 2

Research Design and Methods

The goal of the research presented is to determine the effectiveness of using sparsity constraints and adaptive pursuit techniques for transient detection in power systems signals. The methods will be considered effective if the transient disturbances are represented correctly and sparsely in the time frequency plane. This implies a crisp detection of the transient.

2.1 Test Signals

In the interest of simplicity and computational tractibility, this thesis focuses on a specific type of transient, capacitor switches. Energizing a capacitor bank results in a quick voltage recovery that overshoots and results in an oscillating transient voltage superimposed to the system fundamental. Following Zhu's description of capacitor switches in [26], this thesis focuses on damped sinusoids with frequency ranging from several hundred to a couple thousand Hertz, damping factors of 450-500, and phase angles from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$.

The synthetic signal is designed to test the algorithms' ability to not only detect these transients, but to differentiate between two different disturbances that may occur in close proximity. The signal is designed to model a power system signal after it has been through a high pass filter that takes out the low harmonics, particularly the 60Hz fundamental, and leaves the transients, namely capacitor switching transients. The signal is sampled at N=512 equally spaced points on the interval $[0, .1]$. This means the sampling frequency is 5120 Hz, and frequencies up to 2560 Hz can be accurately resolved. The first signal, f_1, is without noise

$$\begin{aligned} f_1(t) = e^{-460.5(t-.03)} \sin(2\pi 875(t-.03))u(t-.03) + \\ e^{-475(t-.033)} \sin(2\pi 2000(t-.033))u(t-.033) + \\ e^{-500(t-.08)} \sin(2\pi 1500(t-.08))u(t-.08), \end{aligned}$$

while f_2 is affected by noise

$$f_2(t) = f_1(t) + n(t)$$

where n(t) is random Gaussian white noise, ranging from -.3 to .3 with a mean of 0.

2.2 Dictionary

Two dictionaries are considered in this research: a wavelet packet dictionary based on the Daubechies 4 wavelet and a damped sinusoid dictionary. The wavelet packet dictionary is constructed from the db4 wavelet, with $n\log(n)$ atoms. Since N=512, the dictionary includes all of the wavelet packet atoms up to level $\log(n) = 9$, for a total of 4608 waveforms. The Damped Sinusoid Dictionary is built similarly to

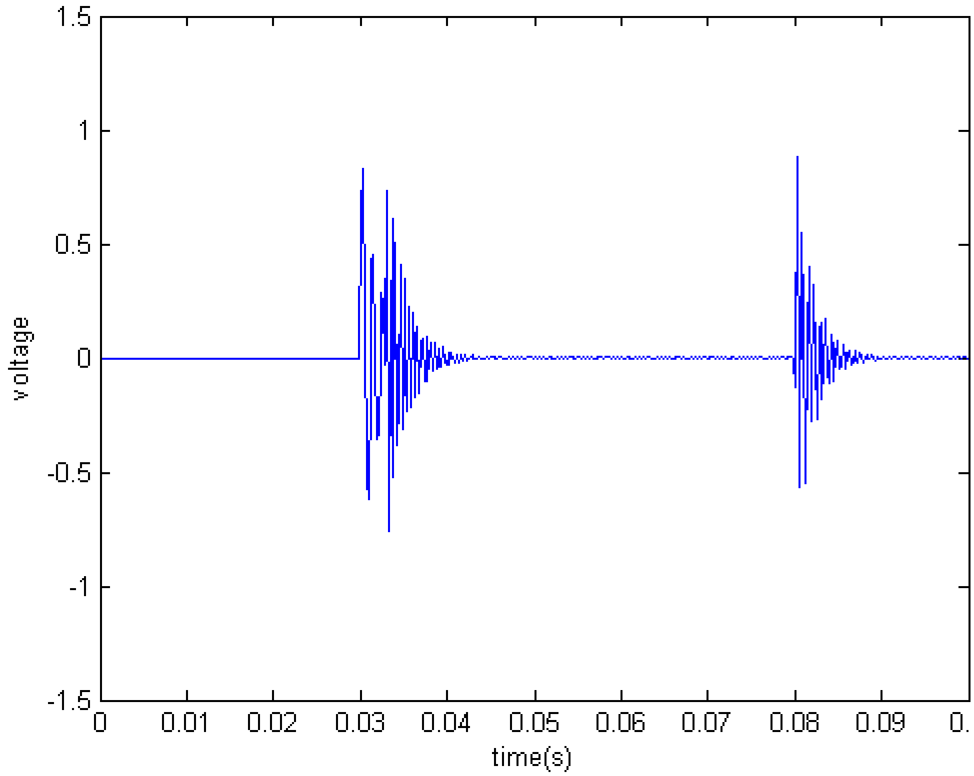

Figure 2.1: The tested model without noise sampled at 512 points from 0 to .1 seconds.

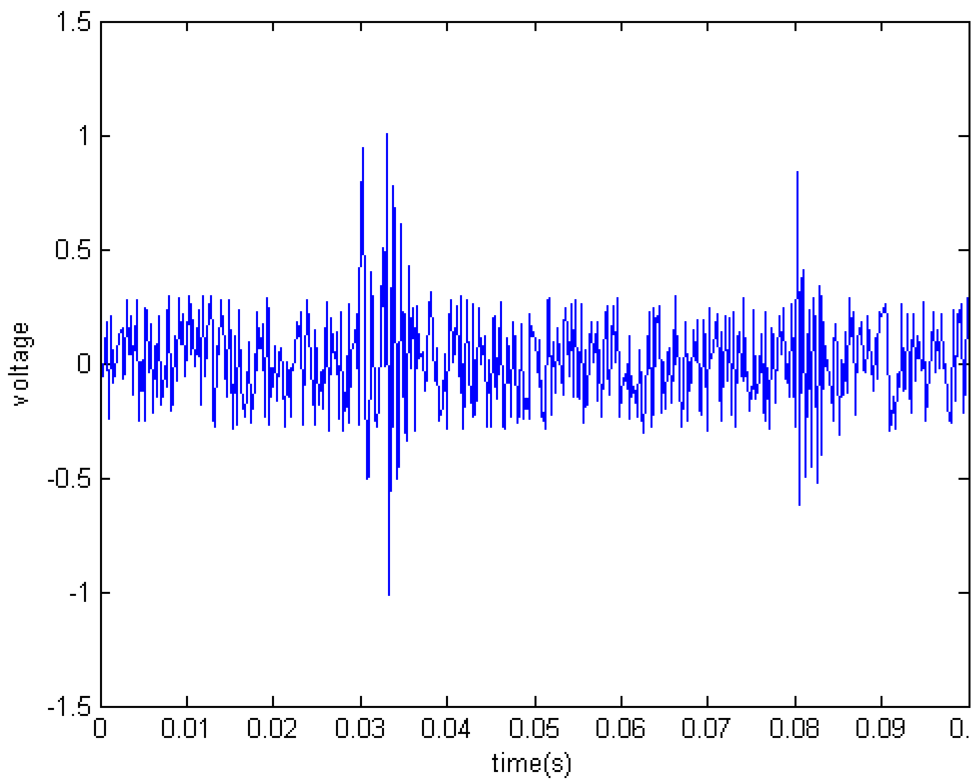

Figure 2.2: The tested model with noise sampled at 512 points from 0 to .1 seconds.

a wavelet packet dictionary by scaling, dilating, and translating

$$\psi(t) = e^{-480t} \sin(2\pi 1250t). \tag{2.1}$$

The parameters of this damped sinusoid are chosen to be average in the range of capacitor switches transients previously given as the focus of this detection scheme. Since the type of transient to be searched for is defined, the idea is that a waveform similar to the transient is the best suited to create a dictionary from.

The respective mutual coherences will be calculated numerically and each will be evaluated on the test signals using both algorithms.

2.3 Motivation for Pursuit Algorithms and Sparsity

In detecting and identifying transients, a sparse solution will associate certain dictionary atoms with certain types of transients. If the solution is sparse in the time frequency display, the transient can be characterized by its frequency resolution, and the time of the disturbance can be identified. The two pursuit algorithms, Matching Pursuit and Basis Pursuit, are used in this thesis instead of the popular wavelet transform [20]. Because the goal of sparsity has been set forth, the wavelet transform will not be sufficient. Using $f_1(t)$, a multilevel wavelet decomposition using db4 wavelets and depth nine gives the coefficient vector shown in Figure 2.3. This wavelet transform method shows a clear ringing and the energy is dispersed over many coefficients. This method will not give as sparse of a solution as the results presented later using Basis Pursuit and Matching Pursuit.

Wavelet shrinkage, a process of denoising by thresholding the detail coefficients

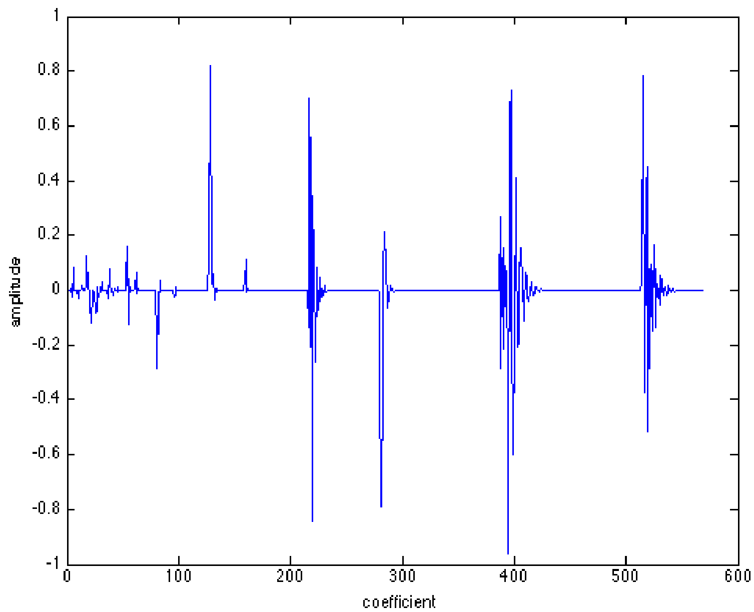

Figure 2.3: The coefficient vector of a wavelet decomposition using db4 wavelets and depth 9.

of the multilevel discrete wavelet transform, can be used to denoise a signal and then analyze it. This concept is not visited in this thesis. Instead the implemented adaptive pursuit algorithms address the sparsity issue directly.

2.4 Comparing Sparsity

The time frequency plane displays the concentration of various dictionary atoms. In this study, the sparsity of a representation will be viewed through this tool as well as a simple plot of the coefficients. In the time frequency plane the boxes representing a particular atom are referred to as Heisenberg boxes, and concentration of the atom is displayed through the darkness of its Heisenberg box. Crisp, sparse representations will have time frequency representations with few dark boxes at the appropriate locations.

2.5 Importance of the Study

Enforcing sparsity constraints has proven to be key in other applications, and this concept has not yet been fully addressed in power system signals. Matching pursuit algorithms for power systems analysis and transient detection have previously been proposed [15, 16, 17], but using basis pursuit has not been properly investigated. With a sparse representation of the transient disturbances, the characteristics of each disturbance have been identified, an important task in power quality monitoring.

Chapter 3

Related Work

3.1 Sparsity Development

This thesis's primary reference for results in sparsity is a sequence of papers from David Donoho and Michael Elad with others [7, 8, 9] along with their comprehensive review article [1]. These articles develop two concepts to use as tools for measuring the ability of pursuit algorithms to find the sparsest solution. First, the spark of a matrix is defined as

$$spark(\Phi) = \text{ the smallest number of columns from } \Phi \text{ that are linearly dependent.} \tag{3.1}$$

Using this concept the following theorem was proven in [7]:

Theorem 3.1 *If a system of linear equations $\Phi\alpha = x$ has a solution α obeying $\|\alpha\|_0 < \frac{spark(\Phi)}{2}$, this solution is necessarily the sparest possible.*

While this was a breakthrough, there is no efficient way to find the spark of matrix other than an exhaustive search. However, it is shown that the spark is related to the mutual coherence defined in 1.13.

Theorem 3.2 *For any matrix $\Phi \in \Re^{n \times m}$, the following relationship holds:*

$$spark(\Phi) \geq 1 + \frac{1}{\mu(\Phi)}$$

Since the mutual coherence can be used as a bound on the spark, the following result is given in [7]:

Theorem 3.3 *If a system of linear equations $\Phi\alpha = x$ has a solution α obeying*

$$\|\alpha\|_0 < \frac{1}{2}(1 + \frac{1}{\mu(\Phi)})$$

then this solution is necessarily the sparsest possible.

Then from here the mutual coherence can be used to get the results of Theorem 1.1 and Theorem 1.2. That is, the ability of the solution of 1.10 to match the solution of 1.9, and the ability of Matching Pursuit to find the sparsest solution. The coherence level theory is extended to deal with stability in finding the sparest solution in noise. The main result from [8] is:

Theorem 3.4 *For a noise bound of ϵ, a bound on $\|x - \Phi\alpha\|_2$ of δ, a dictionary of size $N \times M$, and an arbitrarily small T, if*

$$\|\alpha\|_0 < \frac{1 + \mu(\Phi)}{2\mu(\Phi) + \frac{2\sqrt{N}(\epsilon+\delta)}{T}}$$

then with $\delta \geq \epsilon$ the solution to 1.11 exhibits stability

$$\|\alpha - \alpha_0\|_1 \leq T$$

where α_0 is the sparest solution.

Since T is arbitrary, any can be chosen to balance its effect on the bound of $\|\alpha\|_0$ and its bound on $\|\alpha - \alpha_0\|_1$.

Extending beyond a simple coherence argument, Joel Tropp introduced the cumulative coherence function in [21]:.

$$\mu_1(m) = \max_{|\Lambda|=m} \max_{\psi} \sum_{\lambda=1}^{m} |\langle \psi, \varphi_\lambda \rangle|, \tag{3.2}$$

where Λ is any collection of dictionary waveforms, in this case with cardinality m, ψ ranges over the atoms in the dictionary which are not included in Λ, and the φ_λ are the individual waveforms comprising Λ .

The reasoning behind this measure is to soften the effect of one high correlation in the dictionary, which is always reflected in the mutual coherence. Instead, Tropp noticed that the same results can be reached depending instead on the rate at which the correlation of dictionary waveforms grows.

Theorem 3.5 *If the following condition is met*

$$\mu_1(m-1) + \mu_1(m) < 1,$$

then both orthogonal matching pursuit and basis pursuit will find the correct sparse solution.

This thesis evaluates the mutual coherence of the tested dictionaries before applying them directly to the test signals.

3.2 Algorithm Development

Matching pursuit was developed by Mallat and Zhang in [19]. They use Gabor time frequency atoms as dictionary waveforms and explain this choice because these atoms are optimally localized in time and frequency. A tweaked version of this algorithm is applied to power system signals by Lovisolo and others in [15, 16, 17]. These works use an expanded Matching Pursuit algorithm on a dictionary of damped sinusoids to empirically successful results.

Basis pursuit was introduced in [6]. In this seminal work, the problem is formulated and a primal-dual log-barrier linear programming approach is suggested to solve the problem. Many different test signals are used to show the difference between Basis Pursuit, Matching Pursuit, and what is referred to as the Method of Frames. In the Method of Frames, the l^2 norm of the solution vector is minimized as formulated in 1.6. It is evident that the Method of Frames is not sparsity preserving, as searching for a single time-frequency atom solution instead returns a solution with the energy spread amongst many atoms. Further examples are contrived in which the matching pursuit algorithm preforms poorly because of missteps in its first few iterations, and the value of Basis Pursuit resulting in the global solution to the problem formulation is evident. Basis Pursuit Denoising, 1.11,is also introduced and developed in [6].

3.3 Dictionary Development

3.3.1 Damped Sinusoid

A parameterized damped sinusoid is sometimes referred to as the Laplace wavelet in literature. This wavelet is used in [14, 24, 10]. A version of matching pursuit is used along side a damped sinusoid, or Laplace wavelet, dictionary in [26, 15, 16, 17, 11]. The parameterized waveform is defined explicitly in [26] as

$$g_\gamma(t) = Ae^{-\lambda t} \sin(2\pi f t + \phi) \tag{3.3}$$

for $t \in [0, T]$.

The waveforms are indexed by $\gamma = (f, \lambda, \phi)$ where f is frequency in Hertz, λ is the damping factor, ϕ is the phase angle, A is the initial amplitude, and T is the support range. In [15, 16, 17, 11] the waveform is given time supports to allow for different starting points. The dictionary is formed by adding each combination of waveforms over a mesh of the parameters. In [26] the range used is 851-1220 Hz with 185 elements, 450-500 s^{-1} with 15 elements, and $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ with 45 elements for the phase angle. The different combinations here gives a dictionary with 124,875 waveforms. These parameters were previously mentioned as a guide for the power system capacitor switching transients sought by this thesis. However, the damped sinusoid dictionary used in this thesis is built similarly to wavelet packets by scaling, dilating, and translating a single mother damped sinusoid function instead of including each parameter combination of damped sinusoids as in [26, 15, 16, 17, 11].

3.3.2 Wavelet Packet Dictionary

Wavelet Packets were first introduced by Coifman and Wickerhauser [4]. They were designed to improve computational speed as from a wavelet packet, there are many choices for a wavelet basis. To choose which basis to use an entropy based function is minimized, and the result is known as the best basis. In [18] there is a comprehensive chapter detailing wavelet packets. The key difference between wavelet analysis and wavelet packet analysis is that in a multilevel decomposition wavelet analysis filters the approximation coefficients at each step, while a wavelet packet filters both approximation coefficients and detail coefficients.

Taking all the waveforms from a wavelet packet and adding them to a dictionary, resulting in $n\log(n)$ waveforms, was done in [12, 25, 22]. In [25], a wavelet packet dictionary based upon a Symmlet wavelet is used with Basis Pursuit to detect faults in rolling element bearings. This research model is echoed by the design of this thesis. In [12], a wavelet packet dictionary is again formed, but this time Matching Pursuit is used for ultrasonic inspection. Finally, in [22] a wavelet dictionary is used with matching pursuit for parametric audio coding.

Chapter 4

Results

First, the mutual coherence of the wavelet packet dictionary is calculated at different sizes. Then, the algorithms and dictionaries proposed were tested on the noiseless and noisy signals. The results are presented here as time frequency representations as well as coefficient plots. The representations that show the key components accurately and in a sparse way are preferred, because this implies the algorithm was stable in the noise and the transient features accurately identified.

4.1 Dictionary Coherence

The mutual coherence of the damped sinusoid dictionary is one due to the redundancy and lack of orthogonality in the atoms. None of the theory mentioned as 1.1 or 1.2 can be applied, and the empirical test here will have to suffice as evidence regarding the effectiveness of the dictionary. The mutual coherence of the wavelet packet dictionary is given in the following table. $1 \leq L \leq \log(N)$ represents the maximum level the wavelet packet reaches (giving $N*D$ waveforms in the dictionary), and μ is the mutual coherence. The signal length needs to be at least 512 out of these

$\boldsymbol{N}$	$\boldsymbol{L}$	μ
512	**9**	.7457
512	**8**	.7457
512	**7**	.7148
512	**6**	.7148
512	**5**	.7148
512	**4**	.7148
512	**3**	.7148
512	**2**	.7148
512	**1**	≈ 0
256	**4**	.7148
256	**3**	.7148
256	**2**	.7148
128	**3**	.7148
128	**2**	.7148

Table 4.1: Mutual Coherence Table

options so that it can resolve frequencies as high as are typical with these types of transients. As shown in this table, the coherence drops slightly as fewer waveforms are included in the decomposition, but the coherence is still high. Other than with only a one level split, in which the db4 waveforms are orthogonal, is the coherence low. Even at small signal lengths and levels, the coherence is high. This is because the coherence reflects the extremes of the dictionary. In spite of many orthogonal waveforms in the dictionary, as seen in the one level test, it only takes one redundancy to push the coherence over .7, which renders the theory of Theorem 1.1 or Theorem 1.2 unusable for this wavelet packet dictionary. Again, just as in the damped sinusoid dictionary, the empirical evidence shall serve as the metric for the success in finding sparse solutions. In terms of the theory developed by Donoho and Elad [7], a sharper metric is needed to evaluate the effectiveness of the increasing number of dictionaries used in today's applications. The cumulative coherence function suggested by Tropp in [21] is a step in this direction, but is still not sharp enough to apply in this situation.

4.2 Simulation without Noise

The Matching Pursuit and Basis Pursuit algorithms are run on f_1 using both the wavelet packet and damped sinusoid dictionaries. From the equation for f_1, the time frequency plane should reflect two disturbances early. It should also be able to separate them both by time and by frequency, as the first impulse has a frequency component of 875 Hz and the second impulse has a frequency component of 2000 Hz. Further, the impulse towards the end of the window should be detected and localized reasonably well in time and frequency.

4.2.1 Wavelet Packet Dictionary

4.2.1.1 Matching Pursuit

Figure 4.1 shows the Matching Pursuit algorithm does a good job of finding each of the three impulses, even though the first two are close together. The coefficient plot reveals a sparse representation. Using this dictionary and this algorithm in a noiseless environment appears to work well.

4.2.1.2 Basis Pursuit

Figure 4.2 shows the Basis Pursuit solution. Similar to the Matching Pursuit algorithm in this case it resolved the three different disturbances accurately and with only a few coefficients as shown by the coefficient vector.

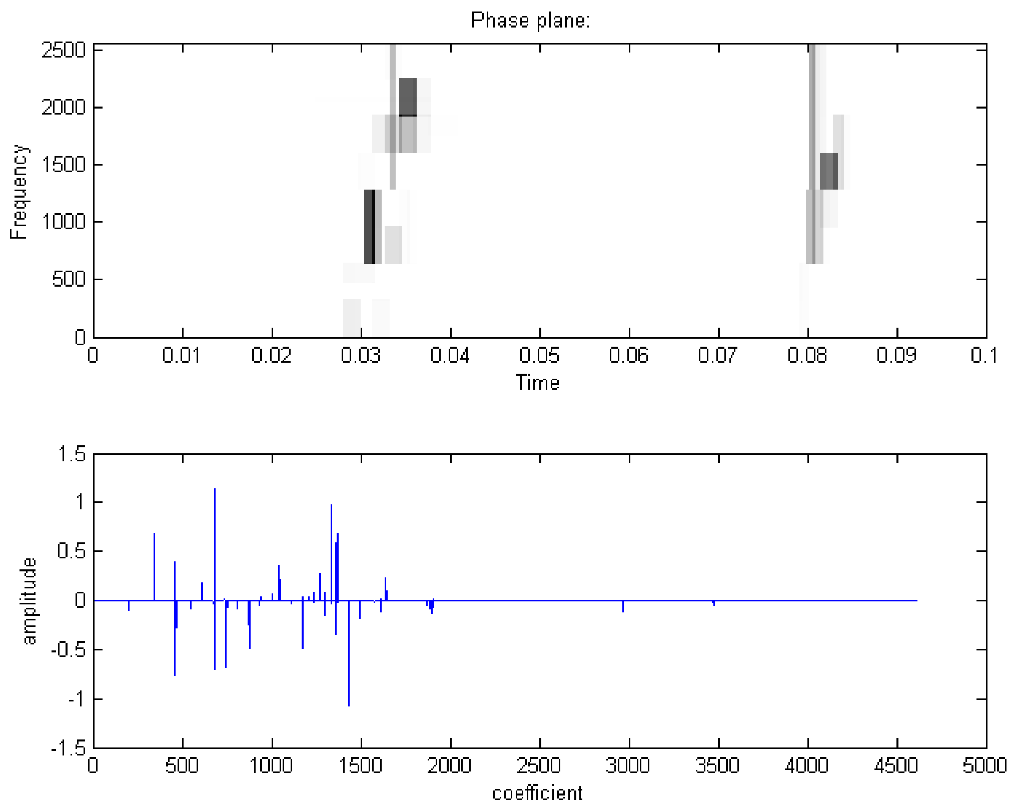

Figure 4.1: Time Frequency representation and coefficient vector after Matching Pursuit on Wavelet Packet Dictionary without noise

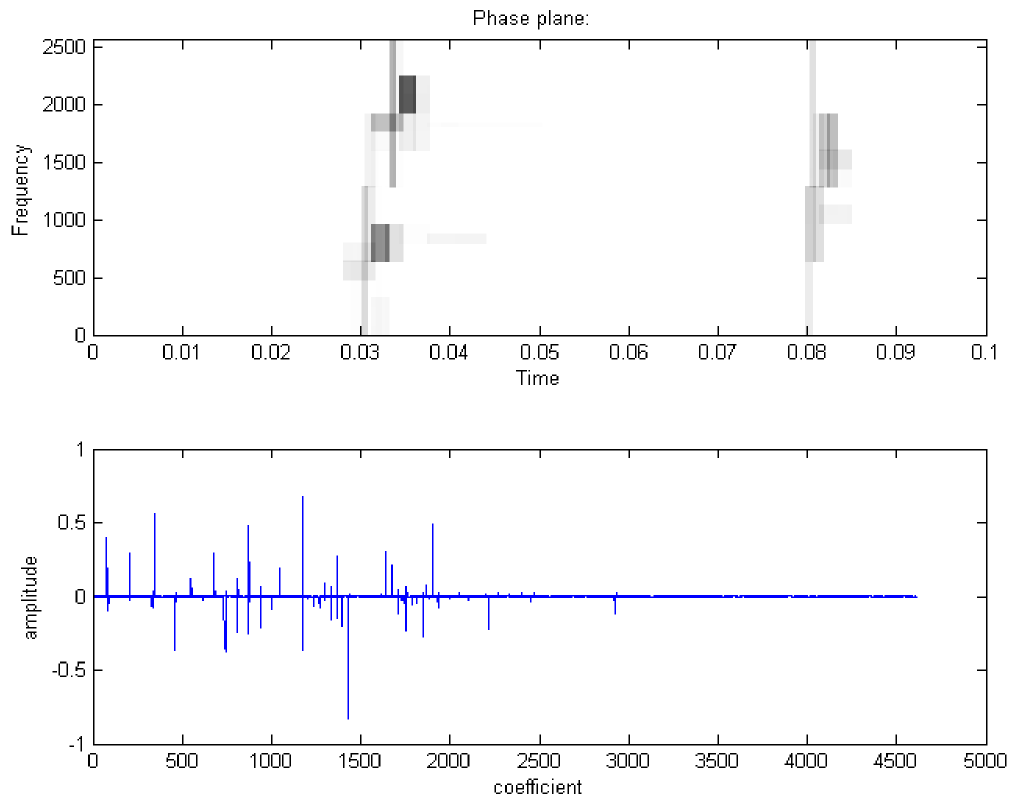

Figure 4.2: Time Frequency representation and coefficient vector Basis Pursuit on Wavelet Packet Dictionary without noise

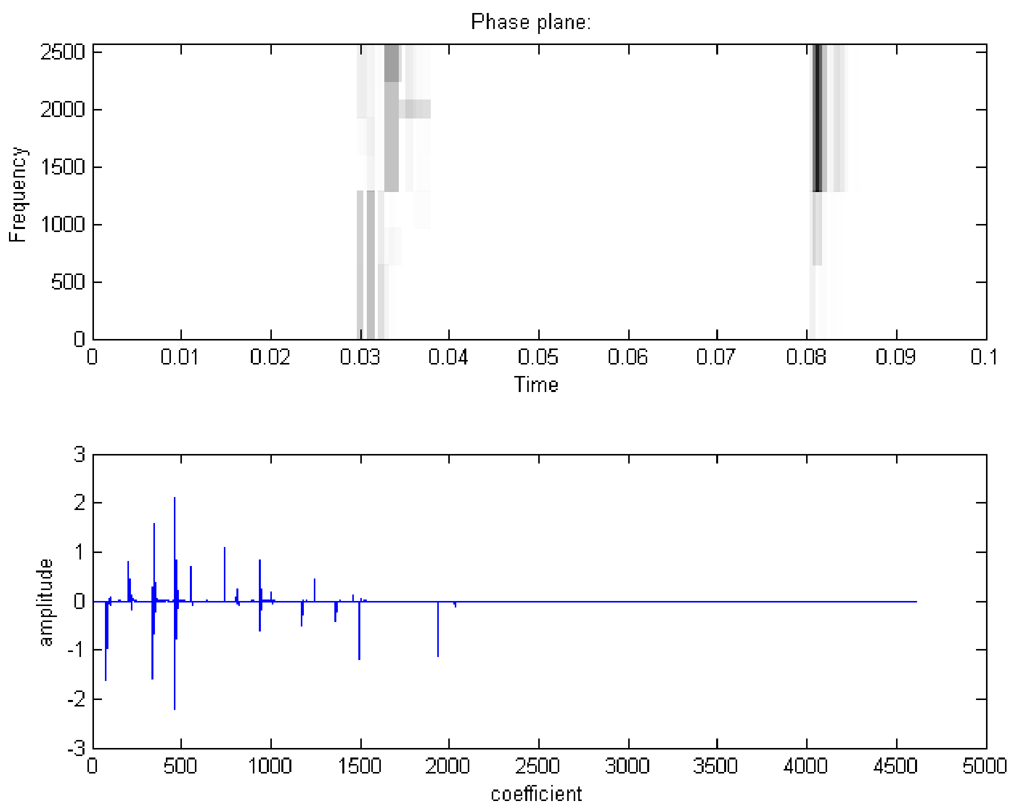

Figure 4.3: Time Frequency representation and coefficient vector Matching Pursuit on Damped Sinusoid Dictionary without noise

4.2.2 Damped Sinusoid Dictionary

4.2.2.1 Matching Pursuit

Figure 4.3 shows that the Matching Pursuit algorithm isolated the three transients in both time and frequency. This time the time-frequency atoms are sharp in time, and less localized in frequency. With the shape of the damped sinusoid waveforms in the dictionary, it is no surprise that a sparse solution is found here, it appears to be even more sparse than the wavelet packet dictionary results.

4.2.2.2 Basis Pursuit

In Figure 4.4, once again without noise the algorithm is able to identify the three impulses. Similarly to Matching Pursuit in the damped sinusoid dictionary, the solution is heavily localized in time and less so in frequency. The overall coefficient vector shows quite a sparse solution.

4.3 Simulation with Noise

The Matching Pursuit and Basis Pursuit algorithms are run on f_2 using both the wavelet packet and damped sinusoid dictionaries. Even with noise added the goal is for these algorithms to identify sparse, accurate, solutions. Because an exact representation no longer makes sense in the presence of noise, a δ value of 1 is set and the solutions will obey $\|x - \Phi\alpha\| \leq \delta$.

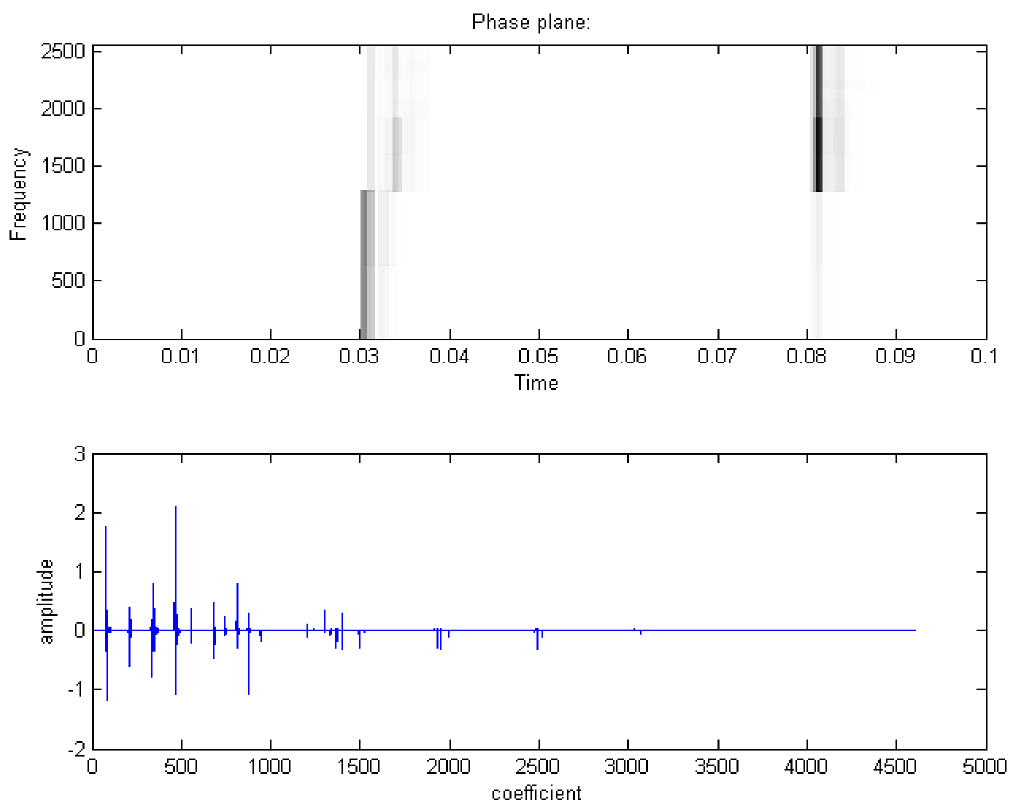

Figure 4.4: Time Frequency representation and coefficient vector Basis Pursuit on Damped Sinusoid Dictionary without noise

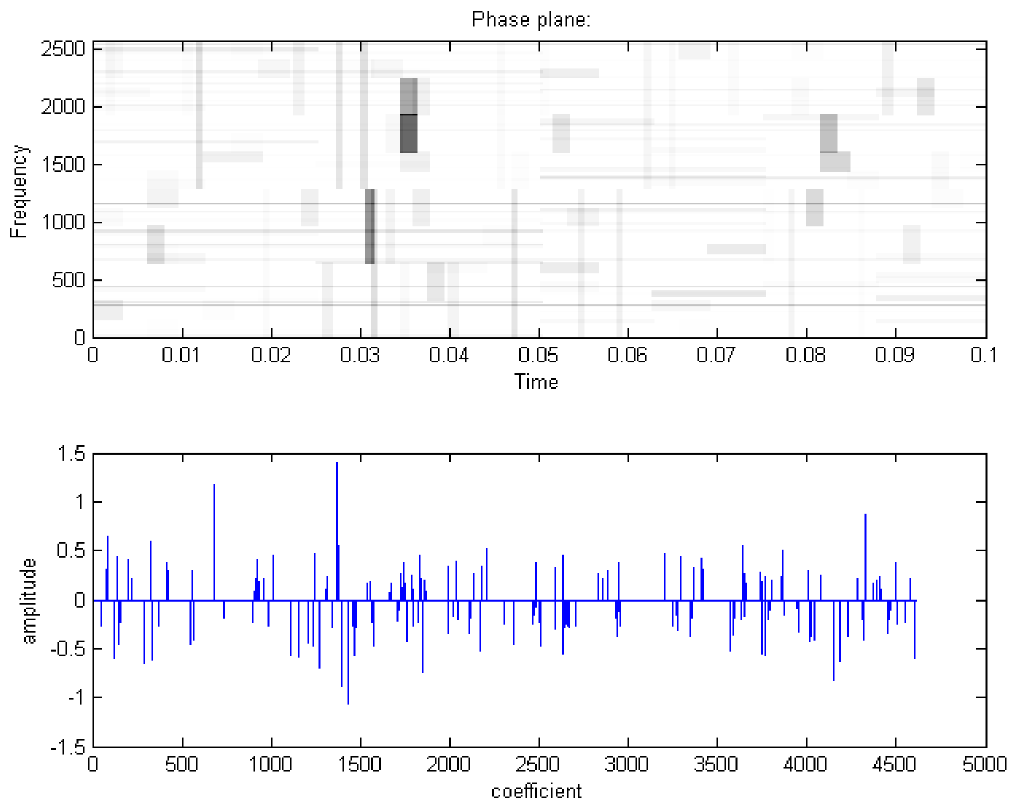

Figure 4.5: Time Frequency representation and coefficient vector Matching Pursuit on Wavelet Packet Dictionary with noise

4.3.1 Wavelet Packet Dictionary

4.3.1.1 Matching Pursuit

Figure 4.5 shows that the Matching Pursuit algorithm picked up a lot of the noise despite only being required to iterate until $\|x - \Phi\alpha\| \leq 1$. The main components of the first two impulses are represented strongly, but the third impulse is barely visible. The coefficient vector suggests a pretty sparse representation of the signal, but the time frequency representation shows that while there may not be that many coefficients there are too many large ones coming from noise. It would be difficult to pick out the impulses from this representation.

4.3.1.2 Basis Pursuit

Figure 4.6 shows that the Basis Pursuit algorithm picked up significantly less of the noise than the Matching Pursuit algorithm even though they were held to the same standard of $\|x - \Phi\alpha\| \leq 1$. However, similar to the Matching Pursuit attempt, the third impulse is weakly represented. The coefficient vector shows a decently sparse solution, albeit with more coefficients than Matching Pursuit.

4.3.2 Damped Sinusoid Dictionary

4.3.2.1 Matching Pursuit

In Figure 4.7, due to its similarity with the mother function that was used to create the dictionary, the third impulse is picked up the heaviest unlike when the wavelet packet dictionary was used. However, other garbage is picked up intensely. Where the wavelet packet dictionary picked up significant amounts of scattered noise, the damped sinusoid dictionary picked up atoms with high amplitude that were not

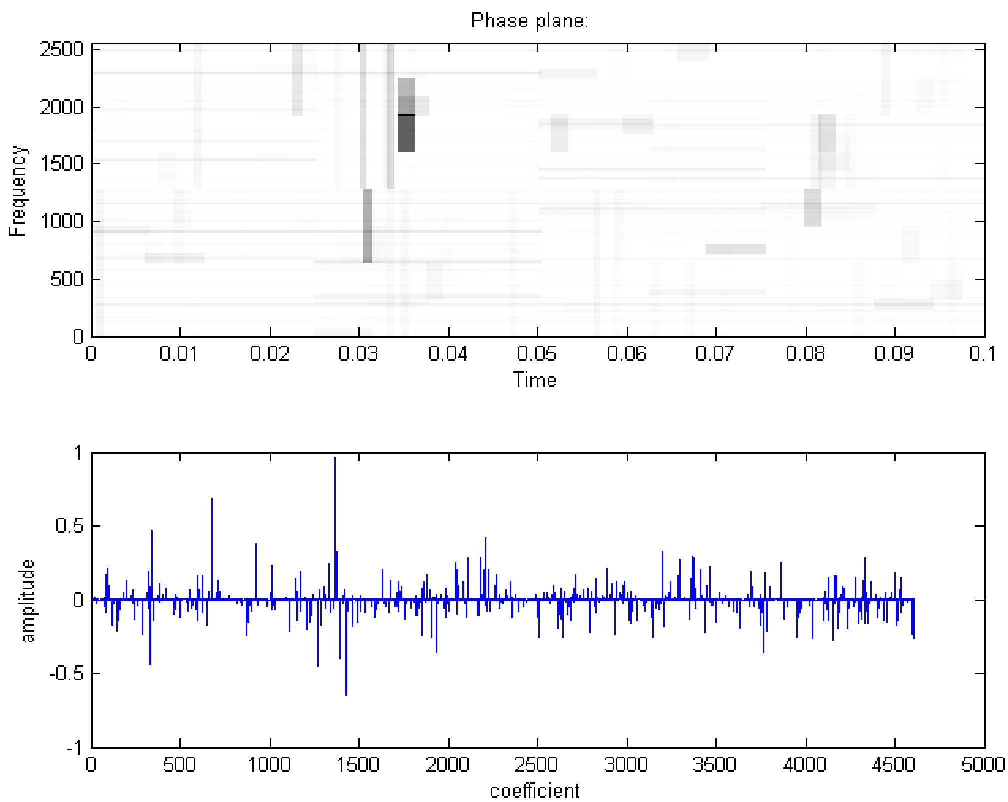

Figure 4.6: Time Frequency representation and coefficient vector Basis Pursuit on Wavelet Packet Dictionary with noise

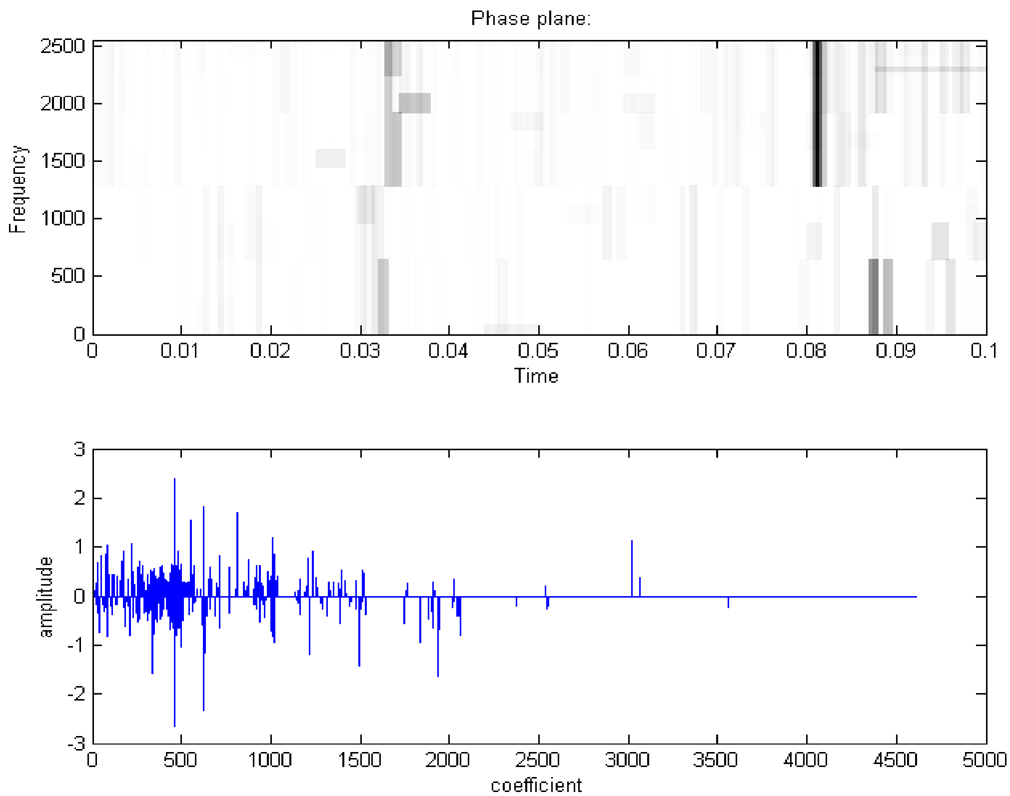

Figure 4.7: Time Frequency representation and coefficient vector Matching Pursuit on Damped Sinusoid Dictionary with noise

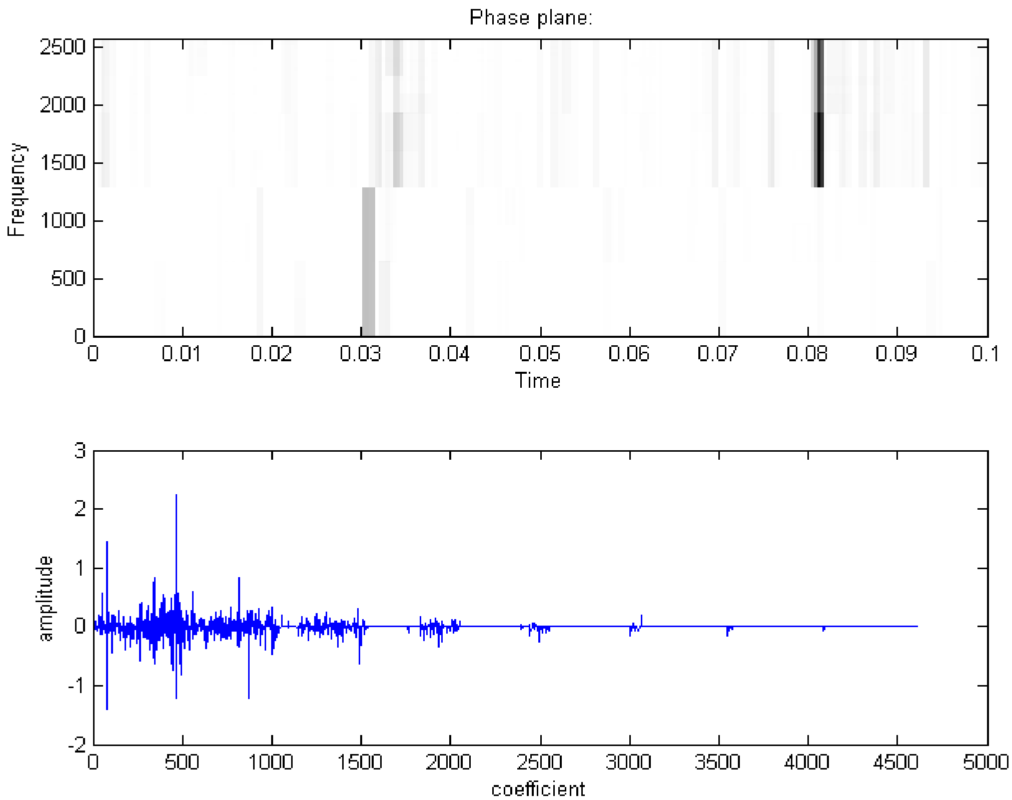

Figure 4.8: Time Frequency representation and coefficient vector Basis Pursuit on Damped Sinusoid Dictionary with noise

impulses. This is not good for a fault detection scheme. At the low levels the solution is not particularly sparse.

4.3.2.2 Basis Pursuit

Figure 4.8 shows less noise than its Matching Pursuit counterpart, and does not pick up any particularly strong false positives. Again, due to similarity to the function used in the creation of the dictionary, the third impulse is picked up the strongest, and the other two are weaker. The time localization of these impulses is stronger than the frequency localization. The coefficient vector is not sparse at low levels.

4.4 Conclusion of Results

The theoretical coherence arguments proved to be irrelevant due to the high coherence of both dictionaries, but the dictionaries do not preform the same. This suggests the need for a better metric. When noise is added, the wavelet packet dictionary resulted in much sparser coefficient vectors than the damped sinusoid dictionary. However, the coherence arguments currently available do not give information to suggest this result, and only the empirical evidence is relied upon. Further comparing the dictionaries, the damped sinusoid dictionary did a superior job in time localization of the impulses, while the wavelet packet dictionary localized in a better compromise between time and frequency.

Without noise present, the Matching Pursuit algorithm and the Basis Pursuit algorithm gave similar time frequency representations, but the Matching Pursuit appears to give slightly better sparsity in the coefficient vector. Once noise is added, the Matching Pursuit algorithm picked up more noise in both dictionary tests, in spite of the fact that the algorithms were held to the same standard of accuracy. Most damaging is the Matching Pursuit algorithm revealed a strong atom where there was no impulse in the damped sinusoid dictionary case.

Chapter 5

Conclusions and Discussion

5.1 Answering the Research Questions

This thesis aimed to compare two dictionaries and two algorithms as they pertain to transient detection in power systems signals, particularly in the presence of noise.

5.1.1 Dictionary

The mutual coherence calculations showed that there was too much redundancy in either dictionary to apply the theory currently developed. However, the structure of the wavelet packet dictionary has less coherence than the damped sinusoid dictionary.

The two dictionaries, wavelet packet and damped sinusoid, preformed with advantages and disadvantages. The conclusion made is that when it is of primary importance to find the exact instants the transients hit the damped sinusoid dictionary does a good job because the dictionary resolved time instants better in both dictionaries, in noise and without noise. However, they did not localize the frequency as well as the wavelet packet dictionary, so if the goal is to compare and identify the transient types then the wavelet packet dictionary is preferable.

In the damped sinusoid dictionary it is hard to tell the difference between the higher frequency transient and the lower frequency transient. This is likely because of the damped nature of these dictionary atoms instead of being pure sinusoids. The damping will lead to broader resonance peaks, meaning the frequencies present will not be sharply, or sparsely, identified. Empirically, this was the case.

Due to the redundancy in the damped sinusoid dictionary it did not resolve very sparse solutions as compared with the wavelet packet dictionary. Even though both dictionaries do not have low enough mutual coherence to apply the theorems referenced, the wavelet packet dictionary certainly has more orthogonality, and as a consequence had more sparse coefficient vectors. In noise, the damped sinusoid dictionary was more likely to find a high correlation with an atom that was not a part of the impulses.

5.1.2 Algorithms

Both algorithms fare well in finding sparse solutions when no noise is present. Either method proved to be acceptable in this case. In the presence of noise, Matching Pursuit picked up a larger spread of the white noise. Neither method found the perfect atoms, but Basis Pursuit included less noise atoms that were unrelated to the simulated impulses. Matching Pursuit used fewer coefficients, so if the goal is for superior compression this algorithm would be better. However, because less noise was picked up the argument can be made that Basis Pursuit solvers should be looked at alongside the plethora of Matching Pursuit algorithms being proposed on the power systems signal analysis community.

5.2 Discussion of future work

A new way to compare the ability of different dictionaries to resolve sparse solutions is needed. The mutual coherence arguments have serious limitations, and the only tool available to compare the two dictionaries in this thesis was simulation and empirical evidence. This is an area where the gap between electrical engineering and mathematics could be bridged if the sparsity theory can be applied to the more generalized dictionaries that work well with power system transients.

Instead of pursuit algorithms, wavelet shrinkage algorithms can be investigated as denoising tools, and further analysis can be preformed on the denoised signal. These types of algorithms have value in larger scale problems. A longer sample of the power system signal would be prohibitively large for the pursuit algorithms due to the size and complexity of the dictionaries.

Common in radar signal processing, the ambiguity function is used with a matched filter to detect signal features in a time-frequency setting, and a similar process could be extrapolated to power system signals. Similarly, the Wigner-Ville transform is another attractive approach for alternate representations in time-frequency for transient identification.

This thesis focused on a specific type of transient disturbance, the ones caused by capacitor switches. Similar analysis should be applied to all forms of power system phenomena, perhaps changing the makeup of the dictionaries to better match the type of transient being searched for. To help with computational complexity, a parallel processing system could be used to simultaneously search hundreds of different dictionaries looking for matches.

The methods applied here were tested on a synthetic signal. This worked well for the purposes of this thesis because a key part of the analysis was knowing what exactly the solution should be. The future of the research should focus on actual samples drawn from power lines.

References

[1] A. Bruckstein, D. Donoho, and M. Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Review*, 51(1), 2009.

[2] M. Chan and R. Donaldson. Amplitude, width, and interarrival distributions for noise impulses on intrabuilding power line communication networks. *IEEE Transactions on Electromagnetic Compatibility*, 21(3), 1989.

[3] S. Chen and H.Y. Zhu. Wavelet transform for processing power quality disturbances. *EURASIP Journal on Advances in Signal Processing*, 2007.

[4] R. Coifman and M. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38(2), 1992.

[5] R.A. DeVore and V. Temlyakov. Some remarks on greedy algorithms. *Advances in Computational Mathematics*, 5(1), 1996.

[6] D. Donoho, S. Chen, and M. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1), 1998.

[7] D. Donoho and M. Elad. Optimally sparse representation in general (nonorthogonal) dictionaries via l1 minimization. *PNAS*, 100(5), 2003.

[8] D. Donoho and M. Elad. On the stability of the basis pursuit in the presence of noise. *Signal Processing*, 86, 2006.

[9] D. Donoho, M. Elad, and V. Temlyakov. Stable recovery of sparse overcomplete representations in the presence of noise. *IEEE Transactions on Information Theory*, 52(1), 2006.

[10] L. Freudinger, R. Lind, and M. Brenner. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12), 1993.

[11] M. Goodwin. Matching pursuit with damped sinusoids. In *In Proc. ICASSP*, pages 2037–2040, 1997.

[12] Y. Guang, Z. Qi, and Q. Pei-Wen. Matching pursuit based adaptive wavelet packet atomic decomposition applied in ultrasonic inspection. *Russian Journal of Nondestructive Testing*, 43(1), 2007.

[13] C.H. Kim and R. Aggarwal. Wavelet transforms in power systems. *Power Engineering Journal*, April, 2000.

[14] H. Li and S. Wang. Matching pursuits with time-frequency dictionaries. *Proceedigns of ISCIT*, 2005.

[15] L. Lovisolo, E. daSilva, M. Rodrigues, and P. Diniz. Coherent decompositions of power systems signals using damped sinusoids with applications to denoising. *IEEE International Symposium on Circuits and Systems*, 5, 2002.

[16] L. Lovisolo, E. daSilva, M. Rodrigues, and P. Diniz. Efficient coherent adaptive representations of monitored electric signals in power systems using damped sinusoids. *IEEE Transactions on Signal Processing*, 53(10), 2005.

[17] L. Lovisolo, M. Tcheou, E. daSilva, M. Rodrigues, and P. Diniz. Modeling of electric disturbance signals using damped sinusoids via atomic decompositions and its applications. *EURASIP Journal on Advances in Signal Processing*, 2007.

[18] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1999.

[19] S. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12), 1993.

[20] D. Robertson, O. Camps, J. Mayer, and W. Gish. Wavelets and electromagnetic power system transients. *IEEE Transactions on Power Delivery*, 11(2), 1996.

[21] J. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Transactions on Information Theory*, 50(10), 2004.

[22] P. Vera-Candeas, N. Ruiz-Reyes, M. Rosa-Zurera, D. Martinez-Munoz, and F. Lopez-Ferreras. Transient modeling by matching pursuits with a wavelet dictionary for parametric audio coding. *IEEE Signal Processing Letters*, 11(3), 2004.

[23] W. Xu. Component modeling issues for power quality assessment. *IEEE Power Engineering Review*, 21(11), 2001.

[24] X. Yan, H. Chengjun, Y. Weiyong, and J. Xiuchen. Research on matching pursuit algorithm based on laplace wavelet for partial discharge pulse detection. *Power Engineering Society General Meeting*, 2005.

[25] H. Yang, J. Mathew, and L. Ma. Fault diagnosis of rolling element bearings using basis pursuit. *Mechanical Systems and Signal Processing*, 19, 2005.

[26] T.X. Zhu. Detection and characterization of oscillatory transients using matching pursuits with a damped sinusoidal dictionary. *IEEE Transactions on Power Delivery*, 22(2), 2007.

[27] M. Zimmermann and K. Dosert. Analysis and modeling of impulsive noise in broad-band powerline communication. *IEEE Transactions on Electromagnetic Compatibility*, 44(1), 2002.