Singular Value Decomposition Applied to Damage Diagnosis for Ultrasonic Guided Wave Structural Health Monitoring

Submitted in partial fulfillment of the requirements for

the degree of

Doctor of Philosophy

in

Department of Civil and Environmental Engineering

Chang Liu

B.S., Civil Engineering, Tongji University
M.S., Civil and Environmental Engineering, Carnegie Mellon University

Carnegie Mellon University
Pittsburgh, PA

August, 2014

In dedication

To my late grandfather
My first and best math teacher

&

To my beloved parents Deming & Xiaomei
For giving me the dream and the best education

&

To my dear wife Xuan
For your support and understanding during these years

Acknowledgements

I am deeply grateful to my advisor, Professor Irving J. Oppenheim, for your guidance and support in my four years at Carnegie Mellon University. You always inspire me with new ideas, provide me great suggestions, and offer me the maximum freedom to explore interesting topics. And thank you so much for your understanding and continuous support when I moved to California to join my wife. The thesis would never have been possible without your guidance and encouragement.

I would also sincerely appreciate the members of my thesis committee, Professors David W. Greve, Mario Bergés, Haeyoung Noh, and Dr. Warren R. Junker. Thank you all for sharing your knowledge with me, for providing valuable and constructive feedbacks to my research and thesis, and for asking insightful questions to help me improve my work.

I would like to thank colleagues and collaborators with whom I had the pleasure to work alongside in the past three years. I would like to thank Professors James H. Garrett, Lucio Soilbelman, José Moura, and Mark Patton for teaching me countless things during the group meetings. And to my fellow graduate student collaborators, Joel B. Harley, Yujie Ying, Nicholas O'Donoughue, Peng Gong, Jun Shi, Yanqing Bao, Matineh Eybpoosh: it has been an honor to work with such intelligent and motivated individuals. To Professors Semiha Ergan, Pine Liu, and Dr. Nestor Gomez, whom I enjoyed being teaching assistants for: I am grateful to learned from you your dedication to teaching and your passion to share your knowledge with students.

A large part of my thesis is experimental work, which would never be done without the support and guidance from Professor Larry Cartwright, Ronald Ripper, Martin H. Altschul and Don Coffelt. Your help makes it so much fun to be doing experiments, and makes me more productive on my research.

I am grateful to the financial support by Westinghouse Electric Corporation and CIT Dean's fellowship. I am very thankful to the collaboration with John P. Lareau, Mike Burke, and Tom Nenno from Westinghouse.

Finally, I would like to thank the wonderful CEE staff and everybody in the CEE community. The four years in this big family is and will remain a very precious experience for me.

Abstract

A structural health monitoring (SHM) system takes frequent monitoring records from permanently installed transducers on structures, and uses the information to identify potential structural degradation and to proactively maintain the structures. Guided wave testing is an attractive technique for structural health monitoring of large structures, because guided waves can propagate long distance and are sensitive to subtle and hidden damage. In guided wave SHM systems, ultrasonic records are often affected by environmental and operational variations, which produce undesired changes and cause false results. Moreover, although continuous monitoring produces sufficient information regarding structural integrity, we lack a data processing tool to extract, store, and utilize the damage-sensitive information to leverage the accuracy and robustness of damage detection and localization.

In this dissertation, we develop a data-driven framework based on singular value decomposition that processes guided wave monitoring records to separate damage-related information from effects of environmental and operational variations. The method decomposes sequential monitoring records to reveal trends of different variations, and identifies the singular vector associated with damage development. Combined with delay-and-sum localization method, we can robustly localize the damage using the right singular vectors, which resemble the scatter signal and are robust to environmental and operational variations. The SVD framework is then refined, by adaptively updating the singular vectors with each arriving ultrasonic record, to achieve online damage detection and localization. The SVD damage diagnosis methodology is applied experimentally to detect and localize damage in plate and pipe structures, both in laboratory tests and in field tests.

Contents

List Of Tables

List Of Figures

Chapter 1 Introduction

1.1 Motivation

Large-scale civil infrastructure systems need to be carefully maintained to ensure their structural integrity and service capability. A structural health monitoring (SHM) systems using ultrasonic guided waves is an attractive technique to diagnose damage for such structures. Traditionally, guided wave records are collected periodically from transducers on the structure and are processed to compensate for undesired variations and compared with intact baseline records to detect and localize damage using different methods. Although these approaches work well in theory and laboratory experiments, they are found less effective in real implementation of such a SHM systems, where guided waves are often affected by variations in environmental and operational conditions (EOC).

This dissertation is concerned with developing a data-processing methodology for practical guided wave structural health monitoring to detect and localize damage. The major challenges to be addressed are robustness in the presence of environmental and operational

variations, and the effectiveness of processing a large volume of monitoring records. We develop a data processing framework based on singular value decomposition to process long-term ultrasonic monitoring records. The framework extracts damage-related information from the records that are robust to benign variations, thereby improving the robustness of damage diagnosis. We evaluated our methods with ultrasonic guided wave records in structural health monitoring applications, in both laboratory and field implementations, and show significant improvement in accuracy and robustness in detecting and localizing small damage on plate and pipe structures. In the following sections, we review the basics of structural health monitoring and guided wave ultrasonics, and then discuss the significant challenges, which we address in this dessertation, for designing and ultilizing guided wave structural health monitoring systems.

1.2 Background

1.2.1 Condition-based maintenance (CBM) and structural health monitoring (SHM)

Conventional infrastructure maintenance usually follows a schedule-based strategy, i.e., the infrastructure systems are examined at mandatory inspection intervals defined by regulations. For example, 'The Pipeline Safety Improvement Act' (Zolet and Moore 2003) requires a 7-year inspection interval, meaning that every pipeline section is reexamined within a 7-year period. However, it is both labor-intensive and time-consuming to inspect the many sections of pipelines manually and/or visually, likewise for other civil and mechanical structures, such as bridges, railroads, wind turbines, and aircrafts. To improve the efficiency and accuracy of inspection, many advanced nondestructive evaluation techniques (NDE/NDT)

have been developed and adopted. However, these techniques are still schedule-based; and the inspection intervals are often long due to the limitations of labor and budget.

Because of these limitations, condition-based maintenance (CBM) has gained increasing attention and interest. Following the CBM policy, the system condition is continuously assessed based on data collected from the structure through continuous monitoring or frequent inspection. Having sufficient information enables us to proactively maintenance the structure before destructive failure, which has proved to minimize the cost of maintenance, improve operational safety and reduce the quantity and severity of in-service system failure (Rao 1996; Davies 1998; Grall et al. 2002; Jardine et al. 2006).

Structural health monitoring (SHM) systems are the key component for implementing the condition-based maintenance policy. SHM collects data from permanently installed transducers on the structure, detects damages using the monitored data, and provides the structure status in real time to help estimate its remaining usable life (Raghavan and Cesnik 2007). Such a system provides more available information that can help shift maintenance strategies from schedule-based to condition-based, thereby reducing maintenance costs and labor requirements. More significantly, it could increase the reliability of structures, because unpredictable structural degradation and damage would be identified and addressed in a timely fashion.

Among many approaches, guided wave testing has emerged as a prominent SHM method, because it could propagate and inspect over long distances with small attenuation. It also provides an effective method to estimate the location, severity and type of the damage, and is a well-established practice in the NDE/NDT industry.

1.2.2 Ultrasonic guided waves

Ultrasonic guided waves, such as Rayleigh surface waves and Lamb waves in thin plates, propagate within solid media with free boundaries (Auld 1973; Graff 1975; Rose 2004). Guided waves can propagate long distances in structures, making it possible to detect flaws over a considerable area with a single transducer (Nienwenhuis et al. 2005). Some guided wave modes also have the benefit of propagating through the entire thickness of the material to detect damage such as material loss or cracks (Alleyne and Cawley 1992; Li and Rose 2001; Demma et al. 2004).

During the past decade, ultrasonic guided waves have been used to monitor and evaluate pipelines in different applications, ranging from transmission pipes to chemical or power generation pipes (Alleyne et al. 2001; Cawley et al. 2003; McGregor et al. 2004; Na and Kundu 2002). The favorable characteristics of these waves, such as long propagation distance and full "illumination" of the pipe cross-section, and high sensitivity to subtle material defects, make guided wave systems promising technologies for continuous and automated damage diagnosis of pipes (Cawley et al. 2003).

However, it is difficult to directly interpret ultrasonic guided waves due to their multi-modal and dispersive characteristics. Dispersion curves relate the velocity of every wave mode as a function of frequency, and thereby are often used to characterize guided waves. Figure 1-1(b) (Cawley et al. 2003) shows the dispersion curves of a steel plate, displaying group velocities against the frequency-thickness product. There are three wave modes, namely, symmetric (S) mode, asymmetric (A) mode, and shear horizontal (SH) mode, corresponding to different particle motions. Except for the SH_0 mode, the group velocity of each guided wave mode varies with the frequency-thickness product, known as dispersion.

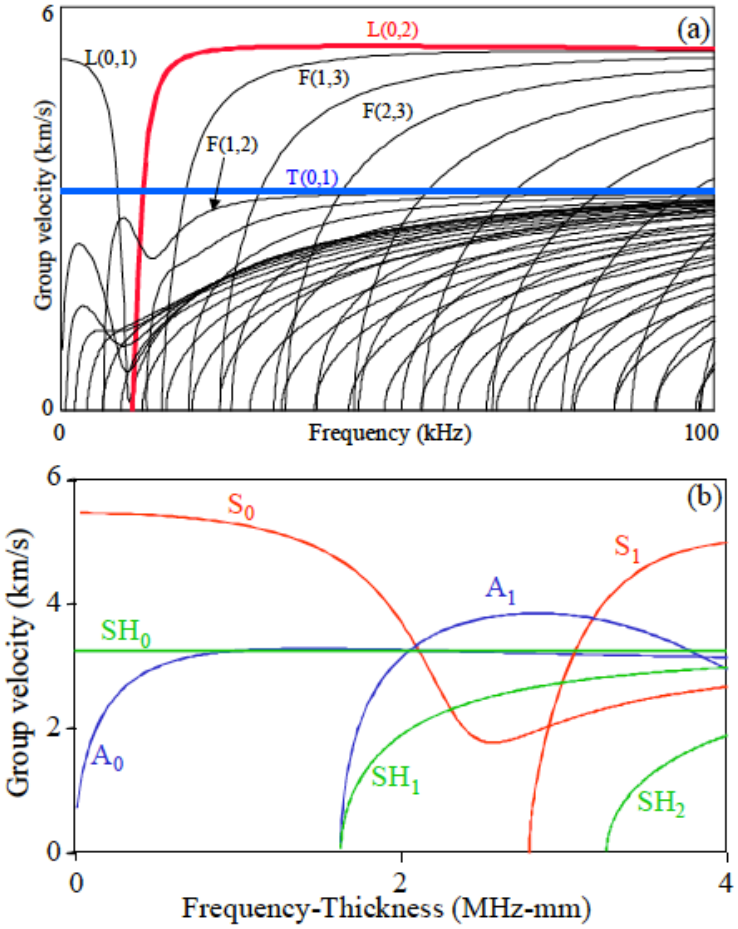


Figure 1-1 Dispersion Curve of (a) 6 inch, schedule 40 steel pipe, and (b) a steel plate [Cawley, 2003]

Similarly, guided waves also propagate in pipes (cylindrical shell), as treated in many textbooks (Auld 1973; Graff 1975; Rose 2004). Because of the closed (circular) cross-section, guided waves in pipes can propagate through multiple paths, increasing the number of wave modes. Corresponding to the Lamb wave modes (in thin plate), pipe guided waves also divide into three categories: longitudinal modes, flexural modes, and torsional modes. Figure 1-2 shows a 3-D finite element simulation of each wave mode (Edwards and Gan 2007). Figure 1-1(a) shows the dispersion curves of a 6-inch, Schedule 40 steel pipe. The dispersion

curve shows more than 50 modes at the frequency range below 100kHZ, many of which are highly dispersive.

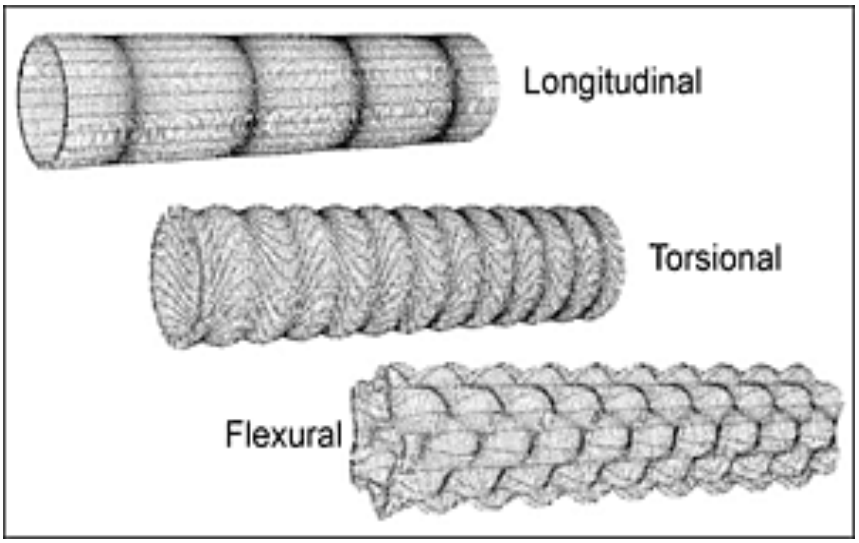


Figure 1-2 Three Guided Wave modes that can propagate in a pipe: longitudinal mode, torsional mode, and flexural mode

As a consequence of these characteristics, guided waves are very complicated. Figure 1-3 shows a typical pitch-catch record propagated in a laboratory pipe specimen. A short (100us) simple excitation becomes a long and complicated signal with many arrivals after propagating a short distance in the pipe. This complexity makes it difficult to identify changes associate with structural damage.

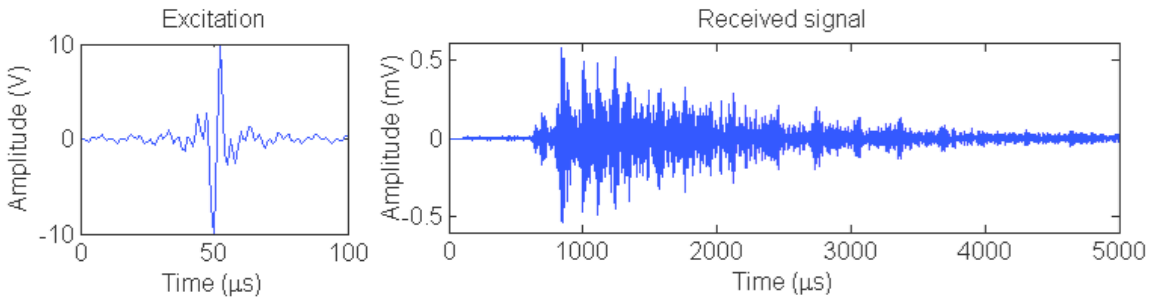


Figure 1-3 typical pitch-catch records collected from a pair of transducers on a pipe in laboratory

A partial solution to this complexity is to design transducers that maximally excite a single mode. Researchers have employed various strategies in the past in order to achieve this goal. Rose (Rose 2000) demonstrated that an angled wedge can be used to convert a longitudinal mode into a particular plate mode by appropriate choice of the incident angle. Rose and his group (Li and Rose 2001) developed finely tuned phased arrays that contain large number of sensors (on the order of 64 or 128) to selectively transmit a single wave mode in a pipe and limit the frequency to a narrow band in which the mode is less dispersive. Figure 1-4 shows a wedge transducer on the left and a (circumferential) comb transducer on the right. The phase velocity of the ultrasonic waves that refract from the wedge to the structure can be calculated following Snell's law. Knowing the wave velocity, a particular mode can be excited by choosing the appropriate frequency. If the array is used, mode selection can be further enhanced by carefully tuning the phase of each individual element (Cawley et al. 2003; Clarke et al. 2009).

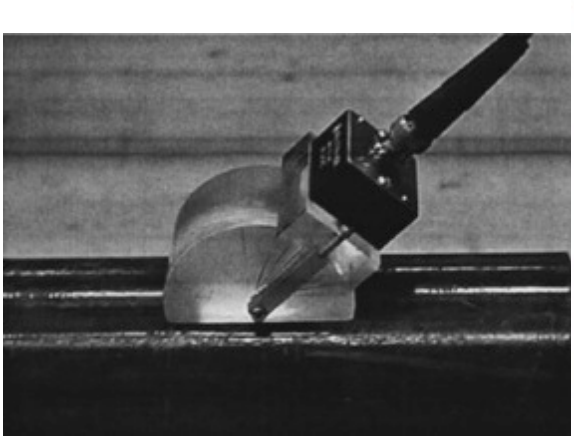


Figure 1-4 Wedged transducer and phased-array

However, the single mode selected for excitation may not be sensitive to all types of damage. Moreover, in practical applications, the EOC effects will shift the dispersion curve, and mode selectivity will deteriorate.

Ultrasonic guided waves can also be effectively excited and picked up by attaching small PZT wafers to a subject structure (Nieuwenhuis et al. 2005). Such a transducer is attractive due to its simplicity and potential low cost. Therefore, in our practice, we take pitch-catch tests using PZT wafers mounted in coarse arrays located tens of pipe diameters apart. We excite a transmitter (one PZT wafer) with a wideband, windowed-sinusoidal signal and listen at multiple receivers in an array for 10 ms, including not only the first arrival but also the later coda waves. We then develop signal processing and machine learning methods to process the long received records and to identify changes in the wave signals. However, there are practical challenges associated with analyzing guided waves for structural health monitoring system, especially when environmental and operational conditions complicate the long-term monitoring records.

1.3 Challenges

In this section, we discuss two major challenges associated with guided wave structural health monitoring, namely, the effect of environmental and operational conditions on guided wave based SHM, and the lack of effective data-processing framework for large volume of long-term monitoring data. Some compensation and data processing techniques are reviewed and their limitations are discussed.

1.3.1 Environmental and operational conditions (EOC)

Despite the increasing published literature on SHM technologies to detect damages in fully operational structures, a large number of these recently proposed SHM methods have been developed under controlled laboratory or simulated conditions, ignoring the effects of environmental and operational conditions on the structures (Sohn 2007; Oh and Sohn 2009). Most existing methods assume that the changes caused by damages will be clearly differentiable from changes caused by EOC variations. However, as proved by several studies, the effects of EOC variations can mask, or adversely affect (through false negatives or false positives) the detection of changes caused by structural anomalies and damages (Croxford et al. 2007; Farrar and Worden 2007; Harley et al. 2012; Sohn 2007). Real-world infrastructure systems, like pipelines, are typically exposed to highly varying EOC, such as flow pressure, flow temperature, flow rate, and external loads. Thus, damage-sensitive features (DSF) identified in a controlled laboratory environment would fail to detect damages of such real-world infrastructures in a robust way.

Generally, detecting and locating damage requires one to isolate ultrasonic scatter signals that originate from damage. This is usually accomplished by baseline subtraction. To

implement baseline subtraction, one measures a single baseline (background) signal when the structure is intact, and then subtracts that baseline from each subsequent observation. One then analyses the resulting (baseline-subtracted) difference signal to determine the presence of damage in a structure. One can further utilize the amplitude and phase information in the difference signal to infer the location of the damage.

Baseline subtraction is unreliable in practical structural health monitoring systems because most structures are located in highly dynamic environments that experience environmental and operational variations over time. These benign variations, both temporal and spatial, may affect the mechanical properties of the structure and thereby distort ultrasonic records, effectively changing the baseline information. As a result, baseline subtraction cannot perfectly remove background signals, and detection and localization frameworks often display false positives results (Sohn 2007; Raghavan and Cesnik 2008; Liu et al. 2012).

One of the most ubiquitous EOC is temperature (Konstantinidis et al. 2007). Variations in temperature change the material density and elastic modulus of the material, thereby altering the velocity of the guided waves (Salama and Ling 1980; di Scalea and Salamone 2008; Raghavan and Cesnik 2008; Clarke et al. 2009). In this case, the conventional baseline-subtraction based methods will likely produce false alarms, because the residual signal after baseline subtraction is dramatically affected by the change in velocity. Researchers have developed various algorithms to compensate for temperature variation. These include methods such as optimal baseline selection (Croxford et al. 2010), local peak coherence (Lu and Michaels 2009; Michaels and Michaels 2005), optimal signal stretch (Croxford et al.

2010; Konstantinidis et al. 2007; Lu and Michaels 2009), and scale transform based temperature compensation (Harley and Moura 2012).

In optimal baseline selection, a library of baseline signals is collected at many different temperatures. To perform baseline subtraction, one first finds the baseline signal that optimally matches the monitoring records and then subtracts that baseline signal to best recover the scatter signal. The method can be easily implemented, but requires a large number of baselines at many different temperatures. Moreover, it is often not practical to collect a complete baseline set for the following reasons (Putkis and Croxford 2013a): (1) There is usually no or little control over the EOC variations that a structure is exposed to, such that a complete baseline set has to be collected while the structure is experiencing real variations. Therefore, a long period of time must be dedicated before the monitoring phase of SHM can be started; (2) The assumption that the structure is undamaged during the baseline acquisition process becomes weaker as the process lengthens; and (3) If the range of EOC variation is unknown prior to the baseline acquisition, it is impossible to determine if the baseline set is complete.

Another group of methods, including optimal baseline stretch and scale transform based temperature compensation, assume that a change in temperature approximately stretches or compresses of measured record. These stretching-based methods estimate an optimal stretching factor and stretch the measurements or the baseline to compensate for the change in temperature. The stretching model of temperature change is approximately true when the temperature variation is small and relatively uniform in space (Salama and Ling 1980; Fukuhara and Yamauchi 1993). However, in many practical scenarios, temperature variations can be large and can vary significantly across a large structure. The spatial temperature

gradients on large structure may negatively affect the accuracy of damage localization, as we will further discuss in *Chapter 5*.

1.3.2 Data-driven damage diagnosis for ultrasonic monitoring records

The field of SHM has been actively benefiting from advanced sensing technologies and data-driven analysis techniques, such as signal processing and machine learning methods. The goal of such data-driven methods is usually to extract damage-sensitive information from data and to detect damage. The frequent or continuous data collection of SHM systems produces a large volume of data that is both a challenge and a benefit. It challenges the conventional data-driven methods for their capability to continuously process the monitoring records, and their robustness to the environmental and oeprational variations contained in the monitoring records. The huge amount of monitoring records also benefits data-driven techniques because SHM produces valuable more information regarding the status of the structure, which, if used properly, can potentially make damage diagnosis more accurate and robust.

Depending on different measurement techniques and the applications, the detection problem can be either a classification problem or a regression problem. Classification, in the statistical learning sense, is to associate a set of measured quantities with a set of labels or categories (discrete outputs), while regression is the task of mapping a set of measured quantities to a set of continuous outputs (Cherkassky and Muller 1998). In learning theory, classification and regression are generally considered as supervised learning algorithms, in which a set of correctly labeled observations are available. In contrast, unsupervised learning tries to find hidden structures from the observations without knowledge of these labels. Examples of unsupervised learning algorithms are clustering and dimensionality reduction.

Many researchers have investigated Machine Learning and Signal Processing algorithms in the context of damage detection in civil structures. Signal processing algorithms and some machine learning algorithms are used to generate the features that summarize the data in a lower dimension or represent them in a latent space. Then machine learning classifiers are used to learn a boundary to separate and label the observations into damaged and undamaged sets (Doebling et al. 1998).

According to a 1998 review from Los Alamos National Laboratories (Doebling et al. 1998), the majority of data-driven damage detection algorithms up until the mid-1990s were based on neural networks. A later 2004 survey conducted by Sohn et al. (Sohn et al. 2004) summarized other machine learning algorithms researchers had studied to perform damage detection. Examples of machine learning algorithms can be found in work done by Sohn et al. (Sohn et al. 2000) using Auto-Regression (AR) on column structure vibration tests, by Posenato et al. (Posenato et al. 2008) applying Moving-window Principle Component Analysis (MPCA) to finite element simulation of bridge structures, and by Vanlanduit et al. (Vanlanduit et al. 2005) applying Singular Value Decomposition (SVD) on vibration tests of a laboratory aluminum beam. Examples of supervised machine learning for structural damage detection include work reported in Figueiredo et al. (Figueiredo et al. 2011) on neural networks and work presented by Ying et al. (Ying et al. 2011; Ying 2012) using a Support Vector Machine (SVM).

Worden and Manson (Worden and Manson 2007) reviewed application of machine learning on structural health monitoring with several case studies. They conclude from their review that the major problem of machine learning-based change detection remains the robustness of damage detection in the presence of EOCs. To address this problem, a PCA-

based damage detector is reported by Yan et al. (Yan et al. 2005) that can successfully detect damage from accelerometer data collected in a laboratory wooden bridge experiment under temperature variation created by a fan. An AR-based damage identification procedure for vibration-based monitoring is reported in Figueiredo et al. 2011 in which environmental variations were simulated by different mass-loading and stiffness conditions. Moving Principle Component Analysis (MPCA) is reported by Laory et al. 2011 to eliminate the effects of seasonal temperature variations in bridge damage detection.

The results presented in these studies confirmed the potential of machine learning algorithms to detect damage under varying EOCs. However, most of these algorithms are illustrated by experiments where environmental variations are simulated. In fact, it is seldom reported how such damage detection algorithms work under realistic environmental and operational conditions. There is limited research focused on application of Machine Learning and Signal Processing algorithms for guided wave SHM applications.

In summary, the state-of-the-art methods for damage detection in mechanical and civil structures have achieved some success in the past few years detecting damage from a small volume of records collected under controlled and/or simulated conditions, and there have been an increasing number of attempts to address the issues caused by environmental and operational variations in specific applications. However, among all issues in this domain, considering the problem of guided wave SHM system, effective extraction and identification of damage-sensitive features from large volume of monitoring records that are robust to varying EOCs remains unsolved. We address this challenge by developing a singular value decomposition method that process batches of guided wave records collected over the monitoring periods. We then extended the method for online monitoring in *0*.

1.4 Dissertation Outline

In the following chapters, we establish a data driven framework based on singular value decomposition to extract uncorrelated components from ultrasonic records, we leverage the framework to detect and localize scatterers and corrosion pits in complex environments, and we then refine the framework to conduct online monitoring. In *Chapter 2*, we demonstrate our field experimental activities, featuring the instrumentation of a pressurized hot-water pipe in real operation, as a case study to better understand the environmental and operational variations associated with practical guided wave monitoring. In *Chapter 3*, we review singular value decomposition and derive a damage-sensitive feature selection process to establish our singular value decomposition framework, and we analyze its performance to extract damage-sensitive information in different simulated scenarios using synthetic ultrasonic records. In *Chapter 4*, we apply our singular value decomposition framework to detect scatterers from field experimental data that we collected from the pipe instrumented as described in *Chapter 2*. We show that despite the complex environmental and operational conditions, our framework successfully detects the scatterer with satisfactory robustness and accuracy. In *Chapter 5*, we integrate the singular value decomposition framework with delay-and-sum localization to conduct robust damage localization. We show that our singular value decomposition framework can remove the baseline record more robustly than the state-of-art temperature compensation and baseline subtraction routine, and significantly improve localization accuracy in certain circumstances. In *0*, we enhance our singular value decomposition framework with the capability to adaptively update singular vectors as new measurements are collected, and we use the enhanced framework to detect and localize damage in experimental data in an online manner. In *Chapter 7*, we will summarize our

contributions, discuss how this framework can be improved, and discuss how this methodology can be extended to other structures, materials, and applications.

Chapter 2 Field Experimental Case Study: Guided Wave Monitoring Of A Pressurized Hot Water Pipe In Operation

In this chapter, we introduce our field experiments instrumenting ultrasonic pitch-catch monitoring on an operating hot water pipe system in its normal operating conditions in discrete time periods spanning seven months. We conducted this field experimental investigation to better understand the real environmental and operational variations that would affect practical guided wave structural health monitoring. In *Section 2.2*, we describe our field experiment environment, our experimental setup and data acquisition, and our data management. Then in *Section 2.3*, we show examples of field ultrasonic monitoring records and establish pre-processing routines to clean the receiving signals. We then investigate the environmental and operational variations that affect the ultrasonic records. We found that such benign variations can drastically change the ultrasonic records, making it very difficult to detect subtle damage on the pipe using conventional methods.

2.1 Motivation

Pipelines and piping systems are critical components of modern infrastructure. They transport gases and liquids that support many industrial and civil activities. For example, energy is transported from production basins to areas of consumption in transmission and distribution pipelines, totaling 2,562,000 miles (PHMSA 2013); pressurized gases or liquids, often hazardous, are conveyed through pipe systems in factories for many manufacturing processes; the urban water pipeline network plays an important role in our living standard. Damage to these pipes would have significant consequences, including economic and environmental costs. Therefore, it is important that pipelines and piping systems be well monitored and maintained.

As an example, transmission pipelines play a critical role in energy distribution in the US (Pipeline & Hazardous Materials Saftey Administration n.d.). Table 1 shows statistics extracted from PHMSA (Pipeline & Hazardous Materials Saftey Administration n.d.), summarizing reported transmission pipeline incidents and associated loss in the twenty-year period from 1993 to 2012.

Table 1 National All Pipeline Systems: Significant Incidents Summary Statistics: 1993-2012

Year	Number	Fatalities	Injuries	Property Damage
Totals	5,613	367	1,465	$6,416,454,070
2013 YTD	99	5	17	$62,919,232
3 Year Average (2010-2012)	264	14	71	$678,422,202
5 Year Average (2008-2012)	269	12	66	$554,158,011
10 Year Average (2003-2012)	282	14	59	$499,181,181
20 Year Average (1993-2012)	281	18	73	$320,822,704

According to Table 1, the 5,613 incidents reported in the past two decades cost a total of $6.4 billion of property damage, with an increasing trend in the yearly loss. Besides, almost one-fifth of the reported incidents involved casualties, causing a total of 367 fatalities and 1,465 injuries in the past twenty years. Moreover, many of these incidents pollute or contaminate the nearby environment. Such secondary damage would often take years to recover from, therefore continuously harming nearby communities long after the incidents.

PHMSA also analyzed the causes of those incidents reported in the twenty-year period. Figure 2-1 shows a breakdown of the causes of the significant incidents, in which over 75% of the "significant incidents" are caused by structural or material failure, causing a total loss of $5.58 billion. Many of these incidents could be avoided if defects on the pipe could be detected before causing failure. Moreover, the losses could be alleviated if the damage, once happened, could be detected and localized in a timely fashion using structural health monitoring system.

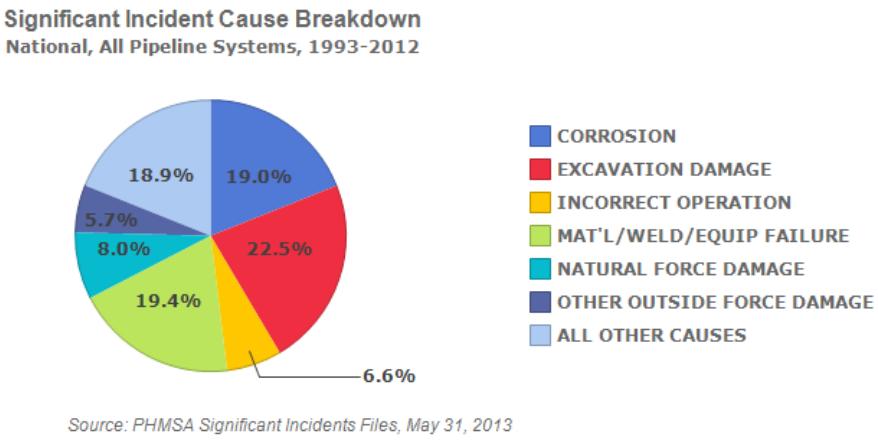


Figure 2-1 Significant incident cause breakdown, 1993-2012 (PSHMA)

However, the real-world application of developed SHM systems is still quite limited. For example, according to PHMSA, among 1,337 release incidents in transmission pipelines

reported during 2010 to 2012, less than 10% of the incidents were identified by data acquisition and testing systems. This fact implies two possibilities: (1) sensing and data-driven methods of SHM are not as widely used as traditional methods; or (2) the currently employed sensing and data-driven SHM systems do not perform reliably as expected in real-world applications. Among many challenges for improving the performance of data-driven SHM systems, identifying and characterizing the effects of various EOC have been known as major difficulties in analyzing data.

In order to better understand the challenges associated with practical implementation for pipe damage detection, we experimentally implemented a guided wave structural health monitoring system on a field pressurized hot-water pipe systems. We first illustrate our experimental setup in *Section 2.2*, and then examine the field experimental records to investigate the effects of environmental and operational conditions in *Section 2.3*.

2.2 Experimental Setup

2.2.1 Field environment

The tested pipe segment is located in a mechanical space (roughly 700 m^2) in Wean Hall, a 27,000 m^2 nine-floor campus building at Carnegie Mellon University (Facility Management Services n.d.). This mechanical space houses large pipes and mechanical equipment that serve various heating, ventilation, and air-conditioning (HVAC) requirements of Wean Hall, as well as Doherty Hall, an adjacent ten-floor office building. The HVAC equipment operates continuously to supply heating, cooling and domestic water to more than a thousand offices and classrooms. Figure 2-3 shows the photos of the second floor of the

mechanical space from two angels, in both of which a black dashed arrow indicates the subject pipe segment.

Figure 2-2 Photo of the mechanical space in Wean Hall, CMU, where field experiments were conducted. The subject pipe segment is labeled with black dashed arrow.

The subject pipe segment is a NPS 10, Schedule 40 steel pipe with 10.75 in. (273.05 mm) outer diameter and 0.365 in. (9.271 mm) wall thickness, covered by 2 in. (50.8 mm) thick fiberglass insulation. Figure 2-3 shows the dimensions of the pipe segments. The test pipe consists of a straight segment of roughly 35 feet with a heat exchanger at one end, and an elbow terminated with a ring flange at the other end. We chose to instrument the 30 ft. straight section of the pipe, which is supported by four roll chairs hung from the floor framing overhead.

The subject pipe conveys pressurized hot water in continuous operation. Because of periodic cycling, the hot water flow rate varies from 200 to 450 gpm, and the water temperature fluctuates from 100 to up to 140 ^{o}F. The mechanical space itself contains high

levels of mechanical and electrical noise created by heating, ventilation, and air conditioning equipment. These environmental and operational conditions (EOC) are not present in a laboratory experiment, thereby making the field instrumentation a more practical testbed.

2.2.2 Data acquisition

We removed short sections of insulation at four locations to expose areas of the pipe surface and to mount piezoelectric transducers. At one such location we observed a welded joint, as indicated by red in Figure 2-3. The weld affects the wave transmission along the pipe, but our experimental observations suggest that the effect is negligible compared with the change produced by environmental and operational conditions. Figure 2-3 also shows a schematic drawing of the placement of PZT transducers. We placed arrays of 4 or 5 transducers at three locations 10 ft. (3.05 m) apart from each other. The wafers are lead zirconium titanate (PZT, PSI-5A4E from PIEZO SYSTEMS, INC (Piezo systems n.d.)) and are permanently adhered to the pipe using temperature-resistant epoxy. The rightmost transducer array shown in Figure 2-3 was not used in our studies because there is no effective acoustic coupling through the ring flange.

The sensors are connected to a National Instruments PXI data acquisition (DAQ) system. The DAQ system was configured to take pitch-catch records in which ultrasonic waves are generated from one transducer and recorded at another some distance away. An NI-5122 arbitrary waveform generator was used to excite a wideband windowed-sinusoidal pulse at the transmitter. Multiple NI-5142 digitizers were used to collect voltage readings from different receivers.

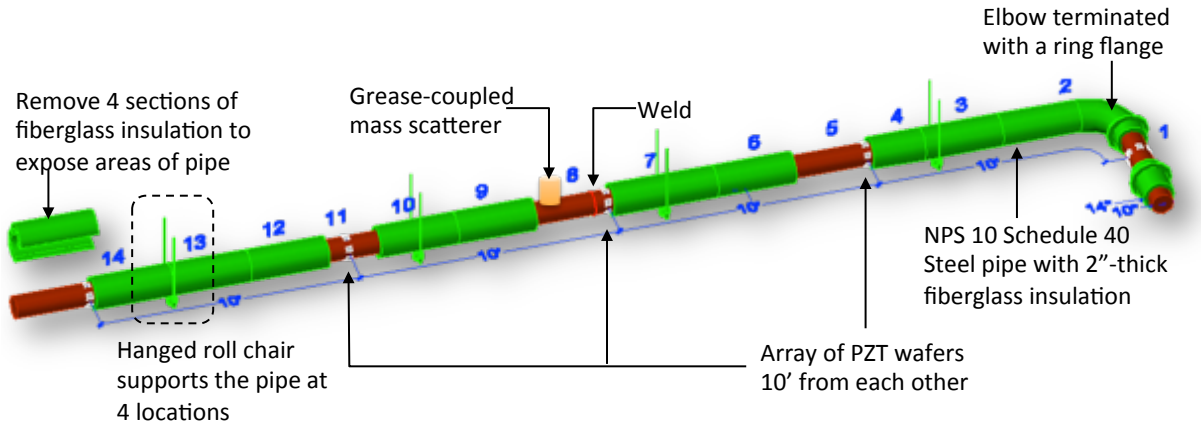


Figure 2-3 Schematic drawing of the instrumented pipe segment shows pipe dimensions and sensor placement.

Figure 2-4 shows some details of the experimental setup. Figure 2-4(a) shows the detail of the hanging roll chair that supports the tested pipe segments at four locations. The supports do not contact the pipe directly because of the layer of insulation, and do not significantly affect the guided wave propagation along the pipe. Figure 2-4(b) shows the grease-coupled mass scatterer we used as a reversible, physical simulation of damage. We apply the mass scatterer (35 mm diameter bronze cylinder) on the pipe surface, and use vacuum grease to provide acoustic coupling between the pipe and the mass. The mass acts as a subtle scatterer similar to a partial thickness crack or an area of corrosion. The mass scatters ultrasonic waves propagating through the grease-coupled area, thereby creating changes in the received records. Figure 2-4(c) is a photograph of an exposed area of the pipe where we observe the welded joint, along with a nearby transducer.

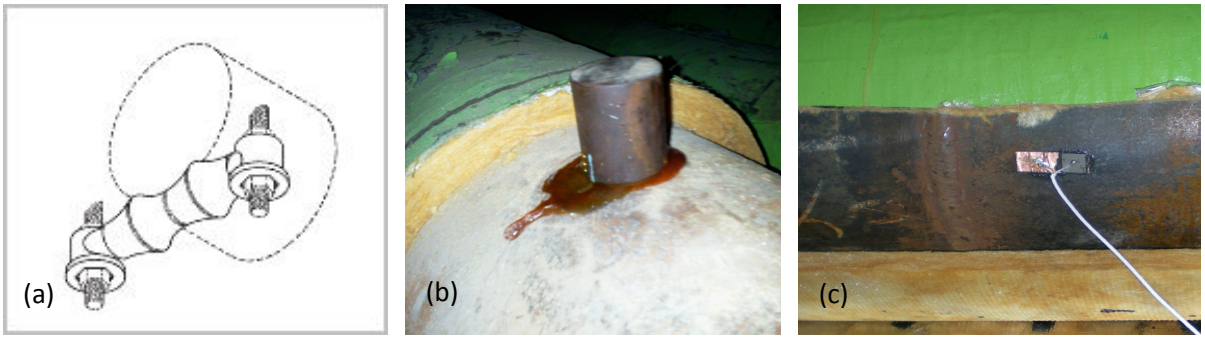


Figure 2-4 Details of the experimental setup. (a) Detail of roll chair that supports the pipe segment; (b) Grease-coupled mass scatterer used as physical simulation of subtle damage; (c) One of the instrumented PZT wafer that is close to a welded joint on the pipe

We transmit from one transducer and record the received signals from each transducer at another array that is 10 ft. (3.05 m) distant. We apply a sinc windowed sinusoidal pulse excitation with a frequency band of 100 kHz to 400 kHz, and record a 10 ms ultrasonic signal (sampling every 1 μs) at each receiver. Each pitch-catch record consists of 10,000 samples. We repeated the process at one-minute time intervals continuously for many hours.

In our experiment, we first took pitch-catch records at 1-min intervals with no scatterer present; we then attached the scatterer and sought to detect a change in the pitch-catch records. Usually, in one experiment, we remove and then re-apply the mass scatterer several times, at the same location or at different locations, to study the performance of our method.

2.2.3 Data management

With the data acquisition system illustrate above, we conducted field experiments from June 2011 to February 2012. In a typical experiment, we took pitch-catch monitoring records at one-minute interval for a period of days, during which time we toggled the mass scatterer several times. We collected a total of over 10,000 hours of data, obtaining over 140,000 pitch-catch records. The ultrasonic records were stored in LabVIEW measurement files (.lvm

(National Instruments Corporation 2010)), and were then fetched and processed using MATLAB.

For convenience, we define two time scales, 'slow time' and 'fast time', to characterize our ultrasonic monitoring records. We use 'slow time' to refer to the time scale associated with the interval (usually in minutes) in between records. In contrast, 'fast time' corresponds to the sampling of voltage readings in one pitch-catch record. For example, in a typical experiment where we collected 1,000 monitoring records at one-minute interval, and each record consist 10,000 samples at 1MHz sampling frequency. Then we have 1,000 ticks in the slow-time dimension, and the total length is 1,000 minutes; we have 10,000 ticks in the fast-time dimension, and the total length is 10 milliseconds. For the rest of the dissertation, we will refer to the two dimensions of our data matrix with these two time scales.

In some experiments, we also measured the temperature and flow rate at the same time when we take ultrasonic monitoring records. We took the environmental measurements with one-minute interval, the same time interval that was used as we collected ultrasonic records. The temperature was measured with a wireless thermocouple that is attached to the pipe wall at a location near the transmitter. The flow rate measurements are made with a Doppler-based ultrasound flow rate gauge (Shenitech STUF-300H), which is attached to the pipe at a downstream location that is 20 feet (6 m) away from the transducer array. The environmental measurements were stored in comma-separated value (.csv) files. The photo in Figure 2-5 shows the flow rate gauge attached to the pipe.

Figure 2-5 Photo of the Doppler-based flow rate gauge on the hot water pipe in field experiments

2.3 Sample Ultrasonic Records And The Effects Of EOC

2.3.1 Sample received records

Figure 2-6 shows an example pitch-catch record. The two plots on the top show a typical excitation signal in both the time and frequency domains. The middle plot shows the raw signal recorded at a receiver, which displays substantial noise caused by EOC. The bottom plot shows the same record filtered by a 10 kHz high-pass Butterworth filter. The filtering successfully eliminated low-frequency mechanical noise that is not relevant. All records were subsequently filtered with the 10 kHz high-pass filter before further analysis. Due to the multi-modal and dispersive characteristic of pipe guided waves, and other EOC effects that are not limited to the low-frequency range, the filtered signal still displays complicated structure with many arrivals, making it difficult to directly interpret the records.

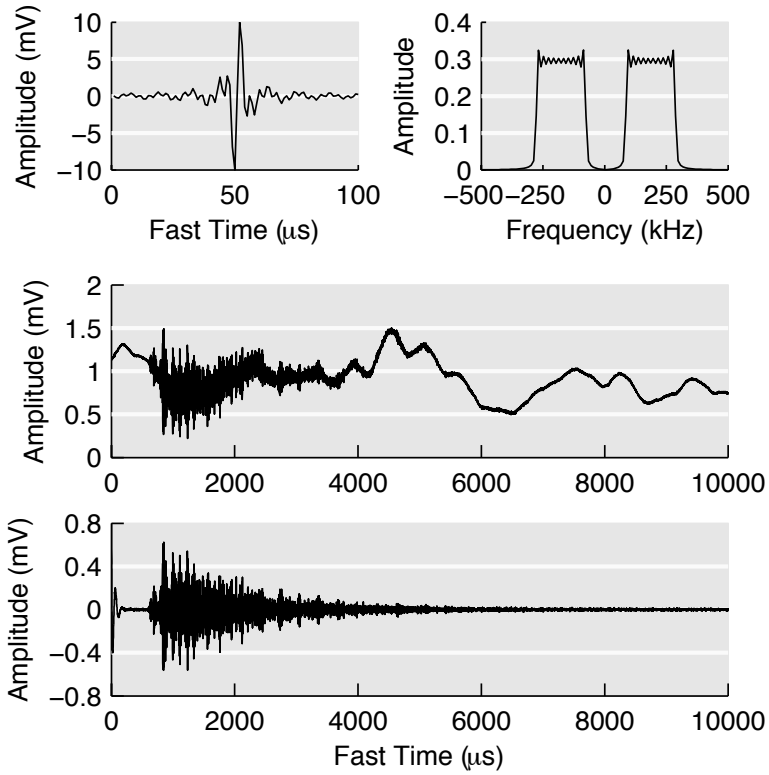


Figure 2-6 Example record from field experiments (a) excitation signal; (b) raw received signal; (c) received signal filtered with 10kHz high-pass filter

2.3.2 Effects of EOC variations

Figure 2-7 shows a portion of the field experimental data collected on August 31, 2011. In that experiment, we collected ultrasonic records and EOC readings every minute for about 26 hours. The top plot of Figure 2-7 shows pipe wall temperature as a solid line, which varied between 103 °F to 109 °F, and the water flow rate as a dash line, which varied between 280 to 420 gpm.

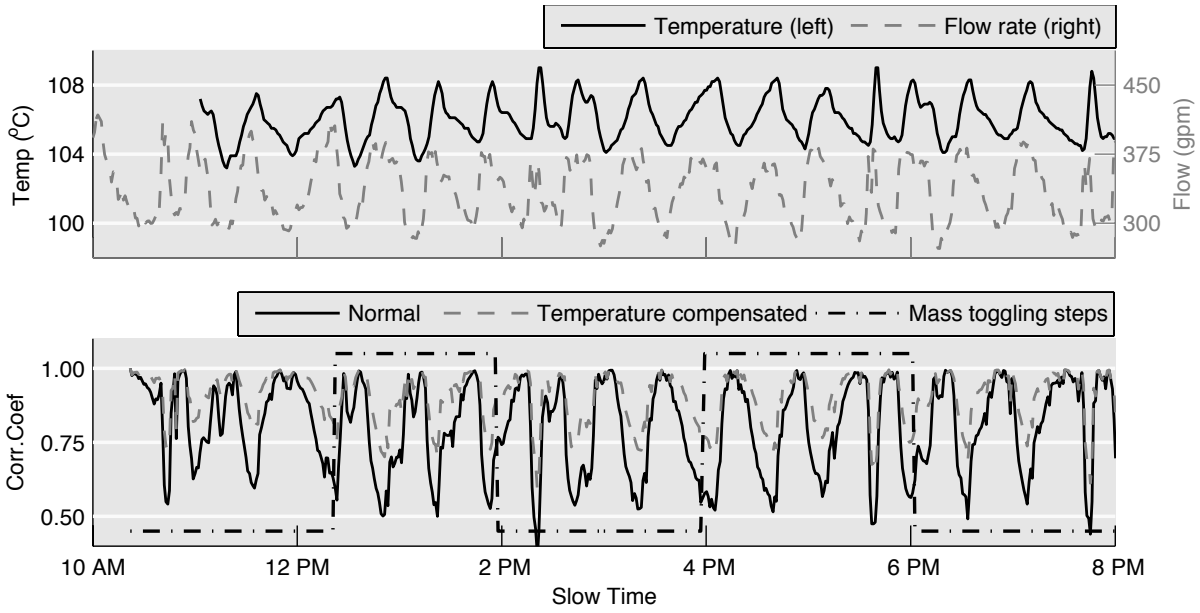


Figure 2-7 Portion of field experimental data collected on 08/31/2011. The correlation coefficients of received signals (before and after temperature compensation) were shown along with field measurements of temperature and flow rate.

One way to characterize the change caused by EOC is to use correlation coefficients, which is the inner product of two waveforms $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, normalized by their norms,

$$C\ (\boldsymbol{x}_1, \boldsymbol{x}_2) \stackrel{\text{def}}{=} \frac{\sum \boldsymbol{x_1 x_2}}{\|\boldsymbol{x_2}\| \|\boldsymbol{x_2}\|}. \tag{2-1}$$

Equation (2-1) is equivalent to the normalized cross-correlation of $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$ at zero-lag, similar to local temporal coherence defined in (Michaels and Michaels 2005)(Michaels 2008), where local indicates that the cross-correlation is computed in the fast-time domain. The correlation coefficient is a scalar value between [-1,1], indicating the similarity between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. A coefficient of 1.0 indicates that $\boldsymbol{x}_1$ is identical to $\boldsymbol{x}_2$, and a coefficient of -1.0 indicates the presence of a 180° phase difference between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. We take thousands of records at one-minute intervals, and then compute the correlation coefficients between each record $\boldsymbol{x}_i$ with the first record of the day $\boldsymbol{x}_0$. These coefficients characterize similarity

between the baseline and every other record over time. In an ideal laboratory environment without EOC variations the plot of correlation coefficients would show a straight line with a constant value near 1.0. In that ideal case, if we were to introduce a change (e.g., attaching the mass scatterer), we would observe a one-step drop in the correlation coefficients, which would indicate damage. However, the presence of EOC variations could also change the ultrasonic records, and thereby change the correlation coefficients.

During the 26-hour period of this experiment, we placed a mass scatterer on the pipe to physically simulate a small, localized defect. From the correlation coefficient, which is sometimes used as a damage indicator in the literature (Michaels and Michaels 2005)(Liu, et al. 2012), we see no indication of the presence of the mass scatterer. Instead, we observe large cyclic effects roughly corresponding to periodic EOC variations. After temperature compensation, the correlation coefficients (in dashed line) increase closer to 1.0, indicating some of the effects were compensated. However, the results with temperature compensation show no indication of the mass scatterer, which was toggled 4 times during the 10 hours, as shown in the dash-dot line in Figure 2-7.

2.4 Summary

In this chapter, we introduced our field experiments instrumenting a pressurized hot water pipe in its normal operation. We described our experimental setup and showed example pitch-catch records collected from this real-world testbed. We established basic signal processing steps to pre-process the received records to remove some environment noise and to obtain recognizable guided wave signals.

We observe from the data that the environmental and operational conditions can vary significantly in the normal working conditions of the tested structures. These benign variations produce drastic changes in the ultrasonic records, and mask the information of interest that is caused by any damage.

In some experiments, we placed a grease-coupled mass scatterer on the pipe to physically simulate subtle damage and sought to detect the presence of that scatterer from the received ultrasonic records. However, the large environmental and operational variations made it impossible to obtain any useful indications of damage using conventional available methods. We tested different methods that have been effective in ideal laboratory experiments, or in experiments with limited environmental and operational noise, but none of them detected the presence of the scatterer in our experimental records from a real-world environment. Therefore, we need to develop a new data analysis method, robust to environmental and operational variations, that can detect scatterer-induced change in our ultrasonic records.

In *Chapter 3*, we introduce the damage detection and localization framework we developed based on singular value decomposition. We demonstrate that the framework can effectively extract damage information from batches of ultrasonic monitoring records. We will use synthetic simulation to validate the framework on a preliminary basis, and we will then apply the framework to the field experiment dataset introduced in this chapter. We will then show successful damage (scatterer) detection in *Chapter 4*.

Chapter 3 Singular Value Decomposition On Ultrasonic Monitoring Records

In this chapter, we introduce the damage detection and localization framework we developed based on singular value decomposition. In *Section 3.1*, we review the basics of singular value decomposition and reveal its effects on a batch of ultrasonic monitoring records – that different variations will be separated into different singular vectors, such that we can easily identify and extract the change produced by damage. In *Section 3.2*, we developed a scoring system based on K-means to help identify damage-sensitive singular vector from the singular value decomposition of ultrasonic records.

Then in *Section 3.3*, we analyze the performance of singular value decomposition. We established the analysis in three scenarios, in which we calculate SVD on records that contain no variations, on records that are affected by changes produced by damage, and on records that are affected by other variations. We then further validate the analysis with examples in *Section 3.4*, where we simulate the three scenarios with synthesized ultrasonic monitoring

records. We demonstrate that our SVD framework can effectively extract damage information from batches of ultrasonic monitoring records in different scenarios, from the case where there is no variations to cases where the monitoring records are affected by large slow-time variations as we observed in the field experimental case study in *Chapter 2*.

3.1 Singular Value Decomposition

Singular value decomposition (SVD) seeks a linear decomposition of a data matrix that creates a set of orthonormal bases to represent the data. It is extensively studied and used in many fields, including face recognition (Turk and Pentland 1991), audio/video compressing (Konstantinides and Yovanof 1995), and signal de-noising (Wei et al. 2001). For a general matrix $\boldsymbol{X}^{\mathrm{M\times N}}$ with dimensions $M \leq N$, the SVD of $\boldsymbol{X}$ is defined by

$$\boldsymbol{X} = \boldsymbol{USV}^{H}. \tag{3-1}$$

where, superscript H is the conjugate transpose. In equation (3-1), $\boldsymbol{U}^{M\times M} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_M]$ is the left singular vector matrix, and each $\boldsymbol{u}_{\mathrm{i}}$ is a $M\times 1$ vector represents the projection of $\boldsymbol{X}$ in one singular subspace defined by $\boldsymbol{v}_i$; $\boldsymbol{S}^{\mathrm{M\times M}}$ is the diagonal singular value matrix with diagonal terms $[\sigma_1, \sigma_2, \dots, \sigma_M]$ that are known as singular values of $\boldsymbol{X}$, typically in a descending order; $\boldsymbol{V}^{N\times M} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_N]$ is the right singular vector matrix, and each of the $\boldsymbol{v}_{\mathrm{i}}$ vectors represents a basis that describes a signal subspace. SVD is directly related to principal component analysis (PCA) by the following equation

$$\begin{aligned} \boldsymbol{XX^H} &= (\boldsymbol{USV^H})(\boldsymbol{VS^HU^H}) = \boldsymbol{U\Sigma U^H}, \\ \boldsymbol{X^HX} &= (\boldsymbol{VS^HU^H})(\boldsymbol{USV^H}) = \boldsymbol{V\Sigma V^H}. \end{aligned} \tag{3-2}$$

in which we utilize the fact that $\boldsymbol{U}, \boldsymbol{V}$ are orthogonal and unitary, such that $\boldsymbol{U^H U = I}$ and $\boldsymbol{V^H V = I}$. From (3-2), we can see that the $\boldsymbol{U}$ captures the variations in one dimension of the data matrix, while $\boldsymbol{V}$ captures the variations in another dimension.

In the context of ultrasonic monitoring, suppose we collect N pitch-catch records over a period of time during an experiment, where each record has D voltage readings. The data matrix $\boldsymbol{X}$ can be written as the collection of N records of length D denoted by row vectors $\boldsymbol{x}_i^T$.

$$\boldsymbol{X}^{N\times D} = [\boldsymbol{x}_1^T, \boldsymbol{x}_2^T, \dots, \boldsymbol{x}_N^T]. \tag{3-3}$$

For convenience, we define the two dimensions of the data matrix as 'slow time', and 'fast time'. 'Slow time' refers to the time scale associated with the interval (usually in minutes) in between records, and is characterized by the number of records N. 'Fast time' corresponds to the sampling of voltage readings in one pitch-catch record, and is characterized by the number of samples D.

When we apply SVD to such a data matrix, we obtain,

$$\boldsymbol{X}^{N\times D} = \boldsymbol{U}^{N\times N}\boldsymbol{S}^{N\times N}(\boldsymbol{V^H})^{N\times D}. \tag{3-4}$$

Each column in $\boldsymbol{U}$ represents some variation among the N records that represents a 'trend' of changes over slow time. Because of the orthogonality of singular vectors, $\boldsymbol{u}_i$ are orthogonal to each other, which means that different trends are separated into different de-correlated singular vectors. On the other hand, the $\boldsymbol{V}$ matrix is an orthonormal basis in fast time for explaining the variation in slow time. We can see this by multiplying both sides of (3-1) by $\boldsymbol{V}$, and use the fact that $\boldsymbol{V}$ is a unitary matrix, i.e. $\boldsymbol{V^H V = I}$.

$$\begin{aligned} \boldsymbol{XV} &= \boldsymbol{USV^H V} \\ \boldsymbol{XV} &= \boldsymbol{US} \end{aligned} \tag{3-5}$$

$$\boldsymbol{X}^{N\times D}\boldsymbol{v}_i^{D\times 1} = \boldsymbol{u}_i^{N\times 1}\sigma_i.$$

Equation (3-5) demonstrates that each pair of left/right singular vectors characterizes one component of the variations of the data in slow/fast time, respectively. Figure 3-1 shows an experimental example of applying SVD to ultrasonic records to further illustrate the behavior of singular vectors. In the experiment, we attached a pair of PZT transducers to a pipe. We collected three baseline pitch-catch records when the pipe was intact, each consisting of 2000 voltage measurements sampled at 1 MHz. We then placed a mass scatterer on the pipe to physically simulate a small flaw and collected three additional records. The top left plot of Figure 3-1 shows each row from the data matrix $\boldsymbol{X}^{6\times 2000}$. The top right plot of Figure 3-1 shows the intact baseline signal $\boldsymbol{x}_b$ and the scatter signal $\boldsymbol{x}_s$, which is the difference signal before and after the mass scatterer was introduced. The records were collected under ideal laboratory environment with little noise or environmental variations. Therefore, the only change in slow time corresponds to the application of the mass scatterer.

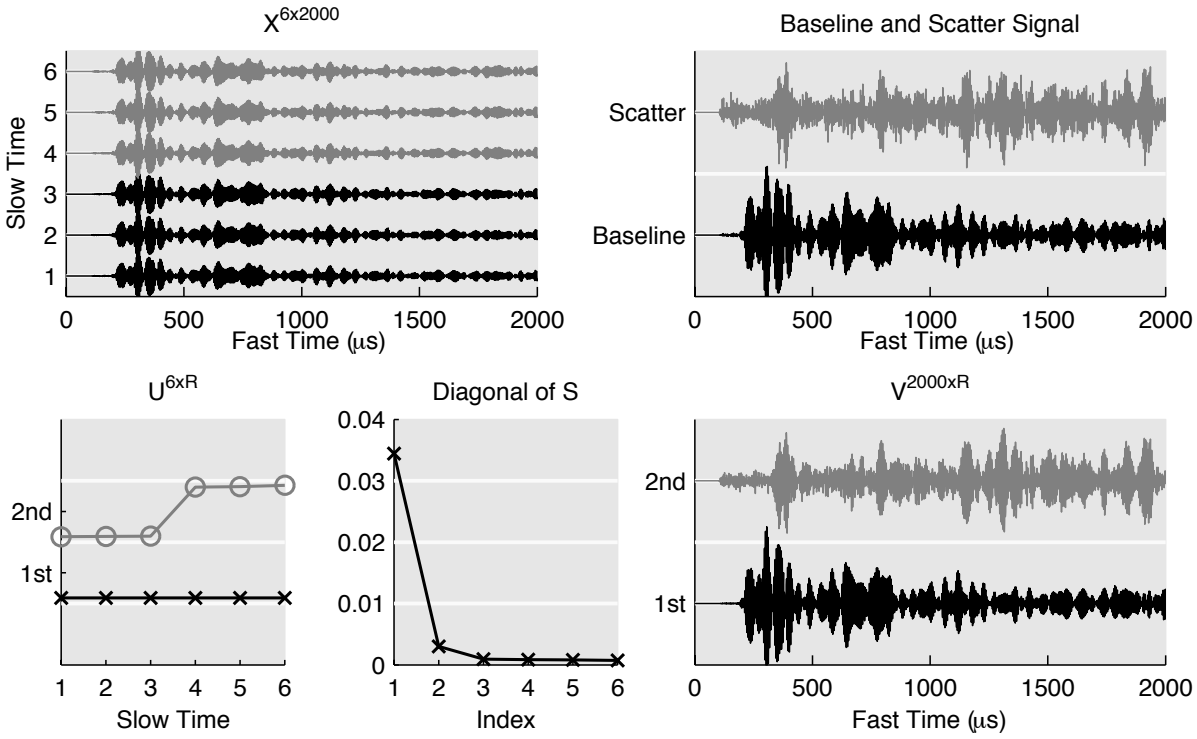


Figure 3-1 SVD on ultrasonic records collected under steady environment. **Top left**: Data matrix $\boldsymbol{X}$ containing 6 pitch-catch records; **Top right**: Intact baseline and scatter signal produced by the mass scatterer; **Bottom row from left to right**: Left singular vector matrix $\boldsymbol{U}$, singular values $\boldsymbol{S}$, and right singular vector matrix $\boldsymbol{V}$

The three plots in the bottom row of Figure 3-1 show, respectively, the $\boldsymbol{U}, \boldsymbol{S},$ and $\boldsymbol{V}$ matrices obtained by applying SVD to $\boldsymbol{X}$. The diagonal of $\boldsymbol{S}$ has only two non-zero values, which means that there are only two meaningful components in the dataset. We therefore plot the first two singular vectors from $\boldsymbol{U}$ and $\boldsymbol{V}$. As a flat line, $\boldsymbol{u}_1$ represents a steady 'trend', which means there is an unchanged component across the six records. Correspondingly, the first right singular vector $\boldsymbol{v}_1$ equals the baseline signal $\boldsymbol{x}_b$. The second left singular vector $\boldsymbol{u}_2$ shows a step that rises between the third and the fourth records, which corresponds to the introduction of the mass. The corresponding second right singular vector $\boldsymbol{v}_2$ equals the scatter signal (the difference signal) $\boldsymbol{x}_s$. The correlation coefficient between $\boldsymbol{v}_1$ and $\boldsymbol{x}_b$ and between $\boldsymbol{v}_2$ and $\boldsymbol{x}_s$ both exceed 0.99.

If we apply SVD to data collected under other environmental and operational variations, because of the orthogonality between singular vectors, we expect to observe the EOC variations and the damage in different left singular vectors. Figure 3-2 shows an example applying SVD to a synthesized experimental dataset. We first collected pitch-catch records from a pair of PZT transducers attached to an aluminum pipe specimen in a laboratory environment, both with and without the presence of a mass scatterer on the pipe. We then synthesize the records with temperature variations by stretching each record with a series of different stretching factors. As presented in temperature compensation literature (e.g. (Clarke et al. 2010; Croxford et al. 2010)), variations in temperature can be approximated as a stretching over time. We synthesized two sets of data. In first set of data, we synthesized 300 intact, damage-free, records with a sinusoidal temperature variation. In the second set of data, we synthesized 150 intact records and 150 records that were stretched copies of records collected with the presence of scatterer. The second data set is given the same sinusoidal temperature variation used in first set. By applying SVD on the datasets, we obtained collections of left singular vectors, each of which represents some variation over slow time. Figure 3-2 shows the first eight singular vectors for the data sets without the mass scatterer (on the left) and with the mass scatterer (on the right).

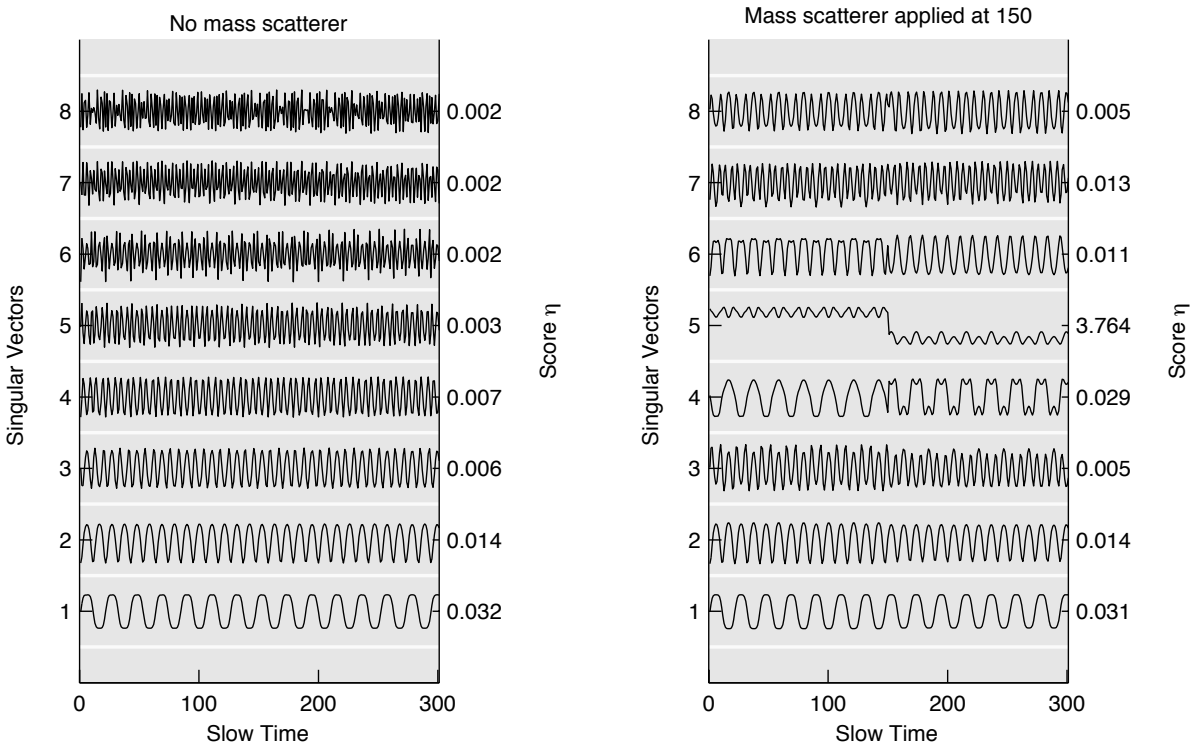


Figure 3-2 First eight left singular vectors on experimental data synthesized with temperature stretching. **Left**: Left singular vectors of 300 records without mass scatterer. **Right**: Left singular vectors of 300 records. The later half of each record was synthesized from a record collected when the mass scatterer was present.

In the right plot of Figure 3-2, the fifth singular vector visually shows a large step when the mass scatterer is introduced. In contrast, this step is not seen in the left plot. This implies that the fifth left singular vector contains significant information about the damage, i.e. it is a *damage-sensitive feature*. We also observe that the other singular vectors correlate to some periodical variations that are related to the temperature variation we synthesized. Because the order of the singular vectors is related to their relative strengths, the damage information will not always be found in the fifth singular vector. In the following section, we develop a method to select the damage-sensitive singular vectors from a collection of singular vectors.

3.2 Damage-Sensitive Singular Vector Selection

From *Section 2.1*, we showed how SVD generates R features from a batch of records, where $R \ll \min(D, N)$ indicates the number of non-trivial singular vectors. In Figure 3-1 and Figure 3-2, we observed that one of the singular vectors is damage-sensitive and robust to EOC changes; however, which singular vector is damage-sensitive cannot be determined analytically. Therefore, we have developed a feature selection scheme based on K-means clustering to identify the damage-sensitive singular vector. Abrupt changes in slow time can be abstracted as a step function, where the step corresponds to the instant in slow time when damage is introduced. We cluster each of the singular vectors into two groups, and then compute a score reflecting the performance of the clustering. The singular vector with highest score best resembles a step function, and therefore is the damage-sensitive singular vector. If the scores do not exceed a prescribed threshold, we conclude there is no damage in the batch of records. Here, we only discuss the simple scenario with only one abrupt damage in the monitoring period. Such a scoring system may be generalized to handle multiple damages or progressive damage.

For each of the left singular vector $\boldsymbol{u_i}^{N\times 1}$, we cluster the N values $\{u_{ij}, j = 1,2,\dots,N\}$ into two clusters ($\boldsymbol{C}_0$ and $\boldsymbol{C}_1$) using K-means with the number of clusters $k = 2$. C_0 contains all the values $\{u_{ij}\}$ in class zero, and C_1 contains values in class one. The sum of lengths of $\boldsymbol{C_0}$ and $\boldsymbol{C_1}$ is N. For convenience, we also hold the class of u_{ij} in a vector $\boldsymbol{\ell}^{1\times N}$, such that $\ell_j = 0\ \forall u_{ij} \in \boldsymbol{C_0}$, and $\ell_j = 1\ \forall u_{ij} \in \boldsymbol{C_1}$. We then compute a score η_i to characterize how well the cluster is separated. The score is defined such that it increases with the between-cluster variation, and decreases with the within-cluster variation. We also include a penalty term as a function of the number of change points within the label vector $\boldsymbol{\ell}$. Because damage

is a permanent change, whereas other transient changes, such as temperature change, fluctuate over slow time, we define the score η as,

$$\eta_i \stackrel{\text{def}}{=} \frac{|\mu_0 - \mu_1|}{\sqrt{\sigma_0^2 + \sigma_1^2}} \times \frac{1}{\left|\dot{\boldsymbol{\ell}}\right|^2}. \tag{3-6}$$

where, μ_i is the expected value of $\boldsymbol{C}_i$, i.e. $\mu_i = \mathbf{E}[\boldsymbol{C}_i]$; and σ_i is the variance of $\boldsymbol{C}_i$, i.e. $\sigma_i = \mathbf{E}[(\boldsymbol{C}_i - \mu_i)^2]$; and $\dot{\boldsymbol{\ell}}$ is the derivative of the label vector $\boldsymbol{\ell}$ over slow time.

We compute η_i for each left singular vector $\boldsymbol{u_i}$ and obtain the vector $\boldsymbol{\eta}$. The scores in $\boldsymbol{\eta}$ characterize how likely is each singular vector to contain a step. We then assign a threshold ϵ that controls the sensitivity of the damage detection. If $\exists \eta_i > \epsilon$, then $\boldsymbol{u}_i$ is a good representation of step function, i.e. there exists an abrupt change in the dataset. If multiple scores exceed the threshold, the singular vector with the highest score is considered the one that represents that change. In our experiments, we found $\epsilon = 0.5$ achieves a good balance between flexibility and accuracy. Using the synthetic datasets in Figure 3-2 as an example, we computed the score η_i for each left singular vector, and show $\boldsymbol{\eta}$ on the right of singular vectors. As shown in the left plot of Figure 3-2, none of η_i exceed the threshold $\epsilon = 0.5$ for the intact dataset, indicating no damage-sensitive feature extracted from the dataset. In contrast, in the right plot of Figure 3-2 that shows dataset with the presence of a scatterer, the maximum score $\eta_5 = 3.764$ correctly identifies the damage-sensitive singular vector.

3.3 Performance Analysis Of Singular Value Decomposition

In this subsection, we analyze the performance of SVD on ultrasonic records, and demonstrate that the orthogonal nature of singular vectors enables us to extract scatter signals that are robust to temperature change and other variations over slow time. Singular value

decomposition produces orthogonal bases that best represent the variations in the data. When our data matrix contains both intact and damage records, the difference signal between the two – the scatter signal – is approximately orthogonal to the direct-path baseline signal. Therefore, the scatter signal will be separated into a single singular vector as an orthogonal component.

We first review basics of singular value decomposition, and then establish our analysis in three scenarios, where the records in the dataset vary.

1) Dataset contains only baseline records,

$$\boldsymbol{X} = [\boldsymbol{x}_b^T, \boldsymbol{x}_b^T, \dots, \boldsymbol{x}_b^T]. \tag{3-7}$$

2) Dataset contains both baseline records and damaged records,

$$\boldsymbol{X} = [\boldsymbol{x}_b^T, \dots, \boldsymbol{x}_b^T, \boldsymbol{x}_b^T + \boldsymbol{x}_s^T, \dots, \boldsymbol{x}_b^T + \boldsymbol{x}_s^T]. \tag{3-8}$$

3) Dataset contains both baseline records and damaged records that are also affected by other slow-time variations.

$$\boldsymbol{X} = [f_1(\boldsymbol{x}_b), \dots, f_n(\boldsymbol{x}_b), f_{n+1}(\boldsymbol{x}_b + \boldsymbol{x}_s), \dots, f_N(\boldsymbol{x}_b + \boldsymbol{x}_s)]. \tag{3-9}$$

Given a real or complex matrix $\boldsymbol{X}$, there always exists a singular value decomposition with positive singular values and singular vectors that span an orthonormal basis (Golub and Loan 1996). The first singular vectors $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$ can be obtained by maximizing the singular value,

$$\sigma_1 = \max_{\boldsymbol{u},\boldsymbol{v}} \boldsymbol{u}_1^T \boldsymbol{X} \boldsymbol{v}_1. \tag{3-10}$$

$$\boldsymbol{u}_1, \boldsymbol{v}_1 = argmax_{\boldsymbol{u},\boldsymbol{v}}\, \boldsymbol{u}^T \boldsymbol{X} \boldsymbol{v}. \tag{3-11}$$

Additional singular vectors and singular values can then be computed by sequentially maximizing the i^{th} singular value σ_i on singular vector space $\boldsymbol{u}, \boldsymbol{v}$ that are orthogonal to all previous singular vectors $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{i-1}$ and $\boldsymbol{v}_1, \dots, \boldsymbol{v}_{i-1}$. (Note that in modern computational practice, SVD is calculated numerically with a two-step iterative procedure (Golub and Kahan 1965) using linear algebra calculation routines such as LAPACK (Anderson et al. 1999). However, for discussion purposes we will use the equations above for our analysis of SVD.) Equation (3-11) indicates we obtain $\boldsymbol{u}_1, \boldsymbol{v}_1$ by maximizing the variation in the projection of $\boldsymbol{X}$ on the bases specified by $\boldsymbol{u}$ and $\boldsymbol{v}$. Additional singular vectors are then obtained by maximizing the variation of the projection of the residual of $\boldsymbol{X}$ on bases that are orthogonal to all previous singular vectors. For a batch of ultrasonic records $\boldsymbol{X}^{N\times D}$, we know that the left singular vectors represent the variation in the slow-time domain, the changes over the N measured records, and that the right singular vectors represent the variation in the fast-time domain, the change in the waveform that corresponding to the slow-time variations [].

3.3.1 Scenario 1: on records containing only direct-path signals

To analyze the approach, we first assume our data matrix only contains N identical baseline records, such that $\boldsymbol{X} = \boldsymbol{X}_{baseline} = [\boldsymbol{x}_b^T, \boldsymbol{x}_b^T, \dots, \boldsymbol{x}_b^T]$. The maximization in equation (3-11) results in,

- $\boldsymbol{u}_1 = \boldsymbol{u}_b = \vec{1}/\sqrt{N}$. The 1st left singular vector is a unit vector of constants;
- $\boldsymbol{v}_1 = \boldsymbol{v}_b = \boldsymbol{x}_b/\|\boldsymbol{x}_b\|$. The 1st right singular vector is a normalized baseline signal;
- $\sigma_1 = \sigma_b = \sqrt{N} \cdot \|\boldsymbol{x}_b\|$. The 1st singular value equals the matrix norm of $\boldsymbol{X}$.

The subscript b indicates that the singular vectors are solely correlated to the baseline signal. The three terms represent the dominant component in the data matrix, where $\boldsymbol{v}_1$ expresses the waveform shape of that component, $\boldsymbol{u}_1$ shows how the component varies over slow time, and σ_1 indicates the weight of that component. Multiplying the three terms together, we can reconstruct the component in terms of a N x D matrix.

$$\boldsymbol{u}_b \sigma_b \boldsymbol{v}_b^T = \frac{\vec{1}}{\sqrt{N}} \cdot \left(\sqrt{N} \cdot \|\boldsymbol{x}_b\|\right) \cdot \frac{\boldsymbol{x}_b^T}{\|\boldsymbol{x}_b\|} = [\boldsymbol{x}_b^T, \boldsymbol{x}_b^T, \dots, \boldsymbol{x}_b^T]. \tag{3-12}$$

After obtaining the first singular vectors, we then maximize the projection of the residual of $\boldsymbol{X}$ onto $\boldsymbol{u}$ and $\boldsymbol{v}$ spaces that are orthogonal to vectors $\boldsymbol{u}_b$ and $\boldsymbol{v}_b$. In this scenario, because the data matrix $\boldsymbol{X}$ can be represented by a replication of a sole component, which is the baseline signal, the residual is zero:

$$\overline{\boldsymbol{X}} = \boldsymbol{X} - \boldsymbol{u}_b \sigma_b \boldsymbol{v}_b^T = \boldsymbol{X} - \boldsymbol{X}_{baseline} = 0. \tag{3-13}$$

Therefore the other singular values and singular vectors are 0, and the SVD of $\boldsymbol{X}$ becomes

$$\boldsymbol{X} = \boldsymbol{u}_b \sigma_b \boldsymbol{v}_b^T + O(\sigma), \tag{3-14}$$

where, $O(\sigma)$ represents the trivial terms, in which the singular values σ are zero, or close to zero if we take computational round-off error into account. The term $O(\sigma)$ also captures random white noise in the records if any is present. This implies that removing terms with singular values smaller than a threshold can effectively remove white noise from records.

3.3.2 Scenario 2: on records containing direct-path and scatter-path signals

In our second scenario, we assume that after the n^{th} measurement the records are affected by a scatter signal $\boldsymbol{x}_s$. Therefore, our data matrix contains n baseline records and $(N - n)$ 'damaged' records, $\boldsymbol{X} = [\boldsymbol{x}_b^T, \dots, \boldsymbol{x}_b^T, \boldsymbol{x}_b^T + \boldsymbol{x}_s^T, \dots, \boldsymbol{x}_b^T + \boldsymbol{x}_s^T]$. In order to recover the

scatter signal from the records, we assume that the scatter signal is approximately orthogonal to the baseline signal, i.e. the inner product

$$< \boldsymbol{x}_b, \boldsymbol{x}_s > \approx 0. \tag{3-15}$$

This assumption is supported by experimental observations that the inner product of $\boldsymbol{x}_b^T$ and $\boldsymbol{x}_s^T$ to be close to 0 if the sensors are sufficiently far from the damage site. We also found this assumption to be approximately valid even if the records are affected by different operational and environmental variations. This is because the scatter signals and the baseline signals propagate through different paths with multi-path reflections that delay the many arrivals of the two signals nonsystematically. If we calculate the inner product of the two signals, the nonsystematic destructive interference tends to cancel out when we sum up the point-wise multiplication, making them approximately orthogonal.

Because the baseline signal $\boldsymbol{x}_b$ is usually much stronger than the scatter signal $\boldsymbol{x}_s$, we will still obtain $\boldsymbol{v}_1 \approx \boldsymbol{x}_b/\|\boldsymbol{x}_b\|$ from equation (3-11). However, that term can no longer represent all the variations in the data. The residual matrix,

$$\begin{aligned} \overline{\boldsymbol{X}} &= \boldsymbol{X} - \boldsymbol{u}_1\sigma_1\boldsymbol{v}_1 \\ &\approx \boldsymbol{X} - [\boldsymbol{x}_b^T, \boldsymbol{x}_b^T, \dots, \boldsymbol{x}_b^T] \\ &= [0, \dots, 0, \boldsymbol{x}_s^T, \dots, \boldsymbol{x}_s^T] = \boldsymbol{X}_{scatter}, \end{aligned} \tag{3-16}$$

contains n columns of zeros, and $(N - n)$ columns of scatter signals. We therefore obtain non-trivial singular vectors from maximizing the projection of $\boldsymbol{X}_{scatter}$ on u and v spaces.

$$\boldsymbol{u}_2, \boldsymbol{v}_2 = argmax_{u,v}\, \boldsymbol{u}^T\boldsymbol{X}_{scatter}\boldsymbol{v}. \tag{3-17}$$

From equation (3-17), we find that $\boldsymbol{v}_2 \approx \boldsymbol{x}_s/\|\boldsymbol{x}_s\|$, and $\boldsymbol{u}_2$ resembles a step function with a step change after the n^{th} records. We again use subscripts s for 'scatter' to indicate that

they characterize the variation produced by the introduction of a scatterer. The reason that $\boldsymbol{v}_s$ only approximates $\boldsymbol{x}_s/\|\boldsymbol{x}_s\|$ is that $\boldsymbol{v}_s$ is obtained by maximizing equation (3-17) in the space that is orthogonal to $\boldsymbol{u}_b$ and $\boldsymbol{v}_b$. If $\boldsymbol{x}_s$ is strictly orthogonal to $\boldsymbol{v}_1 \approx \boldsymbol{x}_b/\|\boldsymbol{x}_b\|$, we can find a $\boldsymbol{v}_2$ that exactly resembles $\boldsymbol{x}_s/\|\boldsymbol{x}_s\|$. Hence, the SVD of $\boldsymbol{X}$ becomes

$$\boldsymbol{X} = \boldsymbol{X}_{baseline} + \boldsymbol{X}_{scatter} = \boldsymbol{u}_b\sigma_b\boldsymbol{v}_b^T + \boldsymbol{u}_s\sigma_s\boldsymbol{v}_s^T + O(\sigma). \tag{3-18}$$

3.3.3 Scenario 3: on records containing direct-path and scatter-path signals that are affected by slow-time variations

In our third scenario, we assume that the data matrix $\boldsymbol{X} = [f_1(\boldsymbol{x}_b), \dots, f_n(\boldsymbol{x}_b), f_{n+1}(\boldsymbol{x}_b + \boldsymbol{x}_s), \dots, f_N(\boldsymbol{x}_b + \boldsymbol{x}_s)]$ not only contains baseline and damaged signals, but that these signals are also affected by other benign variations, such as uniform or non-uniform temperature change, white or color noise, or a slow-time decrease in receiving amplitude because of the degradation of transducer bonding. The time-dependent function $f_t(\cdot)$, for $t = 1,\dots,N$, represents the benign variations. These variations affect wave propagation**,** and further complicate the records in $\boldsymbol{X}$. However, these additional variations do not affect the approximate orthogonality between baselines and scatter signals, as long as they occur in slow time and affect both the baseline and the scatter signals. Therefore, we can still obtain the baseline and scatter signals from different right singular vectors.

Because the baseline is affected by the benign variations over slow time, the first right singular vector is that which resembles the average of each baseline record. More specifically, the first singular vector $\boldsymbol{v}_1$ will optimally represent all of the strong, highly correlated baselines according to (3-11). Similarly, another pair of singular vectors will approximately resemble the average of each scatter signal affected by different variations. Note that these

two singular vectors cannot explain all the variations in the data, because the records are also affected by the other variations over slow time. Therefore, more singular vectors are needed to explain the errors in both the baseline signals and scatter signals to fully represent variations in the data. The singular value decomposition of $\boldsymbol{X}$ becomes,

$$\boldsymbol{X} = \boldsymbol{u}_b\sigma_b\boldsymbol{v}_b^T + \boldsymbol{u}_s\sigma_s\boldsymbol{v}_s^T + \boldsymbol{u}_{f1}\sigma_{f1}\boldsymbol{v}_{f1}^T + \boldsymbol{u}_{f2}\sigma_{f2}\boldsymbol{v}_{f2}^T + \cdots + O(\sigma). \qquad (3\text{-}19)$$

where, subscripts $f1$, $f2$, … indicate the singular vectors corresponding to other variations. Because the singular vectors are ordered by the singular values, the index of $\boldsymbol{v}_s$ depends on the relative strength of scatter signal relative to changes produced by errors produced by other variations. However, because we have prior knowledge of the behavior of the left singular vectors corresponding to scatter signal, specifically an abrupt change will resemble a step-like function, we can identify the right singular vector that resembles the scatter signal, as previously discussed in *Section 3.2*.

In the above subsections, we demonstrated our SVD robust baseline removal and evaluated its performance on ultrasonic records in three different scenarios. We showed that by assuming the scatter signal and baseline signals are sufficiently uncorrelated, we can extract the scatter signal as a separate singular vector. The method is effective as long as our assumption of approximate orthogonality is valid, which is supported by our experimental observation that the scatter signals created by defects are usually sufficiently dis-similar with the baseline records, even if the records are affected by other variations. Because we did not specify a model for the variations on the baseline and scatter signal, we can apply this to virtually any slow-time environmental variations, including temperature variations and temperature gradients. As we have shown in *Section 2.3*, temperature gradients defeat the conventional temperature compensation methods because temperature gradients on the

structure affect the direct-path and scatter signal differently. Because the orthogonal components extracted by SVD are 'averaged' throughout the data matrix, they are less affected by the different variations of temperature, whether spatial or temporal, in any particular record.

3.4 Synthetic Simulations

We now use synthetic simulated pitch-catch records to demonstrate the analysis in the three subsections in *Section 3.3*. We first show the simulated records and the SVD results for scenarios 1 and 2, where the data matrix contains direct-path or scatter-path signals without other variations. Then for scenario 3, we synthesize various slow-time variations on the data matrix to demonstrate the robustness of our SVD method. The goal of these simulations is to show that with the assumption of approximate orthogonality between baseline signal and scatter signal, SVD is effective in extracting the scatter signal despite the change caused by other variations.

We first simulate one pitch-catch record consists of a direct-path signal from a transmitter to a receiver some distance away, by delaying a Gaussian-windowed sinusoidal excitation according to the dispersion curve of guided waves on an aluminum plate, as shown in Figure 3-3.

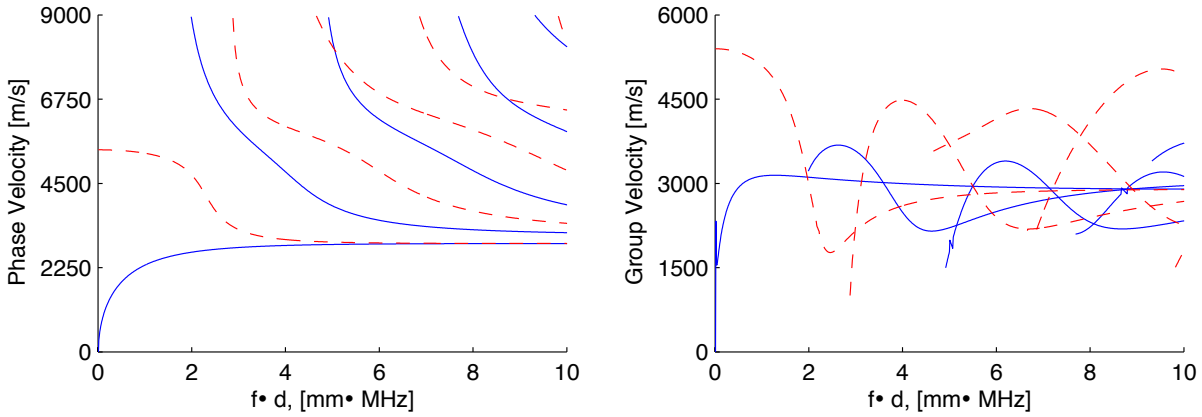


Figure 3-3 Phase and group velocity dispersion curves of guided waves on an aluminum plate

We also simulated multi-path reflections, which are simplified by adding delayed wave packets that propagate through longer distances. Note this is an over-simplified simulation for evaluating the real wave propagation in a structure, but this serves our purpose for this demonstration, which is to mimic the complex received records with multiple arrivals. With similar simplification, we produce arrivals that represent the wave propagated along a scatter path from the transmitter to the scatterer and then to the receiver. (We omit the mode conversion that will occur when a wave interacts with the scatterer. However, that mode conversion only makes the signals more complicated, and will not negatively affect our assumption of approximate orthogonality.) Figure 3-4 shows example multi-path direct and scattered signals from A0 and S0 modes. Note that in this simulation we did not take into account of mode conversion, namely the conversion of energy between different modes when guided waves interact with damage. The mode conversion effect further complicates the signals, such that our assumption of orthogonality becomes more valid.

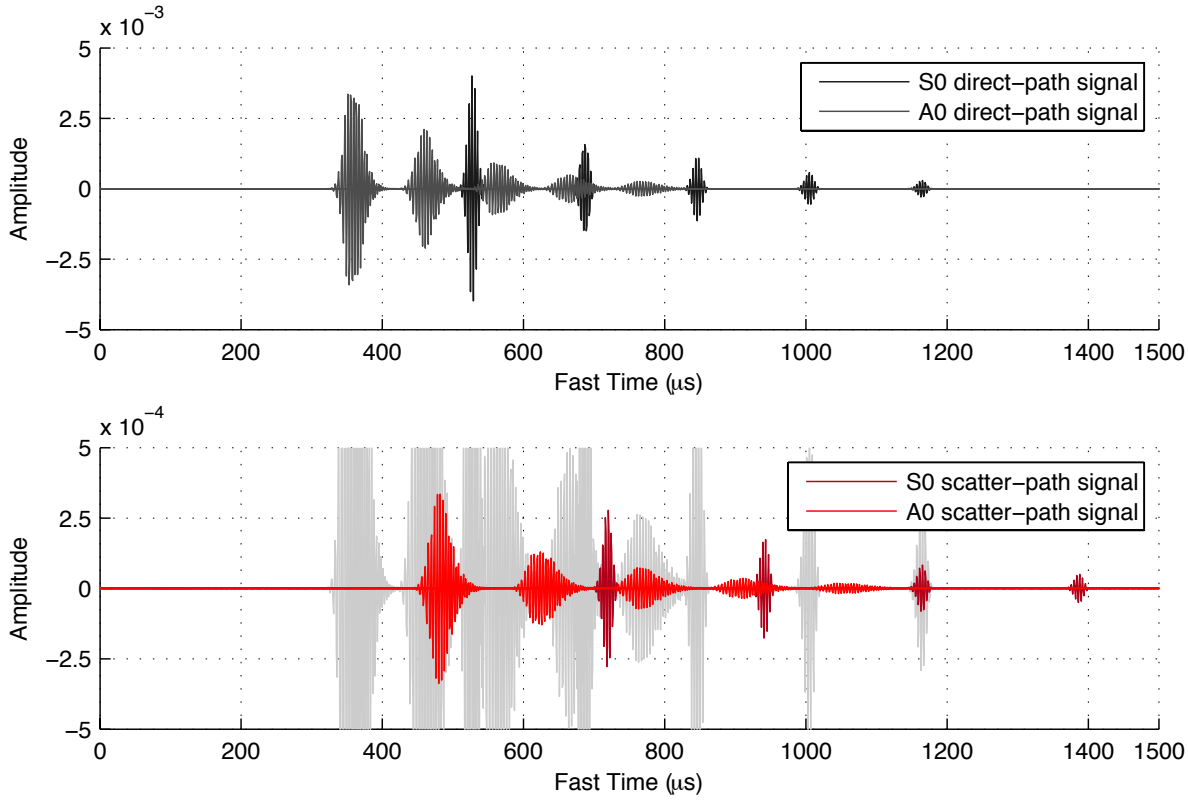


Figure 3-4 Synthetic simulations: direct path and scatter path signals.

3.4.1 Simulation for scenario 1 and 2: no slow time variations

For Scenario 1, we construct a data matrix $\boldsymbol{X} = [\boldsymbol{x}_b^T, \boldsymbol{x}_b^T, \ldots, \boldsymbol{x}_b^T]$ with 100 replications of the direct-path signal $\boldsymbol{x}_b^T$. We also add a small white noise to the data matrix, because otherwise the singular vectors with large indices will be distorted by the round-off error, which gets amplified because the singular vectors are energy-normalized. We apply SVD on $\boldsymbol{X}$, and show the obtained singular vector and singular values in Figure 3-5.

In our analysis in *Section 3.3.1*, equation (3-12) and (3-13) shows that the first singular vector $\boldsymbol{v}_1$ perfectly recovers the direct-path signal $\boldsymbol{x}_b(t)$, and the residual is zero: $\overline{\boldsymbol{X}} = \boldsymbol{X} -$

$\boldsymbol{u}_b \sigma_b \boldsymbol{v}_b^T = 0$. Our simulated result confirms our analysis, that the correlation coefficient between $\boldsymbol{v}_1$ and $\boldsymbol{x}_b(t)$ is 1.0, and that $\|\overline{\boldsymbol{X}}\|$ is 90 dB smaller than $\|\boldsymbol{X}\|$. We also observe that all singular values except the first one are close to 0, indicating that only the first left and right singular vectors are non-trivial, and that the other singular vectors are merely random noise.

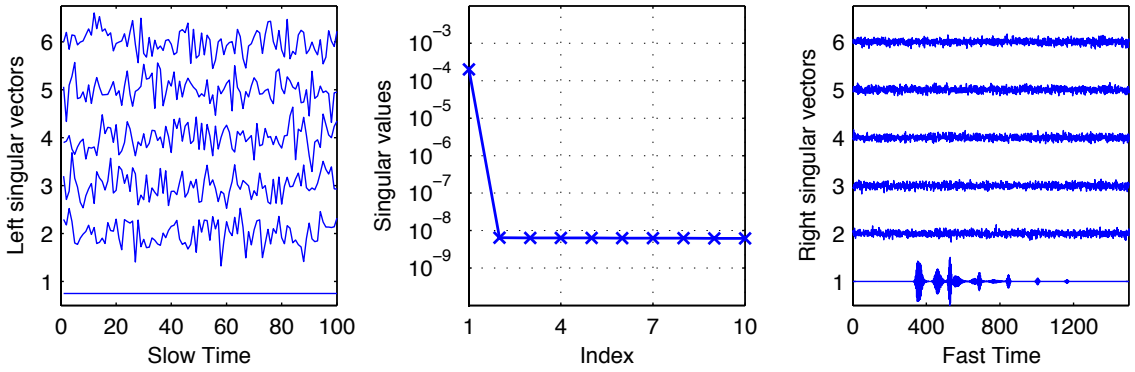


Figure 3-5 SVD of data matrix contains only direct-path records

For scenario 2, we construct a data matrix $\boldsymbol{X} = [\boldsymbol{x}_b^T, \dots, \boldsymbol{x}_b^T, \boldsymbol{x}_b^T + \boldsymbol{x}_s^T, \dots, \boldsymbol{x}_b^T + \boldsymbol{x}_s^T]$ with 80 replications of the direct-path signal $\boldsymbol{x}_b^T$, and 20 replications of the damage affected signal $\boldsymbol{x}_b^T + \boldsymbol{x}_s^T$. This simulates the scenario where during a monitoring phase, we collected 80 records when there is no damage on the structure, and the damage occurs before we collect the 81st record. We then calculate the SVD of $\boldsymbol{X}$ and plot the first 6 singular vectors in Figure 3-6.

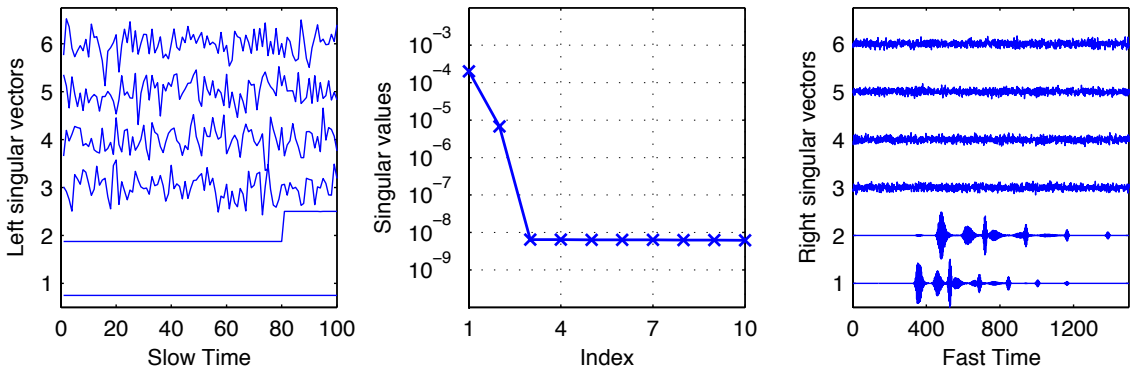


Figure 3-6 SVD of data matrix contains both direct-path and scatter-path signals and no other variations.

We can see from the middle plot in Figure 3-6 that the first two singular values are non-zero, indicating the first two pairs of singular vectors are meaningful. The first left and right singular vectors ($\boldsymbol{u}_1$, and $\boldsymbol{v}_1$) are the same as in Figure 3-5, indicating that they still resembles an optimal baseline record. However, in this case, the residual matrix $\overline{\boldsymbol{X}} = \boldsymbol{X} - \boldsymbol{u}_b\sigma_b\boldsymbol{v}_b^T$ is non-zero. In fact, as shown by Equation (3-16), the residual matrix contains 0 for the first 80 columns and the scatter-path signal for the last 20 columns. We show the simulated results in Figure 3-7. The three images from left to right show, respectively, the data matrix $\boldsymbol{X}$, the first singular vector term $\boldsymbol{u}_b\sigma_b\boldsymbol{v}_b^T$, and the residual matrix $\overline{\boldsymbol{X}}$.

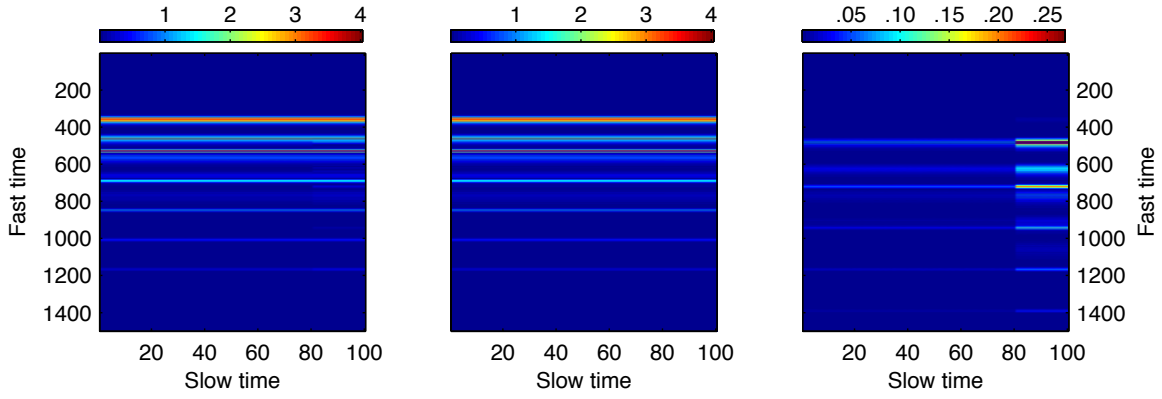


Figure 3-7 Images of: data matrix $\boldsymbol{X}$, first singular vector term $\boldsymbol{u}_b\sigma_b\boldsymbol{v}_b^T$, and residual $\overline{\boldsymbol{X}}$

The rightmost plot in Figure 3-7 shows the residual matrix $\overline{\boldsymbol{X}}$, which is the difference between the two images on the left – the records matrix $\boldsymbol{X}$ and the first singular vector terms $\boldsymbol{u}_b \sigma_b \boldsymbol{v}_b^T$ (note the difference in scales for the three images). The similarity between $\boldsymbol{X}$ and $\boldsymbol{u}_b \sigma_b \boldsymbol{v}_b^T$ indicates that the first singular vector terms captures most variations in the data – the baseline records. The residual $\overline{\boldsymbol{X}} = \boldsymbol{X} - \boldsymbol{u}_b \sigma_b \boldsymbol{v}_b^T$, as shown in the rightmost plot, has zeros in the first 80 columns, indicating that the baseline is mostly removed. The rest 20 columns show signals that are significantly different from the baseline records, and they are in fact the scatter-path signals. Therefore, when we maximizing Equation (3-17) with this residual matrix that contains only zeros and scatter-path signals, we obtain the second left and right singular vectors that only correspond to the scatter-path signal. Figure 3-8 shows a comparison of the actual scatter signal and the second right singular vector. For clarity, we plot their envelopes instead of the waveform, but note that the correlation coefficient between the two waveforms exceeds 0.99.

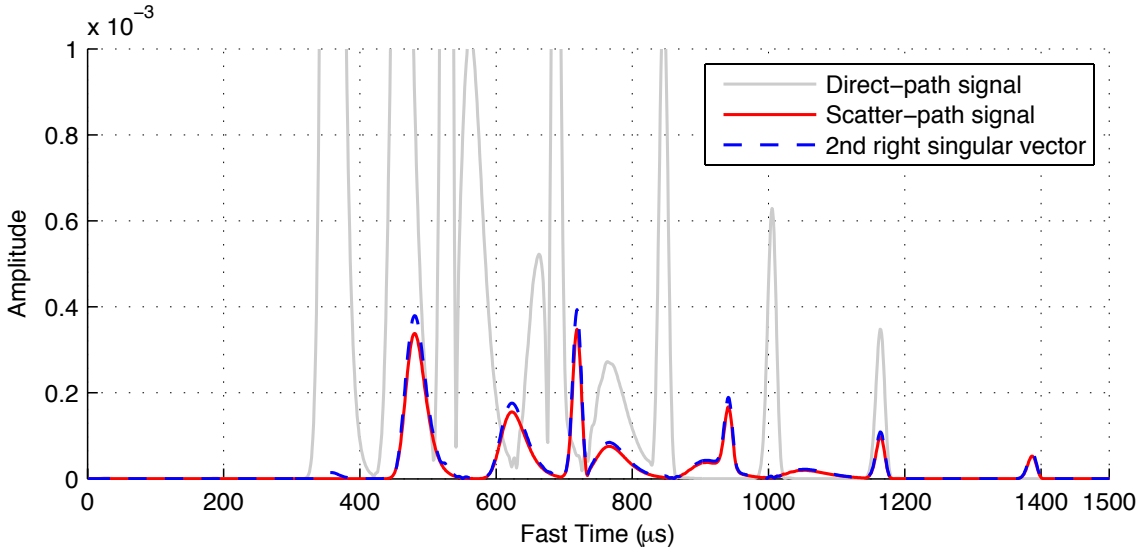


Figure 3-8 A comparison between scatter signal and extracted singular vector.

3.4.2 Simulation for scenario 3: various slow-time variations

In this subsection we simulate the scenario where the data matrix of N records $\mathbf{X} = [f_1(\boldsymbol{x}_b), \dots, f_n(\boldsymbol{x}_b), f_{n+1}(\boldsymbol{x}_b + \boldsymbol{x}_s), \dots, f_N(\boldsymbol{x}_b + \boldsymbol{x}_s)]$ consists not only the baseline and scatter signals as in the previous subsection, but the records are also affected by other variations. The time-dependent function $f_t(\cdot)$, for $t = 1,\dots,N$, represents the benign variations, such as uniform or non-uniform temperature change, white or color noise, or a slow-time decrease in receiving amplitude because of the degradation of transducer bonding. We will show multiple simulations where the records are affected by different slow time variations. In each simulation scenario, we will show the extracted singular vectors and compare them with the actual baseline and scatter signals to reveal their physical meaning. We will show that the SVD method not only separates the baseline from the scatter signals, but also captures the effects of the variations on them, as we predicted in *Section 3.3.3* by equation (3-19).

We first simulate the baseline and scatter signals as we did in the previous section, which is shown in Figure 3-4, and then design the time-dependent function $f_t(\cdot)$, which abstracts some slow-time variations. For example, if we were to simulate a temperature variation over N records, we define the function as $f_t(\cdot) = \boldsymbol{T}_\alpha(\cdot)$, where $\boldsymbol{T}_\alpha(\boldsymbol{x}(t)) = \boldsymbol{x}(\alpha t)$ uses a time-domain stretch to simulate the primary effect of a temperature change. We specify the stretching factors $\boldsymbol{\alpha}^{1\times N}$ for each of the N records to simulate temperature varying over the slow time. We then apply $f_t(\cdot)$ to each of the simulated record to synthesize the variations, obtaining $\mathbf{X} = [f_1(\boldsymbol{x}_b), \dots, f_n(\boldsymbol{x}_b), f_{n+1}(\boldsymbol{x}_b + \boldsymbol{x}_s), \dots, f_N(\boldsymbol{x}_b + \boldsymbol{x}_s)]$.

3.4.2.1 **Case 1: Periodic, two-stage temperature variation**

In the first case, we synthesize records with a two-stage, periodic temperature variation, where $\boldsymbol{\alpha}^{1\times N}$ contains only two distinct values. We synthesize five records at temperature T_0

and then five records at T_1, and repeat the process until we obtain 100 records. We introduce the scatterer at the 81st record, as we did in Scenario 2 in the previous subsection. Figure 3-9 shows three images to demonstrate the synthesized process. Each image is normalized such that each pixel ranges from 0 (blue) to 1 (red). The horizontal axis is slow time, where the values 1 to 100 indicate 100 synthesized records. The vertical axis is fast time, where the values 1 to 1500 indicate 1500 samples in each record. The two plots on the left show the simulated baseline signals $[\boldsymbol{x}_b^T, \boldsymbol{x}_b^T, \dots, \boldsymbol{x}_b^T]$ and scatter signals $[0, \dots, 0, \boldsymbol{x}_s^T, \dots, \boldsymbol{x}_s^T]$ (the scatter signals only exist in the last 20 records). The plot on the right shows the 100 records synthesized with two-stage temperature variations. We can clearly identify the change produced by temperature because it stretches the records and creates ripples. We can barely see any indication of the scatter signal, whose energy is 40 dB smaller than that of the baselines.

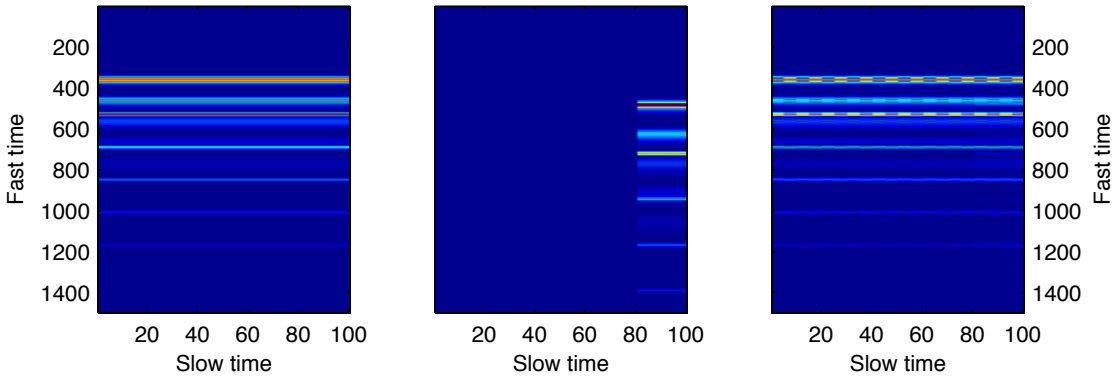


Figure 3-9 Normalized images of direct-path and scatter-path signals used in the synthesis. Left: 100 direct-path records; Middle: 20 scatter records; Right: 100 synthesized records that experience two-stage periodic temperature variations

Figure 3-10 shows the correlation coefficients of the 100 records to further illustrate the synthesized temperature variation. As defined in equation (2-1), correlation coefficients

characterize the similarity between two waveforms. If two waveforms have the identical shape, the correlation coefficient is 1.0. The less correlated the two waveforms are, the closer the coefficient is to 0. In Figure 3-10, each point on the blue curve shows how similar that record is to the baseline record, which is the first record at room temperature T_0. Each point of the green curve shows the similarity between the record and the scatter signal. We see that the blue curve shows no indication of a change at the 81st measurement, but shows only variations caused by the periodic temperature toggling. On the other hand, we observe from the green curve that there is relatively little correlation between the scatter signal and the first 80 baseline records, which confirms the assumption that the baseline and scatter signals are approximately orthogonal to each other.

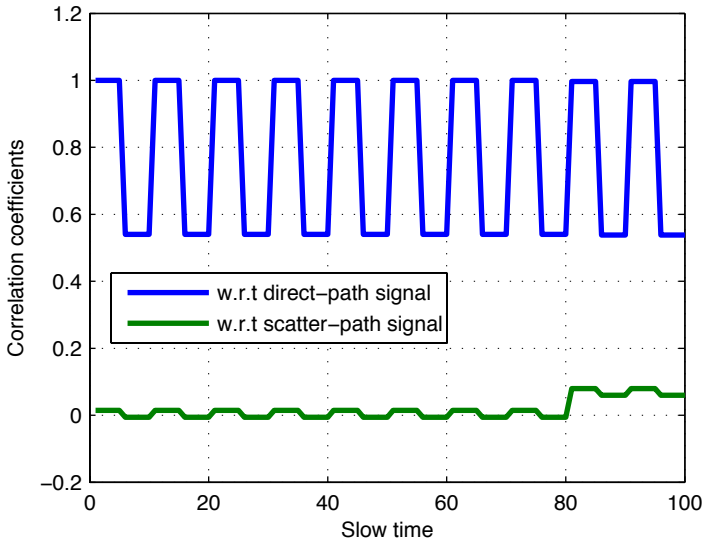


Figure 3-10 Correlation coefficients of simulated records synthetized with two-stage temperature variation

We calculate the SVD of the data matrix containing these 100 records, and show the result in Figure 3-11. We can make several interesting observations here. First, the first four singular values are non-zero. Intuitively, we can see each non-trivial term as some kind of

variation in the dataset. Therefore, the four non-zero singular vector indicates that the dataset is decomposed into four kinds of variations.

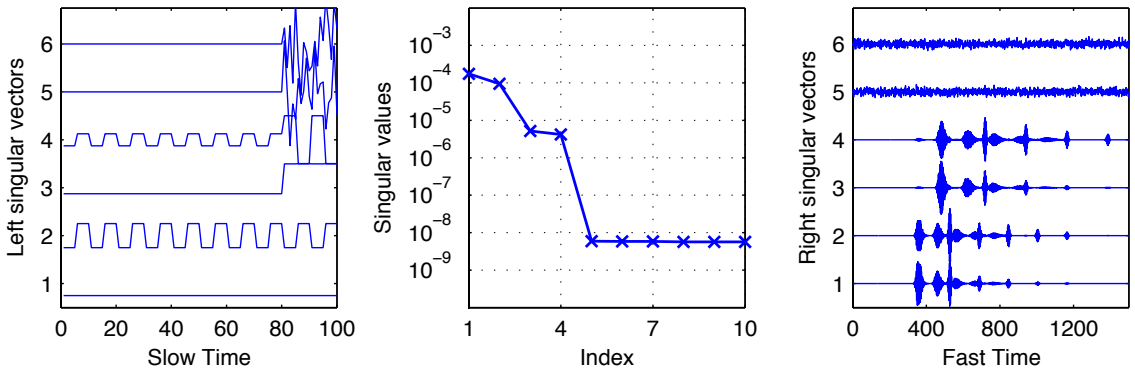


Figure 3-11 SVD result of simulated records synthetized with two-stage temperature change

We then examine the first four singular vector pairs one-by-one.

- The 1st left and right singular vectors, as usual, are associated with the baseline record. This can be confirmed by the fact that the 1st left singular vector is a straight line, representing a steady trend exists across the *N* records. We can also confirm this relationship by calculating the correlation coefficients between the 1st right singular vector and the actually baseline record. The coefficient is 0.99, very close to the value of 1.0 that would indicate the two waveforms have identical shape.

- The 2nd right singular vector has all the arrivals exhibited in the 1st singular vector and the baseline record. The only difference is that the 2nd singular vector shows later arrivals with relatively higher amplitude, indicating that it captures the change produced by the temperature change, which becomes more significant as the wave propagates further. The 2nd left singular vector further confirms this – it shows periodic changes corresponding to the temperature change.

- The 3rd left singular vector shows a clear step at the 81st record, indicating that this term captures the change of interest – the scatter signal. The 3rd right singular vector confirms this observation, as the correlation coefficients between the 3rd right singular vector and the actual scatter signal is 0.98.
- The 4th right singular vector looks visually similar to the 3rd one, but again shows the later arrivals to have relatively higher amplitude. This indicates that this singular vector represents the temperature effect on the scatter signals. The 4th left singular vector shows small variations in its first 80 values and large periodic variations in the last 20 values.

The analysis above shows that the four meaningful singular vector terms represent, respectively, the baseline signal $\boldsymbol{x}_b$, the temperature change on baseline signal ($\boldsymbol{T}\{\boldsymbol{x}_b\} - \boldsymbol{x}_b$), the scatter signal $\boldsymbol{x}_s$, and the temperature change on scatter signal ($\boldsymbol{T}\{\boldsymbol{x}_s\} - \boldsymbol{x}_s$). The reason that we have four components is because we have two signal components $\boldsymbol{x}_b$ and $\boldsymbol{x}_s$, collected at two different temperature T_0 and T_1. The combination of the two types of variations creates four distinguishable effects. Because $\boldsymbol{x}_b$ and $\boldsymbol{x}_s$ are approximately orthogonal, they can be well separated into two singular vectors, i.e. the 1st and the 3rd. However, the temperature effect does not make the signals uncorrelated, in that $\boldsymbol{T}_{T_0}\{\boldsymbol{x}_b\}$ is not orthogonal to $\boldsymbol{T}_{T_1}\{\boldsymbol{x}_b\}$. As a consequence, the 2nd and the 4th singular vector resemble the difference signal produced by the temperature change. We summarized the general form of the singular vectors on a data matrix in Equation (3-19). In this particular case with two-stage temperature change, Equation (3-19) becomes,

$$\boldsymbol{X} = \boldsymbol{u}_b\sigma_b\boldsymbol{v}_b^T + \boldsymbol{u}_{bT}\sigma_{bT}\boldsymbol{v}_{bT}^T + \boldsymbol{u}_s\sigma_s\boldsymbol{v}_s^T + \boldsymbol{u}_{sT}\sigma_{sT}\boldsymbol{v}_{sT}^T + O(\sigma). \tag{3-20}$$

3.4.2.2 **Case 2: Periodic sinusoidal temperature variation**

In the previous, simplified case, where temperature can take only two discrete values, SVD represents the changes produced by temperature with one additional singular vector term for the baseline and another one for the scatter signal. However, in reality, temperature is a continuous function of time such that each record would be affected by different temperature. In that case, more singular vector terms will be needed to represent this variation. We illustrate that with another simulation, in which temperature values are sampled from a sinusoidal function. The sample interval is not a fraction of π, and therefore the temperature readings used in the synthesis are all different from each other. Figure 3-12 shows the correlation coefficients of the 100 records in this synthesized simulation.

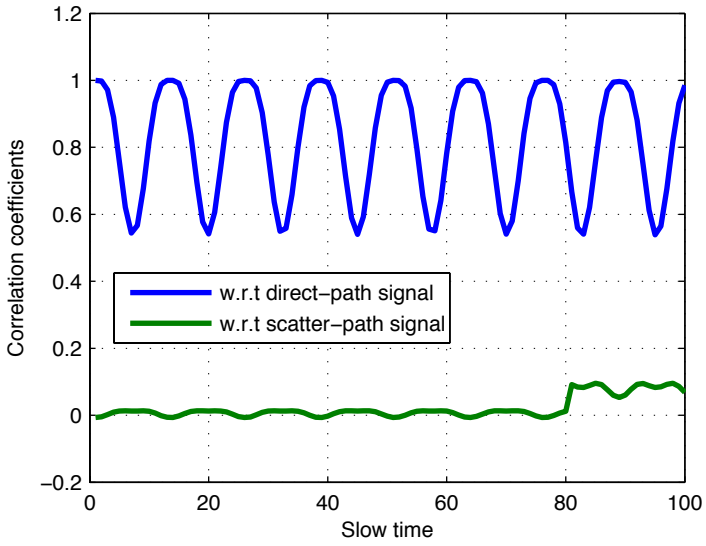


Figure 3-12 Correlation coefficients of simulated records synthetized with sinusoidal temperature change

We show the SVD result in Figure 3-14. In this case, we have more non-zero singular values that indicate the data matrix is experiencing more temperature variations. We further look at the singular vectors to examine what they represent.

By looking at the right singular vectors, we can identify that four of them (1st, 2nd, 3rd, and the 6th) resembles the direct-path signals. We can also recognize a clearly trend that from the first to the fourth right singular vector, the amplitude of later arrivals increases, and the amplitude of the earlier arrival decreases. Unlike the previous case where temperature can only take two discrete values, in this case temperature creates difference in signals that cannot be simply represented by a single difference signal. Therefore, SVD produced more singular vectors with increasingly amplified later arrivals to characterize the temperature variations. We can analogize this phenomenon with Fourier transform, in which sinusoidal bases with different frequencies are used to refine the representation of the original signal, and eventually the summation of the bases recovers the original signal. In our case, the singular vector terms $\boldsymbol{u}_i\sigma_i\boldsymbol{v}_i$ gradually refines the change produced by temperature by increasing the amplitude of later arrivals of the signals, where temperature changes significantly. Eventually, the sum $\sum_R \boldsymbol{u}_i\sigma_i\boldsymbol{v}_i$ approximates all the changes produced by temperature.

In the slow-time domain, the corresponding 2nd, 3rd, and the 6th left singular vectors resemble a set of harmonic bases that characterize the sinusoidal temperature variations. From left to right, Figure 3-13 shows the following: the data matrix $\boldsymbol{X}$, the first singular vector term $\boldsymbol{u}_1\sigma_1\boldsymbol{v}_1$, the summation of the 1st, 2nd, 3rd, 6th singular vector terms $\Sigma_{i=1,2,3,6}\boldsymbol{u}_i\sigma_i\boldsymbol{v}_i$. We can see that the original data matrix has ripples representing the sinusoidal temperature change. The first singular vector term captures the averaged baseline, exhibiting little variation. The 2nd, 3rd and 6th singular vector terms gradually add these variations such that the variations in the data matrix are fully explained. Therefore, the third image that shows $\sum_{i=1,2,3,6}\boldsymbol{u}_i\sigma_i\boldsymbol{v}_i$ is almost identical to the original data matrix. The

correlation coefficient between the two matrices exceeds 0.98. The differences between the two matrices are the scatter signals in the last 20 records, which are characterized in the 4th and 5th singular vector terms.

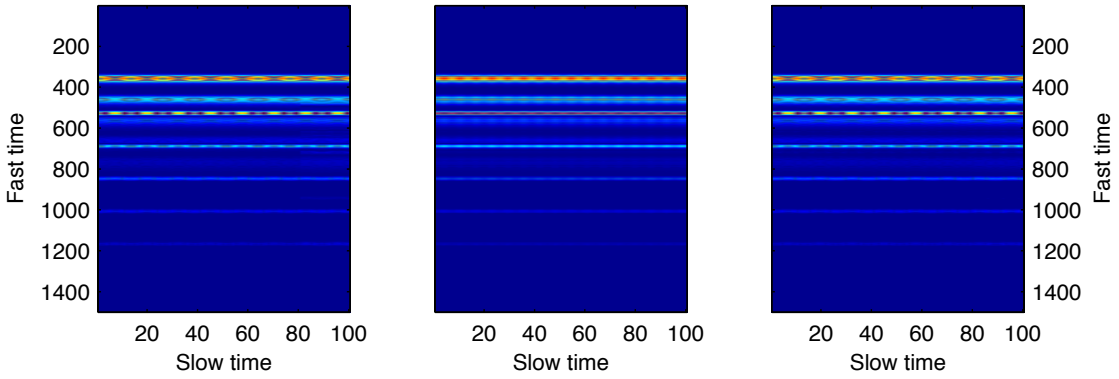


Figure 3-13 Images from left to right: data matrix $\boldsymbol{X}$, the first singular vector term $\boldsymbol{u}_1\sigma_1\boldsymbol{v}_1$, summation of the 1st, 2nd, 3rd, and the 6th singular vector terms $\sum_{i=1,2,3,6}\boldsymbol{u}_i\sigma_i\boldsymbol{v}_i$

The 4th and 5th singular vectors are associated with the scatterer. The 4th singular vector optimally represents the averaged scatter signal, and the 5th singular vector shows later arrivals with larger amplitude to characterize the temperature effect. If we plot more singular vectors, we will see more singular vectors associated with the temperature change on the baseline and scatter signals. There would be as many singular vectors as possible to characterize the variation until the residual of $\overline{\boldsymbol{X}} = \boldsymbol{X} - \sum_R \boldsymbol{u}_i\sigma_i\boldsymbol{v}_i$ is dominated by noise or round-off error.

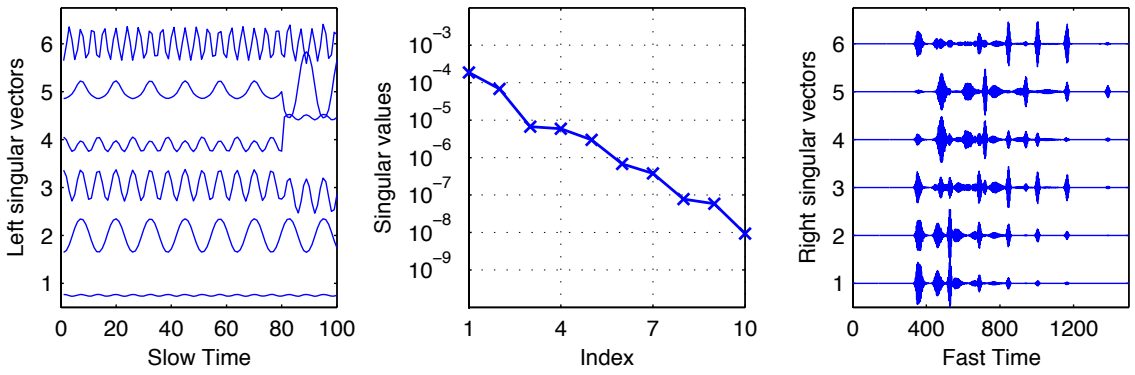


Figure 3-14 SVD result of simulated records synthetized with sinusoidal temperature change

In the next case, we added another type of variation on top of the sinusoidal temperature variation shown above. We decrease the amplitude of the received signal to simulate the degradation of the transducer bonding condition, which causes a drop in received signal strength.

3.4.2.3 **Case 3: Periodic sinusoidal temperature variation combined with decreasing received signal strength**

In the practical implementation of a long-term monitoring system, it is always important to ensure the stability of the transducer performance. Nonetheless, the received signal strength may decrease over time due to degradation and aging. In this case, it is more difficult to identify the scatter signal because the difference signal will include the long-term amplitude change of the direct-path signals, which could become dominant. Figure 3-15 shows the direct-path and scatter-path signals, as well as the synthesized 100 records that experience both sinusoidal temperature variations and a decreasing received signal strength. Because of the decreasing amplitude of the received signals, there is almost no indication of the existence of the scatter signals in the last 20 records.

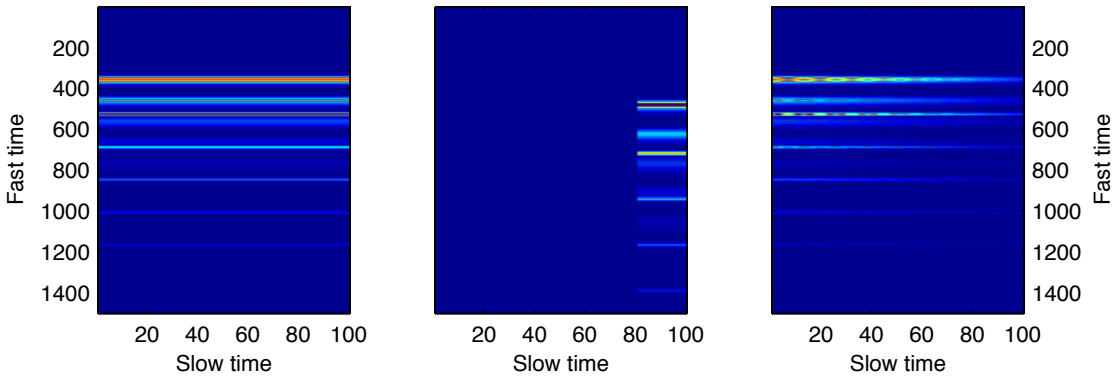


Figure 3-15 Normalized images of direct-path and scatter-path signals used in the synthesis. Left: 100 direct-path records; Middle: 20 scatter records; Right: 100 synthesized records that experience both sinusoidal temperature variations and a decreasing receiving signal strength

We then use SVD on the data matrix $\boldsymbol{X}$, as shown in the right plot of Figure 3-15, and then show the results in Figure 3-16. The result is very similar to the result of the previous case, shown in Figure 3-14. In fact, the right singular vectors are almost identical to those shown in that previous case. The interesting observation here is that now the left singular vectors all show some decreasing trends associated with the decreasing amplitude of the received signals. The 1st left singular vector monotonically goes up, indicating that the amplitude of baseline signals decreases. Similarly, the decreasing trend shown in the last 20 values of the 4th left singular vector indicates that the scatter signal is also affected by the amplitude reduction over slow time. The other singular vectors, exhibiting sinusoidal change associated with the temperature variations, are also affected by the decreasing amplitude of the received signals.

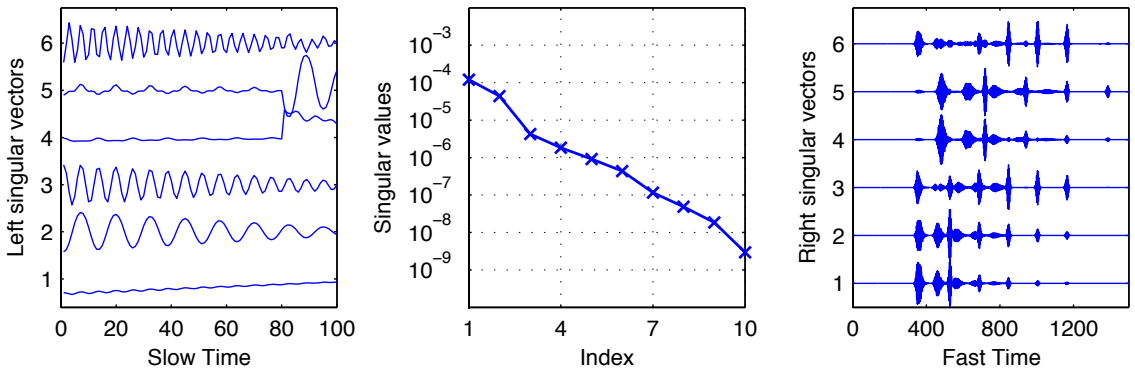


Figure 3-16 SVD result of simulated records synthetized with sinusoidal temperature change and a decreasing receiving energy

The reason that the temperature effects are exhibited in different singular vector, but that decreasing amplitude is largely affecting the left singular vector, is rooted in the orthogonality characteristic of singular vectors. The temperature variation creates non-linear changes, such that each record affected by a different temperature cannot be reconstructed by linear combinations of the existing singular vectors. However, the decreasing received signal strength only reduces the amplitude of the signals. Therefore, the change between two records is not orthogonal to the existing right singular vectors. Instead, the change can be represented by a simple point-wise multiplication of the extracted right singular vectors, which is equivalent to a change in the left singular vectors. As a consequence, the right singular vectors remain largely unchanged when compared to the previous case, because they represent the orthogonal components produced by the temperature variations; however, the left singular vector captures that change caused by the decreasing received signal strength. The result can be generalized if the changes in received signal strength are frequency dependent, as long as that change does not affect our assumption of orthogonality.

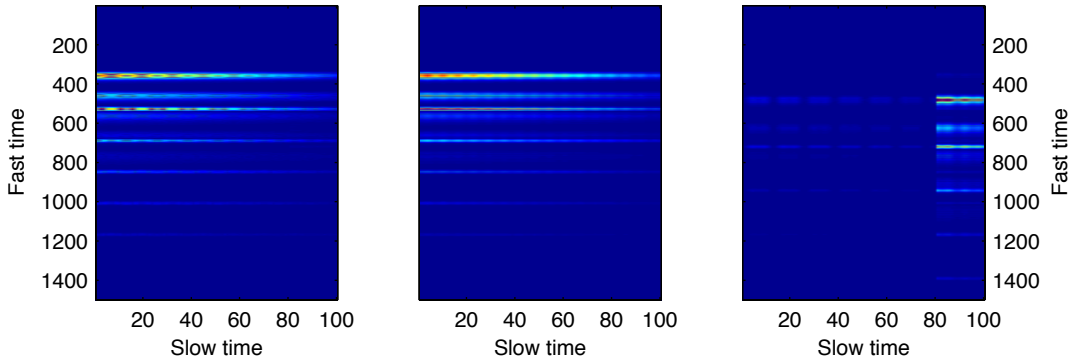


Figure 3-17 Normalized images. Left: data matrix $\boldsymbol{X}$; Middle: the singular vector term associate with baselines $\boldsymbol{u}_1\sigma_1\boldsymbol{v}_1^T$. Right: the singular vector term associate with scatterer $\boldsymbol{u}_4\sigma_4\boldsymbol{v}_4^T$

3.5 Summary

In this chapter, we demonstrated the damage detection framework we developed based on singular value decomposition. With the assumption that the signal produced by the damage – the scatter signal – is approximately orthogonal to the baseline records, singular vector decomposition is an effective tool to extract the scatter signal from a batch of ultrasonic monitoring records.

We illustrated the physical meaning of the result of singular value decomposition, that the left singular vectors exhibit the trend of slow-time variations in the records, and that the right singular vectors resemble the change in the waveform that corresponds to the slow-time variations. We conducted synthetic simulations in different scenarios where the ultrasonic records are experiencing different variations, and show successful separation of the damage information using our singular value decomposition method.

In the next chapter, we will apply our SVD framework on the experimental data we collected from the field testbed we described in *Chapter 2*. We will show successful detection of a small scatterer despite the large environmental and operational variations that

prevent us from getting any indication of damage from the ultrasonic monitoring records even with the help with state-of-art compensation techniques.

I will finish this chapter with a quote from the well-known book *Introduction to Linear Algebra* by Strang (Strang 2009):

"I give you my opinion directly. The SVD is the climax of this linear algebra course. I think of it as the final step in the Fundamental Theorem. First come the dimensions of the four subspaces. Then their orthogonality. Then the orthonormal bases, which diagonalize $\boldsymbol{A}$. It is all in the formula $\boldsymbol{A} = \boldsymbol{U\Sigma V}^T$. More applications are coming – they are certainly important-but you have made it to the top."

Chapter 4 Robust Damage Detection Using SVD

4.1 Motivation

Guided wave ultrasonics is an attractive monitoring technique for damage diagnosis in large-scale plate and pipe structures. Guided waves propagate with low attenuation and can interrogate large areas with only a small number of sparsely distributed, low-voltage transducers (Cawley et al. 2003). Ultrasonic guided waves are characterized by a dispersive and multi-modal nature, which complicates the received ultrasonic signals and makes it challenging to extract information about the damage (Graff 1975).

Due to the complexity of guided waves, many damage detection methods rely on baseline comparison to remove static, background information. In these scenarios, a set of baseline records is collected when the structure is known to be intact. The differences between any new records and the baseline signal can then be used to monitor for structural damage (Michaels and Michaels 2005). Alternatively, structural damage can be detected by analyzing

the cross-correlation coefficients, a measure of similarity, between a baseline and each new time-record (Gao et al. 2005).

However, ultrasonic waves are vulnerable to changes in environmental and operational conditions (EOC) (SOHN 2007) that are inevitable in the normal operation of civil and mechanical structures. Therefore, the baseline information is generally not static over time. In active pipes, for example, ultrasonic waves are often influenced by variations in temperature, pressure, and flow rate. These effects complicate analysis and mask damage-related information (Liu et al. 2012).

Among the common EOCs to affect ultrasonic monitoring systems, temperature is the most ubiquitous and widely studied. Researchers have developed many algorithms to compensate for temperature variations. Optimal baseline selection methods were first developed in (Konstantinidis et al. 2006), which overcome the effects of temperature variation by comparing the new record with a library of baselines collected at different temperatures. Researchers then developed local peak coherence (Michaels and Michaels 2005; Lu and Michaels 2009), and optimal signal stretch methods (Clarke et al. 2010; Croxford et al. 2010), assuming temperature change has a stretching effect on the signal. With that assumption, the amount of temperature change can be determined by comparing new records with stretched version of a baseline signal. One can then stretch the new record so that it is comparable to the baseline records. The estimation of stretching factors can also be done in the scale transform domain to achieve higher resolution and efficiency, as shown in (Harley and Moura 2012). Other researchers argue that having a fixed baseline or baseline set may be not sufficient in a dynamic environment, and developed the continuously growing baseline temperature compensation method to avoid collecting a comprehensive library of

baselines before the monitoring phase (Putkis and Croxford 2013b). As demonstrated in both laboratory and practical experiments, these temperature compensation methods can be used to adjust each record and improve our capability to detect damage.

However, stretch-based temperature compensation methods have certain limitations. First, these temperature compensation methods model temperature changes as a stretching effect on the ultrasonic signals, which is only an approximation. This model does not hold for large changes in temperature. Second, the methods assume that temperature variation is uniform across the structure. Many structures are instead affected by temperature gradients. Last, temperature compensation methods generally require a set of baseline records to be collected either before or during the monitoring, which can be difficult to manage and/or update in a dynamic environment. Moreover, in a practical implementation of ultrasonic monitoring systems, temperature variation is often accompanied by other EOCs that affect the ultrasonic record. Our field experiments show that EOCs contribute to many variations of the ultrasonic records, and that temperature compensation only addresses a portion of them.

Because analytically modeling the effect of EOCs on the ultrasonic is challenging, researchers have developed various data-driven methods that extract useful information from large datasets of ultrasonic records (Lu and Michaels 2009; Ying et al. 2012; Haynes et al. 2013). Data-driven methods extract useful features from data and then use those features to classify the status of the structure. A data-driven damage detection procedure used (explicitly or implicitly) by many researchers includes pre-processing, feature extraction, damage-sensitive feature selection, and damage classification. Detection or classification is accomplished with well-developed methods in the literature such as support vector machine (Cortes and Vapnik 1995), neural network(Bishop 1994), and Fisher's discriminative

analysis (Scholkopft and Mullert 1999). However, reliable damage-sensitive features are usually application-specific and are difficult to find.

In *Chapter 3*, we have developed a novel damage-sensitive feature extraction and selection procedure based on singular value decomposition (SVD) to detect structural damage with ultrasonic pitch-catch records. We demonstrated with synthesized data records that by applying SVD on ultrasonic monitoring records, we can separate the change produced by damage from the change caused by EOCs, and thereby robustly detect damage in a complex environment.

In this chapter, we further test our SVD framework described in *Chapter 3* with experimental data – the ultrasonic pitch-catch monitoring records collected on pipes. We first, in *Section 4.2*, apply our SVD framework on the field experimental data that is described in *Chapter 2*. We show that despite the challenges from large environmental and operational variations, our SVD framework successfully detects a scatterer on the pipe, which we used as a physical simulation of damage. We also show that our method is consistently robust and accurate on experimental monitoring records collected spanning over seven months. We extend the framework to detect progressive damage. In *Section 4.3*, we examine our method on experimental data collected on a pipe that experienced pitting corrosion, temperature variation with time, and spatial temperature gradients. We found that we can clearly identify the trend of the gradual degradation associated with corrosion, even though the temperature changes distorted the record significantly.

4.2 Detecting A Scatterer On A Pipe

In this section, we show successful scatter detection results on experimental data collected in the field testbed that is described in *Chapter 2*. In *Section 4.2*, we show the detection result of our SVD framework on one example dataset that has been analyzed in detail in *Section 2.3*. We show the singular vectors decomposed from the ultrasonic records, and the scores that indicate the damage-sensitive singular vector using the K-means method. We also discuss the effect of temperature compensation as a preprocessing step on SVD results. Then in *Section 4.3*, we summarize the detection result of our framework on field experimental datasets collected over a 7-month period to demonstrate the accuracy and robustness of our method.

4.2.1 Results on one single dataset

We applied singular value decomposition as described in *Chapter 3* to ultrasonic records collected in our field experiments that was described in *Chapter 2*. We use 28 hours of records collected on August 31st, 2011. The first 10 hours of the data has been characterized in Figure 2-7. The dataset consists of 1,481 pitch-catch records collected over 28 hours from 10:22, August 31st to 14:16, September 1st, with a 3.2-hour pause starting from 20:01, August 31st. (The pause was due to a suspension in the data acquisition system, and affected neither the operation of the piping system nor the configuration of the data acquisition system for later data collection.) In Figure 4-1, we illustrate the first 15 left and right singular vectors decomposed from the records. We observe that the 11th singular vector shows steps that are consistent with the presence of a scatterer. The step-wise dark lines indicate the placement and removal of the scatterer, with high and low values representing, respectively, the presence and absence of the mass scatterer. The numbers on the right of the left singular

vectors are the scores $\boldsymbol{\eta}$ computed with the K-means method using equation (3-6). We see that all of the scores except one are smaller than 0.50. The score of the damage-sensitive singular vector is 2.20, which is not only the largest of scores, but also much larger than the threshold we set ($\boldsymbol{\eta}_i > 0.50$) to determine the presence of damage in this batch of records. The right singular vectors have the same length (10,000 points) as the ultrasonic records and show structures with wave packets. As has been shown in Figure 3-1, the right singular vector corresponding to the damage-sensitive left singular vector is a replication of the scatter signal. We have previously shown that we can use the extracted right singular vectors as baseline-removed signals to localize damage under complex environment (Liu et al. 2014).

We also computed the singular vectors after performing temperature compensation, and show the results in Figure 4-2. The damage-sensitive singular vector remains the same shape, showing good consistency, except that the index of the damage-sensitive singular vector changes from 11 to 10. Singular vectors are sorted by singular values, which weight the variations represented by the singular vectors. This indicates that temperature compensation successfully eliminated some variation. However, if temperature compensation could perfectly remove the baseline signal, the damage would become the only variation in the records, and we might expect to see the steps corresponding to the mass in the 1st singular vector.

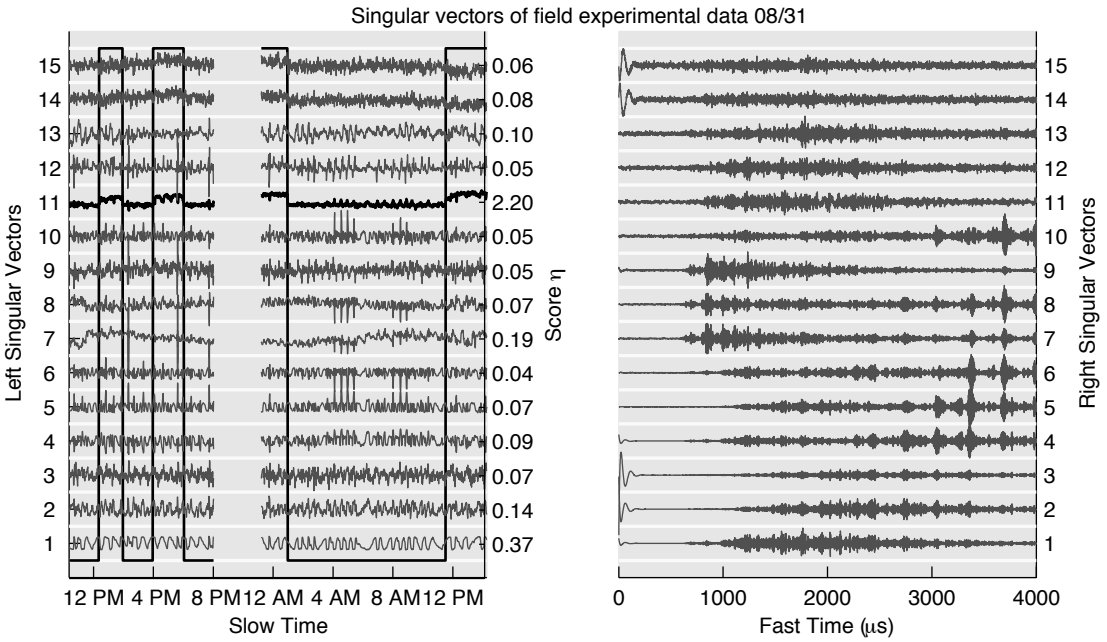


Figure 4-1 Left and right singular vectors decomposed from field experimental data collected on 08/31/2011. Values besides the left singular vectors indicate the damage-sensitive score we compute using Equation 6. The 11th singular vector is corresponding to the mass toggling steps.

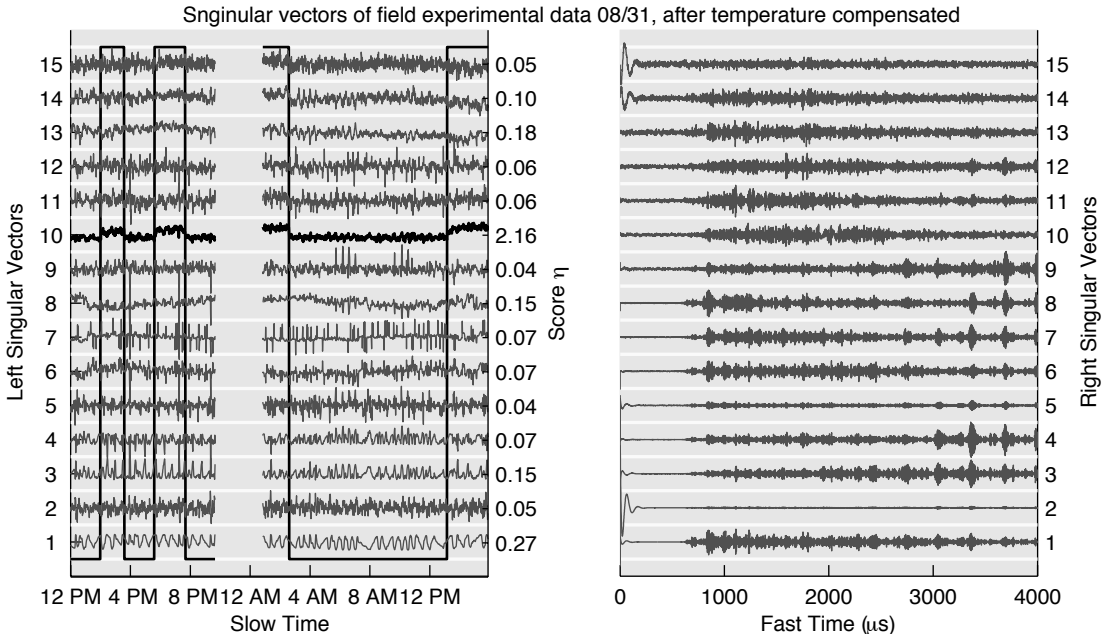


Figure 4-2 Left and right Singular vectors decomposed from temperature compensated data collected on 08/31/2011. The damage-sensitive singular vector is the 10th.

4.2.2 Summarized results on multiple datasets over 7 months

We illustrate the performance of damage detection using our SVD damage-sensitive feature across data collected during a period spanning seven months, as described in *Chapter 2*, by plotting the receiving-operating characteristic (ROC) curves of classification using the SVD damage-sensitive feature. The ROC curve is widely used to evaluate a classifier system (Bradley 1997). It shows the relation between the true-positive rate (TPR) and the false-positive rate (FPR) as the discrimination parameter varies. In our case, we used a one-dimensional threshold classifier on our damage-sensitive singular vector selected by the K-means method. If $\boldsymbol{u}_k^{1\times N}$ denotes the selected singular vector – our damage-sensitive feature – then the threshold classifier outputs a vector $\boldsymbol{\gamma}^{1\times N}$ that indicates presence of damage at each time step.

$$\gamma_i = \begin{cases} 1 & u_{ki} \geq \lambda \\ 0 & otherwise \end{cases} \tag{4-1}$$

We varied the discrimination parameter, λ, from the minimum to the maximum of the singular vector, and compared the output $\boldsymbol{\gamma}$ with the ground truth (the time at which we applied or removed the mass scatterer during experiments) to compute TPR and FPR as defined below

$$TPR = \frac{True\ Positive}{True\ Positive + False\ Negative},$$
$$FPR = \frac{False\ Positive}{False\ Positive + True\ Negative}. \tag{4-2}$$

The ROC curve is then obtained by plotting TPR against FPR, both ranging between [0,1]. The best possible prediction method would yield a point in the upper left corner or coordinate (0,1) of the ROC space, representing 100% sensitivity (no false negatives) and 100% specificity (no false positives). The (0,1) point is also called a perfect classification. A

completely random guess would give a point along a diagonal line (the so-called line of no-discrimination) from the left bottom to the top right corners. In our case, we use a simple classifier to demonstrate the performance of two aspects of our proposed methods: (1) the performance of the selected singular vector as a damage-sensitive feature, and (2) the performance of the K-means selection process, using a fixed threshold $\eta > 0.50$.

Figure 4-3 shows the ROC curves on seven field experimental datasets collected during the seven-month period. The black dashed diagonal line indicates the line of no-discrimination. For each of the seven datasets, we plotted one solid curve and one dashed curve in the same shade. The solid ROC curve was generated from the damage-sensitive singular vector selected by the K-means selection process. The dashed ROC curve was generated with the same threshold classifier on the correlation coefficients computed with the same datasets. We illustrate that correlation coefficients are not a good damage indicator under EOC variations; therefore, as expected, the ROC curve of the correlation coefficients all appear close to the line of no-discrimination.

The area under the ROC curve (AUC) can be used to interpret the ROC curve (Bradley 1997). In general, better classification systems have large AUC; a random guess will have $\mathrm{AUC} = 0.5$; and a perfect classification system will have $\mathrm{AUC} = 1.0$. We summarize in Table 2 the AUCs for the two sets of curves generated from the seven datasets. The first column shows the AUC of our damage-sensitive singular vectors, which are all above 0.90. It shows that the damage-sensitive singular vectors are good features to classify the presence of damage in all seven datasets. The result is also consistent with our previous classification results reported in (Liu et al. 2013a), which applied linear discrimination analysis (LDA) classification on randomly grouped trained and tested data from the datasets. The second

column shows the AUCs of using the correlation coefficients as a feature. The correlation coefficients indicate the similarity between each measured record with the first record of the day. The values in the second column are all close to 0.5, indicating that the correlation coefficients are not a good feature to determine the presence of damage. The third column shows the difference between AUC_1 and AUC_2, showing the superior performance of our SVD damage-sensitive feature compared to correlation coefficients.

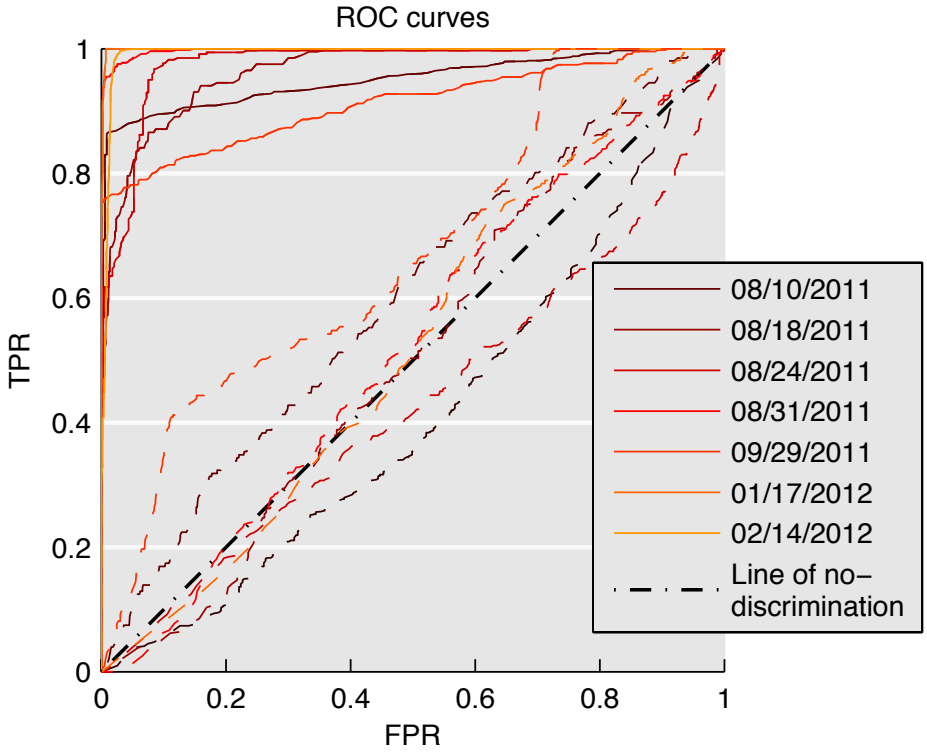


Figure 4-3 ROC curves using selected singular vectors from multiple datasets (shown in different color/shade), compared with ROC curves using correlation coefficients (in dashed lines), along with the line of no-discrimination in dash-dot line.

Table 2 Area under ROC curve for damage-sensitive singular vector and correlation coefficients.

Date	AUC_1	AUC_2	ΔAUC
08/10/2011	0.95	0.41	0.54
08/18/2011	0.96	0.60	0.37

08/24/2011	0.98	0.51	0.47
08/31/2011	1.00	0.43	0.57
09/29/2011	0.91	0.52	0.39
01/17/2012	1.00	0.67	0.33
02/14/2012	0.99	0.52	0.47

*AUC_1 denotes the area under ROC curve generated with damage-sensitive singular vector; AUC_2 denotes the area under ROC curve generated with correlation coefficients;

$\Delta AUC = AUC_1 - AUC_2$.

We also found that our SVD method is robust to different operating conditions of the piping system. The seven datasets displayed in Figure 4-3 spanned 7 months from August 2011 to February 2012, during which time period the EOC vary considerably. Figure 4-4 shows the 1st and the 12th singular vectors of two datasets, collected under two very different operating conditions of the piping system. Note that, by definition, the first singular vector always corresponds to the largest variation in the system, which is dominated by the variation of temperature in our experiments. Comparing the first singular vector in the two plots, we observe that the temperature varies quickly and sharply in the 08/24 dataset, whereas it varies slowly and mildly in the 02/14 dataset. The reason is that in the August dataset the heating system worked periodically because the demand for domestic hot water was low, and therefore for much of the time the boiler was idle. However, in the February dataset there is a much higher demand on the system for space heating, and therefore the heating system approximated continuous operation. Therefore, the large periodic characteristic disappears in February, leaving only mild variations over days that largely depend on the daily weather. Nonetheless, in both cases the 12th singular vector clearly indicated the presence of the mass scatterer.

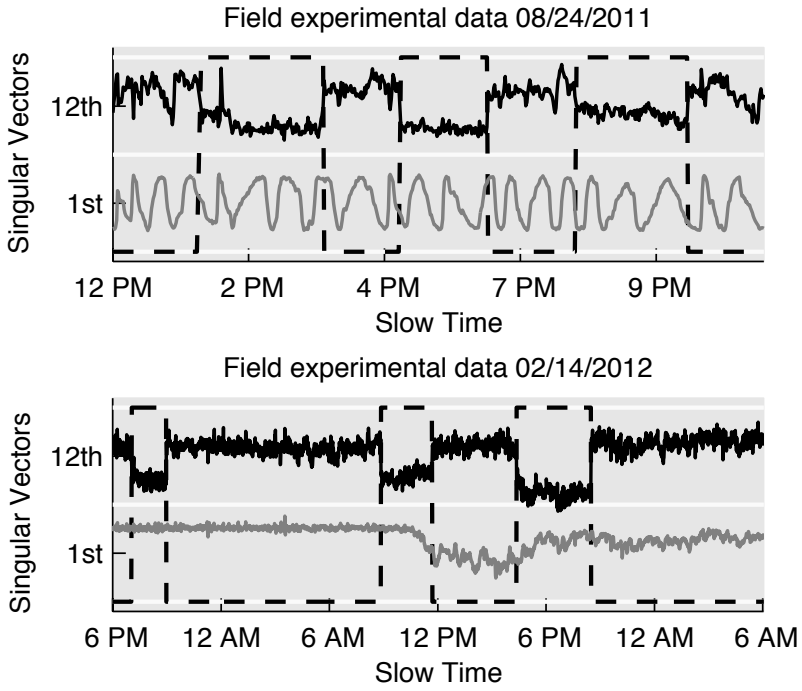


Figure 4-4 Damage-sensitive SV (12) of two datasets, one collected during summer and the other collected during winter, correspond well with the ground truth of mass toggling (black dashed line), along with the 1st SV corresponding to temperature variation.

4.3 Identifying Progressive Degradation In A Pipe

4.3.1 Problem statement

Pipes and pipelines are critical components in civil infrastructure systems transporting water, natural gas, or gasoline. Among the many mechanisms that may damage the structural integrity of pipe-like structures, corrosion is one of the most common and dangerous (Fessler 2008). However, corrosion could be very difficult to detect in pipe structures because of its hidden and inaccessible nature. It often occurs on the interior of pipe walls or on pipes that are buried under the ground.

Researchers have developed many ultrasonic techniques to detect pipe corrosion, taking advantage of their capability to propagate through media or long distance to inspect otherwise inaccessible area. Early work focused on the use of time-of-flight methods, in which transducers are moved along the pipe to measure the wall thickness, by calculating the time delay between transmission and reception of the ultrasonic wave propagating through the pipe wall (Krautkrämer and Krautkrämer 1983). The transducers are either hand-held or guided by a robot, known as a pipe inspection gauge (PIG), traveling in the pipe (Tiratsoo 1992). In either case it takes a long time and significant labor to inspect the many sections of pipes.

Guided wave techniques gained interest for corrosion detection due to their long propagation distance and sensitivity to small defects. Three corrosion sensitive features were proposed and examined in (Zhu et al. 1998): wave mode elimination, time delay, and attenuation. Many other researchers (Tuzzeo and di Scalea 2001; Silva et al. 2003; Su et al. 2006) have discussed different applications and methodologies to determine the presence of hidden corrosion from these features. These features are sensitive to large corrosion areas, because the thinning effect changes the frequency-thickness product of the media, thereby inducing a shift in guided wave dispersion curves. As a consequence, the number of wave modes in a given frequency range may change and wave velocities may shift.

Although these features can be extracted from records on plate-like structures, they are not as effective for pipes. Due to the cylindrical shell geometry of pipes, there exist a much greater number of wave modes at each frequency (Lowe et al. 1998; Rose 2004). This makes it difficult to directly interpret the mode elimination, if any exists. Also, if the corrosion is highly localized(a condition also known as pitting corrosion) it may not produce sufficiently

large changes in the dispersion curve. Finally, because the features are based on dispersion characteristics of guided waves in pipes, they are very sensitive to environmental and operational conditions (EOC) and can easily produce false alarms if no compensation is made for those benign variations (Konstantinidis et al. 2006; Sohn 2007).

In this section, we further extend our SVD method to identify progressive degradation on a pipe that is experiencing temperature variations.

4.3.2 Proof of concept (synthetic) experiment

We first conducted a synthetic experiment of a pipe undergoing both temperature variation and degradation. We collected one baseline signal of a laboratory pipe specimen while it was intact. We then obtained one pitch-catch record while there was a grease-coupled mass scatterer on the pipe, physically simulating damage. The scatter signal was obtained from the difference between the two signals. We then simulated temperature variation by stretching the baseline signals with 500 stretching factors calculated from a set of sequential field experimental data under real operating conditions. The stretching factors were periodically varied from -0.006 to 0.006. After obtaining the 500 synthetic temperature-affected signals, we simulated a gradually increasing degradation by gradually increasing the amplitude of the scatterer signal in the synthetic signals.

Figure 4-5 shows the synthetic SVD results, in simulation, detecting degradation under temperature variation with time. The bottom curve (in green) shows the temperature variation. The upper dash line (in red) describes the degradation in terms of the ratio of energy between the scatter signal and the baseline signal. As shown in the curve, in 400 records, the damage signal energy increases from 0 to 2% of baseline energy, which is significantly smaller compared to the energy change caused by temperature variation. The top solid curve (in blue)

is the extracted singular vector. Despite the small ripples corresponding to the temperature variation, the singular vector clearly follows the trend of the degradation.

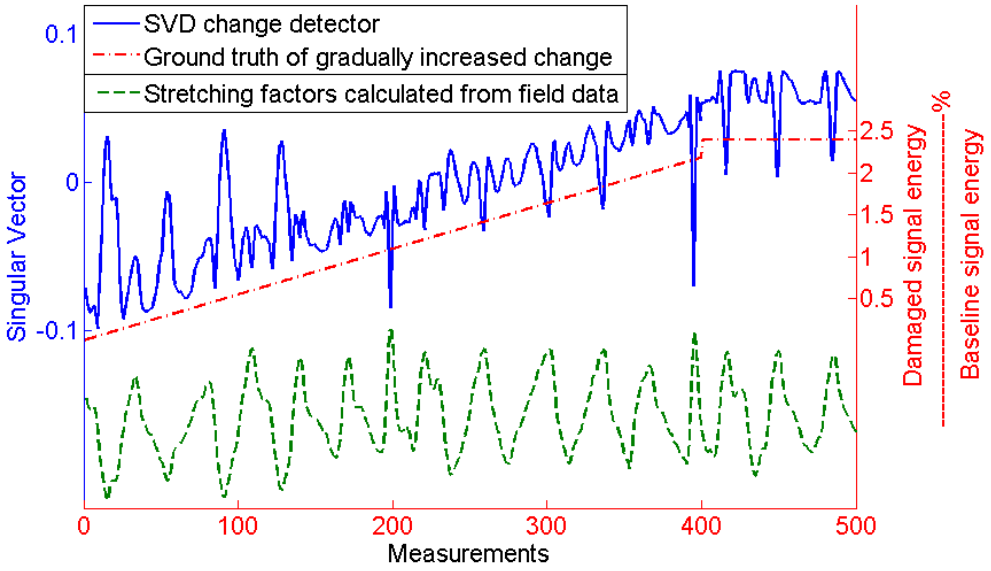


Figure 4-5 SVD degradation detection using synthetic simulation data

4.3.3 Experimental setup

We investigated the pipe degradation problem in a laboratory experiment where we physically produce pitching corrosion damage on an aluminum pipe specimen. Figure 4-6 shows the tested aluminum pipe specimen with 89 mm diameter, 1320 mm length, and 3 mm wall thickness. We bonded lead zirconium titanate (PZT) wafer transducers (7x6x1mm) at two different locations on the pipe. Each array is located 228 mm from the pipe end with 8 PZT transducers uniformly distributed around the circumference. The wafers were bonded to the surface of the pipe with superglue.

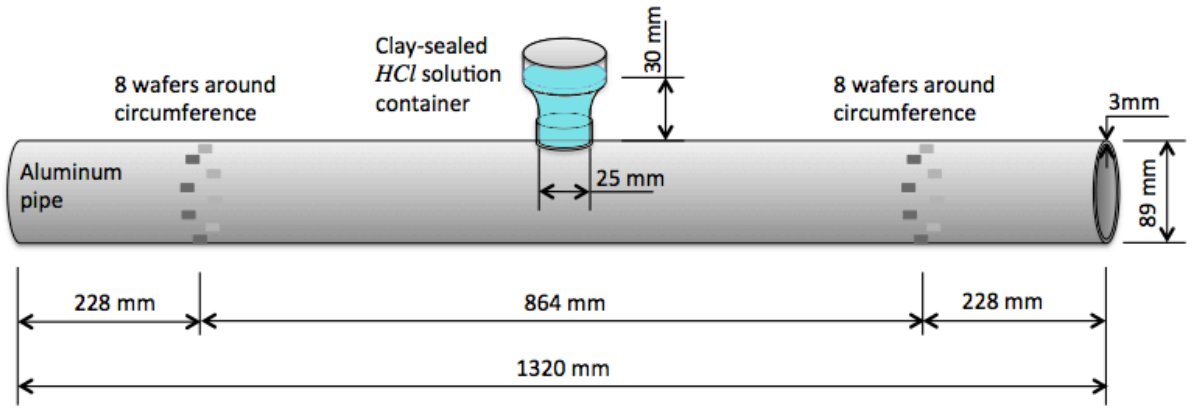


Figure 4-6 Specimen dimension and experimental setup for corrosion simulation.

We took pitch-catch records using a National Instruments data acquisition system. We transmitted a windowed sinusoidal signal from one transducer and recorded signals at multiple receivers at the far array. The excitation was a 100 microsecond sinc-windowed sinusoid with a wide frequency band from 50 kHz to 450 kHz. At the receivers, we recorded 2.5 milliseconds at a sampling rate of 4MHz, resulting in a 10,000-sample signal per receiver per pitch-catch records. We took one such record every minute during the 50-hour experiment period, during which we physically simulated structural degradation (corrosion) and environmental variations (temperature). We then sought to identify and differentiate the two trends by processing all sequential pitch-catch records, a total of 3000, with our SVD-based DSF extractor.

Figure 4-7 shows a photograph of the experimental setup. We toggle the heater to create temperature variations on the pipe while we collect pitch-catch records. Because the aluminum pipe has thin wall and that aluminum has good heat conductivity, the heater can effectively increase the temperature on the pipe from the distant position as shown in the photo. The heater swings during the heating phase to cover a larger area of the pipe, but there are inevitable temperature gradients on the pipe because of the thermal conduction both

along the length of the pipe and around the circumference of the pipe. The complex temperature profile on the pipe significantly changes the ultrasonic records and challenges the identification of corrosion damage.

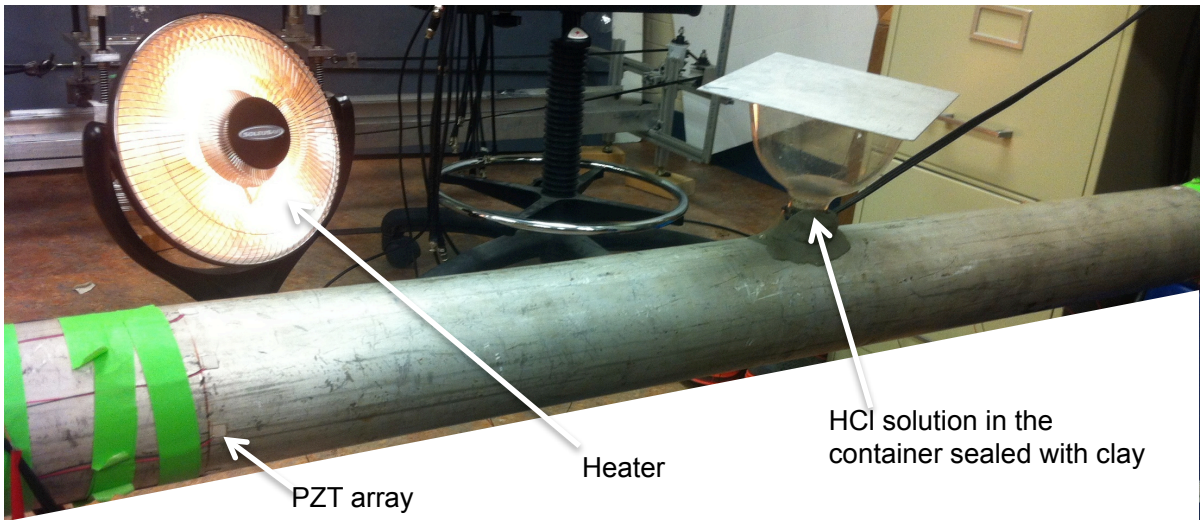


Figure 4-7 Photo of experimental setup

We applied 1.0 normal hydrochloride acid solution on the pipe to gradually induce a pitting corrosion at the center of the pipe segment, midway between the two arrays. Figure 4-7 shows (on the right) the plastic container that is sealed with plastic clay to hold the acid solution. The container is then covered with a cap to prevent the acid from evaporating into the environment. By controlling the amount of acid, corrosion was limited to a partial thickness. Therefore, in the entire experiment period (~50 hours) it is reasonable to assume that the mass loading on the pipe changed only negligibly, and that the degradation was the reduction in wall thickness caused by corrosion. Figure 4-8 shows a photograph of the pipe after our 50-hour experiment, in which we can see the pitting corrosion damage created by the acid solution. We measured the dimension of the corrosion pitch with a dial caliper, obtaining a length of 0.665” (16.9 mm) axially, and a width of 0.494” (12.5 mm)

circumferentially. We molded the corrosion pitch with plastic clay and estimated the depth of the corrosion pit to be 0.037” (0.94 mm), roughly 30% of the pipe wall thickness.

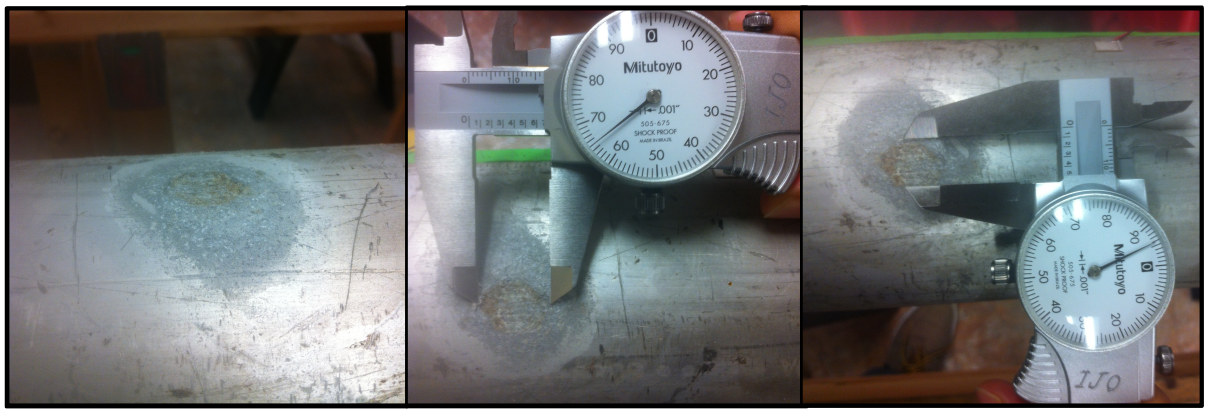


Figure 4-8 Photo of the pitching corrosion damage after 50 hours of corrosion process

4.3.4 Experimental procedure

We then examined our experimental data and showed that our SVD method is effective and robust in identifying degradation from benign variations. Figure 4-9 shows three plots for typical received signals and their Fourier spectra, collected under different testing scenarios: when the specimen 1) was intact, 2) was at a different temperatures, and 3) after corroding. Because of the complexity of guided waves, changes cannot be observed by directly comparing the three waveforms or their frequency spectra. It is also impossible to distinguish the signal affected by temperature from the signal affected by physical degradation.

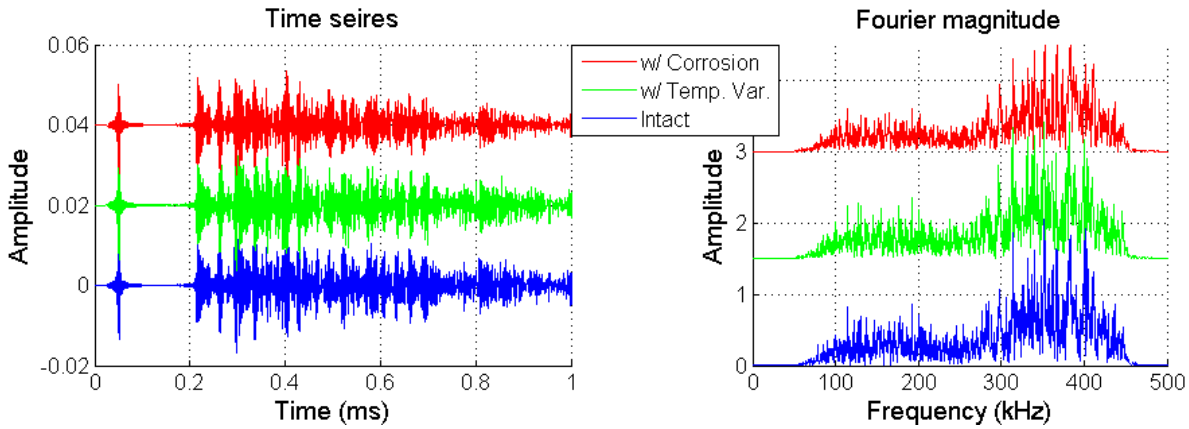


Figure 4-9 Typical signals. From bottom to top: intact signal; signal affected by temperature variation; signal affected by corrosion

We first examine correlation coefficients of the records to characterize the experimental process. Figure 4-10 shows the correlation coefficients of all the records we collected in the experiments with respect to the first record, which serves as a baseline. As defined in equation (2-1), correlation coefficients characterize the similarity between two waveforms. Therefore, the correlation coefficients of a batch of records collected over a period of time with a baseline signal will indicate how the received records change over the period of time, and reveal variations on the subject being monitored. In laboratory experiments performed in a static environment, the change in correlation coefficients could (in principle) indicate the change caused by damage. However, environmental variations also produce significant change in the signal, thus making correlation coefficients ineffective as a damage feature.

In test period 1, we first applied one single heating cycle to the pipe specimen, and then started the corrosion process after the pipe was fully cooled. In test period 2, superimposing the corrosion that was continuously induced in the pipe, we also applied heating cycles within the experiment, indicated by shading in Figure 4-10.

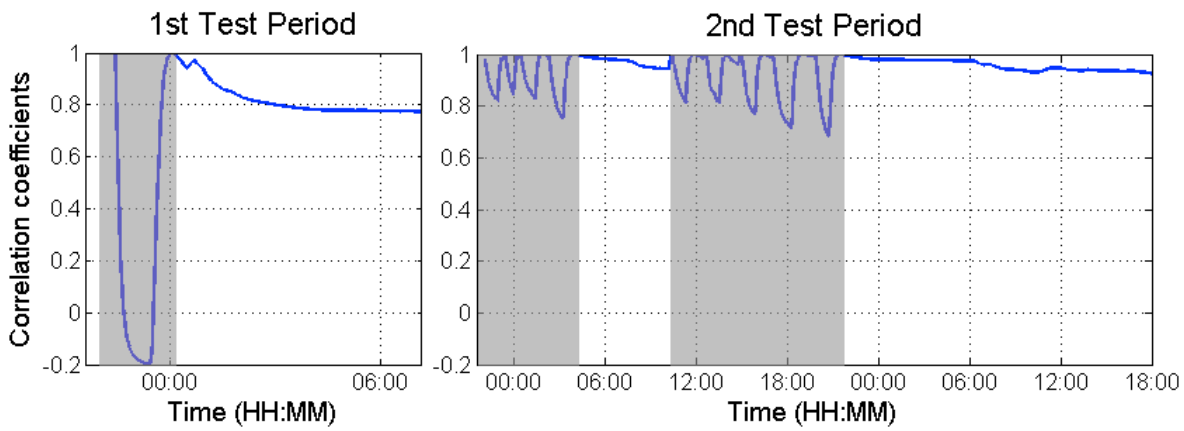


Figure 4-10 Correlation coefficients of data, describes how measured signals change over 72 hours of time. Shading indicates the presence of temperature variations.

We make three observations from the correlation coefficients shown in Figure 4-10:

(1) In the first test period, we observe correlation coefficients decrease to -0.2 as we heated the pipe, and then increase back to 1.0 as the pipe cooled down. We also observe that the correlation coefficients decrease as we apply acid on the pipe to produce corrosion damage while we keep the temperature steady. We conclude from this that we can identify the slow degradation from correlation coefficients when temperature change is absent.

(2) We also observe that temperature variation creates more changes in the correlation coefficients than corrosion, indicating that the temperature variations produce more changes in the records, which can mask the change of interest.

(3) In the second test period, although in the un-shaded time period we can identify the degradation from the decrease of correlation coefficients, we see no indication of degradation at all when temperature variation exists. We cannot identify the degradation because the changes are masked by the temperature change.

In none of the test scenarios can we easily tell the presence of degradation from correlation coefficients, because the changes created by temperature variations and by degradation are coupled. In order to accurately identify the corrosion degradation, we need to separate the two coupled processes. In the next subsection, we show that our SVD method successfully decouples the two variations, making it possible for us to identify the degradation in the presence of temporal and spatial temperature variations.

4.3.5 SVD results

We then applied SVD to the 3000 pitch-catch records and extracted the singular vectors, which indicate the different types of variations. Figure 4-11 and Figure 4-12 compare the correlation coefficients and the extracted singular vectors for the two testing periods.

In test period 1, temperature variation and corrosion were separately induced, yet both manifest themselves as a drop in correlation coefficients. However, as we plot the singular vectors extracted from the data matrix, we obtain two uncoupled singular vectors describing the trend of temperature and degradation separately. As shown in the solid line in the right plot in Figure 4-11, the 1^{st} singular vector captures the temperature variation that was induced before 00:00. It then stays constant for the remainder of the test period indicating that no other temperature variation was involved in the experiment. The dashed line plots the 6^{th} singular vector, which shows little variation before 00:00, followed by a monotonically varying trend describing the corrosion process. The decoupling essentially makes singular vectors a robust damage-sensitive feature.

In test period 2, the pipe underwent continuous corrosion and several heating cycles were induced. Due to discontinuity of heating, the tests could be divided into four intervals in Figure 4-12: two shaded intervals in which both temperature variation and corrosion were

present, and two unshaved intervals featuring only corrosion. It is very difficult to identify the degradation from the correlation coefficients, especially when the temperature is cycling. The right-hand plot of Figure 4-12 plots the 5th singular vector extracted from the records. Visually, we observe a consistent increasing trend with the degradation. In contrast, the 1st singular vector in Figure 4-12 still shows large periodic changes similar to those of the temperature variations.

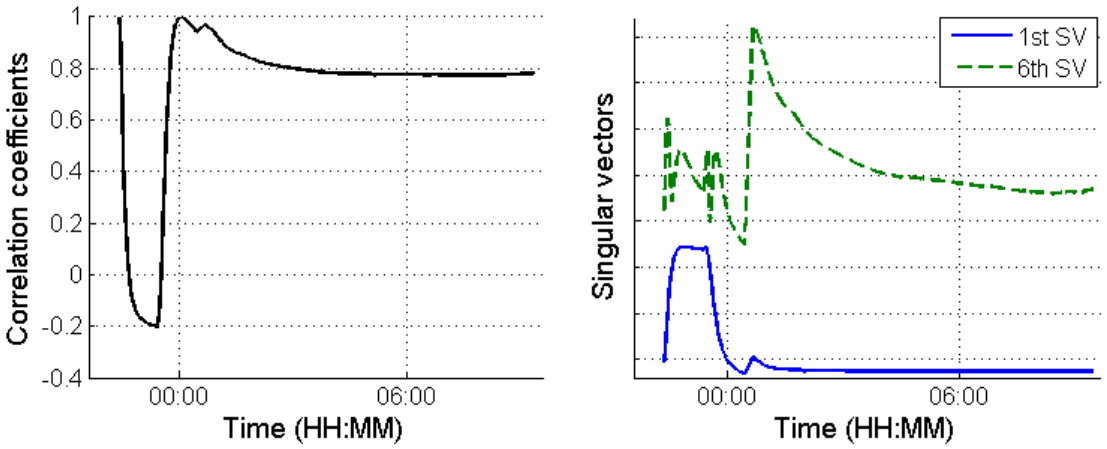


Figure 4-11 Comparison of correlation coefficients and SVD on test period 1. SVD decoupled temperature variation and degradation.

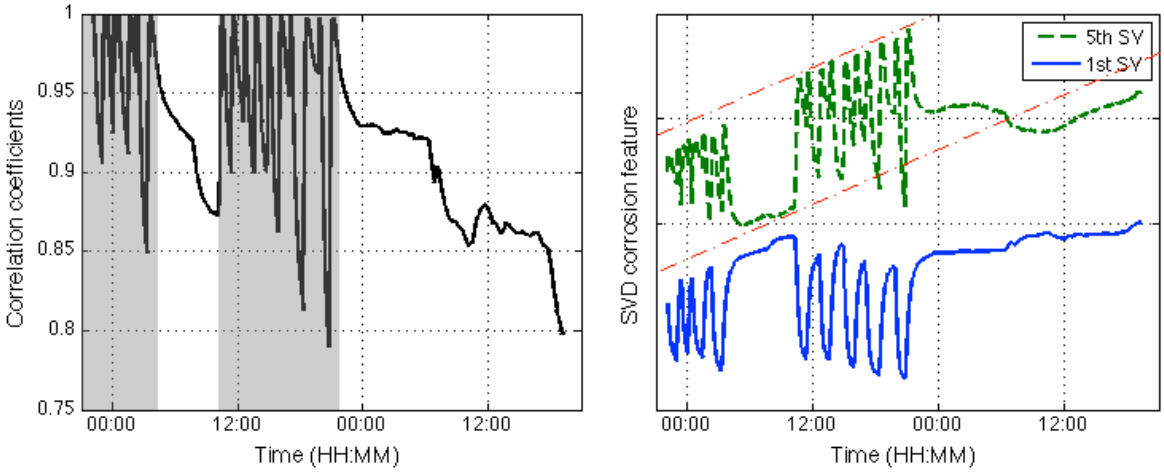


Figure 4-12 Comparison of correlation coefficients and SVD on test period 2. Red dot-dashed auxiliary lines indicate the monotonic trend in singular vector.

More specifically, if we draw two auxiliary trend-lines to cover the 5th singular vector in Figure 4-12, we see that the monotonic increasing trend has a consistent slope in all four intervals, despite the presence or absence of temperature variation. We observe that the 5th singular vector is still affected by the temperature variation; one possible explanation is that temperature variation was also influencing the reaction rate of the corrosion process.

4.4 Summary

Guided wave based damage detection and localization is attractive as a cost-efficient, long-term structural health monitoring technique. However, environmental and operational variations, including temporal temperature variations and spatial temperature gradients, remain challenges for practical implementation of guided wave damage diagnosis systems. Although various temperature compensation methods have been developed to target this problem, they do not successfully remove the effects of temperature gradients on measured data. In this chapter, we applied our singular value decomposition method on real experimental monitoring records to robustly detect damage when the records are affected by environmental and operational variations.

In *Section 4.2* we first show the damage detection result on ultrasonic records from our field experiments described in *Chapter 2*. We show that despite the large environmental and operational variations that defeat conventional damage detection methods, our SVD method successfully detects the events when we applied or removed a small, grease-coupled mass scatterer on the pipe. We then further illustrated the efficacy of the method on multiple datasets collected spanning over seven months, and showed that the selected sensitive feature performs consistently well under different operating conditions

In *Section 4.3*, we extended the method to identify progressive degradation on a pipe, and demonstrated that singular vectors are effective degradation-sensitive features and are robust to temporal and spatial temperature variations. We conducted an experiment in which we gradually induced pitting corrosion damage in an aluminum pipe while undergoing temperature variations. We took ultrasonic pitch-catch records at one-minute interval throughout the experimental period (~50 hours) and then used singular value decomposition to detect degradation. We obtained promising results, separating changes caused by degradation from those caused by temperature variation. We showed that a conventional method (correlation coefficients with respect to a baseline record) can detect the degradation in a steady environment, but fails when there exist temperature variations. We then applied our SVD method on the monitoring records and showed that our method can successfully decouple the two variations and reveal the trend of the progressive degradation.

The purposes of the experimental investigations in this chapter were (1) to provide a practical implementation of the singular value decomposition method and the damage-sensitive feature selection process introduced in *Chapter 3*, and (2) to demonstrate that we can use the left singular vector as a damage-indicator that is robust to environmental and operational variations. The experimental results agree very well with our simulation results in *Chapter 3*, and confirm that our singular value decomposition method can effectively and robustly extract useful information from monitoring records to detect damage and identify degradation even in the presence of large environmental and operational variations.

In the next chapter, we further extend our singular value decomposition method to damage localization. We will show that current delay-and-sum localization relies on accurate removal of baseline information, which is often unreliable when there exist environmental

and operational variations. By applying singular value decomposition on monitoring records and utilizing the damage-sensitive right singular vector, we can effectively remove both the baseline information and the variations, to accurately recover the scatter signals originating from the damage site. We can then use the scatter signals to localize the damage. Due to the removal of the variations, we can significantly improve the robustness and accuracy of damage localization.

Chapter 5 Robust Baseline Removal For Guided Wave Damage Localization

5.1 Motivation And Initial Analysis

Guided wave localization relies on extracting the accurate phase and amplitude information from the scatter signals that originated from the damage site. Scatter signals are usually obtained by subtracting baseline signals from subsequent observations. However, baseline subtraction is not ideal when the baseline signals are affected by environmental and operational effects, among which temperature variation is most significant and widely studied. Variations in temperature change the ultrasonic wave velocities and therefore make baseline subtraction imperfect. In the optimal baseline selection method, one keeps a library of baseline records and selects a baseline that optimally matches the monitoring records to perform the subtraction. A large group of methods, including optimal baseline stretch (Konstantinidis et al. 2007; Clarke et al. 2010) and scale transform based method (Harley and Moura 2012), assume that a change in temperature approximately stretches or compresses the

measured record. Those methods compensate for the temperature change by estimating the stretching factor and reverse the stretching effect caused by temperature change. In the following subsections, we show that the efficacy of these methods diminishes when there exists a spatially temperature gradient.

5.1.1 Spatially uniform temperature variation

We collect baseline records at known temperature, and compare each new measurement record, which may be taken at a different temperature, with the baseline(s) to identify damage. If the temperature change between the baseline and the measured records is small and spatially uniform across the structure, stretch-based temperature compensation methods have been shown to effectively compensate temperature change and successfully extract the scatter signal (Michaels 2008)(Putkis and Croxford 2013b).

We assume a baseline record $\boldsymbol{x}_b(t)$ is collected from an intact specimen at a spatially uniform temperature T_0. Some time later during the monitoring process a defect is introduced to the specimen. Our new record $\boldsymbol{x}_i(t)$ will contain an additive signal that is produced by the defect as a scatter signal $\boldsymbol{x}_s(t)$,

$$\boldsymbol{x}_i(t) = \boldsymbol{x}_b(t) + \boldsymbol{x}_s(t), \tag{5-1}$$

If the temperature is steady, then we can obtain the scatter signal by simple baseline subtraction.

If the record is collected at a different temperature $T_0 + \Delta T$, then

$$\boldsymbol{x}_i(t) = \boldsymbol{T}\{\boldsymbol{x}_b(t) + \boldsymbol{x}_s(t)\}, \tag{5-2}$$

where, and $\boldsymbol{T}\{\cdot\}$ is the change caused by temperature. When there is a temperature change, baseline subtraction can no longer perfectly remove the baseline, and temperature compensation methods are needed.

Temperature compensation methods remove the effects of temperature in one of two ways. Temperature compensation methods compensate for temperature effects in the measured data $\boldsymbol{x}_i(t)$,

$$\boldsymbol{x}_s'(t) = \boldsymbol{T}^{-1}\{\boldsymbol{x}_i(t)\} - \boldsymbol{x}_b(t), \tag{5-3}$$

Or, temperature compensation methods modify the baseline record $\boldsymbol{x}_b(t)$ to experience the same temperature effect as $\boldsymbol{x}_i(t)$,

$$\boldsymbol{x}_s'(t) = \boldsymbol{x}_i(t) - \boldsymbol{T}^{*}\{\boldsymbol{x}_b(t)\}. \tag{5-4}$$

In (5-3) and (5-4), $\boldsymbol{x}_s'(t)$ is the recovered scatter signal, and $\boldsymbol{T}^{-1}\{\cdot\}$ is the inversed temperature change. Baseline selection methods take the form of equation (5-4) where the function $\boldsymbol{T}^{*}\{\cdot\}$ selects the optimal baseline from a baseline library. Stretching based temperature compensation methods can be implemented in either way. For these approaches, the temperature effect $\boldsymbol{T}\{\cdot\}$ is approximated by a stretching model, which assumes that a small temperature change stretches/compresses ultrasonic signals,

$$\begin{aligned}&\boldsymbol{T}_{\alpha}\{\boldsymbol{x}_i(t)\} \approx \boldsymbol{x}_i(\alpha t),\\ &\boldsymbol{T}_{\alpha}^{-1}\{\boldsymbol{x}_i(t)\} \approx \boldsymbol{x}_i(\alpha^{-1} t),\end{aligned} \tag{5-5}$$

where, α is the stretching factor between the two records and is proportional to ΔT when ΔT is small.

If temperature compensation of $\boldsymbol{x}_i(t)$ in (5-2) is implemented according to (5-3), then

$$\boldsymbol{x}_s'(t) = \boldsymbol{T}^{-1}\{\boldsymbol{T}\{\boldsymbol{x}_b(t) + \boldsymbol{x}_s(t)\}\} - \boldsymbol{x}_b(t)$$
$$\approx \boldsymbol{x}_b(t) - \boldsymbol{x}_b(t) + \boldsymbol{x}_s(t). \tag{5-6}$$

Similarly, if temperature compensation of $\boldsymbol{x}_i(t)$ in (5-2) is implemented according to (5-4), then

$$\boldsymbol{x}_s'(t) = \boldsymbol{T}\{\boldsymbol{x}_b(t) + \boldsymbol{x}_s(t)\} - \boldsymbol{T}^*\{\boldsymbol{x}_b(t)\}$$
$$\approx \boldsymbol{T}\{\boldsymbol{x}_b(t) - \boldsymbol{x}_b(t)\} + \boldsymbol{T}\{\boldsymbol{x}_s(t)\}. \tag{5-7}$$

Equations (5-6) and (5-7) show that if the inverse temperature effect is exact, or if the baseline selected experienced the same temperature effect as the measured record, then we can remove the baseline signal and recover the scatter signal after conventional temperature compensation.

5.1.2 Spatial temperature gradients

For equations (5-6) and (5-7) to be true, and therefore for conventional temperature compensation to be effective, we implicitly assume in equation (5-2) that the baseline signal and the scatter signal are affected by the same temperature variation. In many applications, this is not necessarily a valid assumption. For example, if a large structure experiences a spatial temperature gradient, the direct and scatter signals will experience different temperature changes. Because many structures are of large scale and are located in an outdoor environment, complex spatial temperature profiles are common. For example, outdoor shell or plate structures experience different sunshine conditions at different locations; long pipeline segments experience large temperature gradients and large water or oil tanks experience temporal and spatial temperature variations associated with changes in fluid level and with different exposure to insolation and to wind effects. In this subsection,

we first demonstrate that conventional temperature compensation methods are not effective when temperature gradients are present.

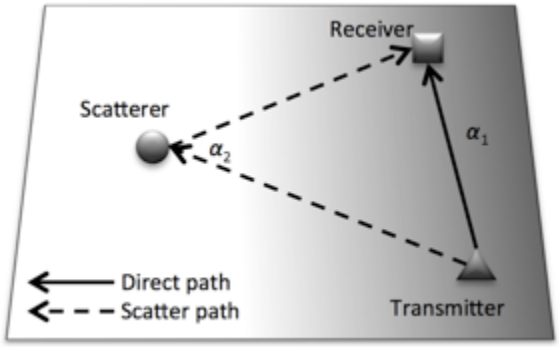


Figure 5-1. Schematic drawing of a plate with spatial temperature variation.

In Figure 5-1, we illustrate a plate with a shade-coded temperature distribution to show how spatial temperature gradients affect conventional temperature compensation methods. In this example, we consider two piezoelectric transducers attached to the plate. When the plate was intact and at room temperature T_0, we collect a baseline pitch-catch signal $\boldsymbol{x}_b(t)$. If the plate is heated along the right edge, a horizontal temperature gradient will develop across the plate as shade-coded in Figure 5-1. When a scatterer is then placed on the plate, the measured record $\boldsymbol{x}_i(t)$ will contain signals from both the direct path and the scatter path. The baseline signal $\boldsymbol{x}_b(t)$ propagates from transmitter to the receiver along the direct path indicated by solid arrow. The dashed arrow indicates the scatter path, which is from the transmitter to the scatterer and then to the receiver. Because the temperature on the plate is uniform when the baseline was collected but exhibits a gradient after heating, the two paths experience different temperature changes, i.e., $\Delta T_{direct} \neq \Delta T_{scatter}$. Therefore the assumption made in equation (5-2) is not valid. Instead we have,

$$\boldsymbol{x}_i(t) = \boldsymbol{T}_1\{\boldsymbol{x}_b(t)\} + \boldsymbol{T}_2\{\boldsymbol{x}_s(t)\}. \tag{5-8}$$

In this scenario, the temperature compensation process shown in (5-6) becomes,

$$\boldsymbol{x}_s'(t) = \boldsymbol{T}^{*-1}\{\boldsymbol{T}_1\{\boldsymbol{x}_b(t)\} + \boldsymbol{T}_2\{\boldsymbol{x}_s(t)\}\} - \boldsymbol{x}_b(t), \tag{5-9}$$

and temperature compensation process shown in (5-7) becomes

$$\boldsymbol{x}_s'(t) = \boldsymbol{T}_1\{\boldsymbol{x}_b(t)\} + \boldsymbol{T}_2\{\boldsymbol{x}_s(t)\} - \boldsymbol{T}\{\boldsymbol{x}_b(t)\}. \tag{5-10}$$

From equation (5-9), we can analyze the performance of the baseline selection method under a spatial temperature gradient. Assuming the method successfully selects the optimal baseline, i.e. $\boldsymbol{T}^*\{\cdot\} = \boldsymbol{T}_1\{\cdot\}$, the baseline is perfectly removed. The recovered scatter signal $\boldsymbol{x}_s'(t)$ would then equal the true scatter signal $\boldsymbol{x}_s(t)$ affected by the temperature change along the scatter path $\boldsymbol{T}_2\{\cdot\}$. When temperature change is uniform across the structure, $\boldsymbol{T}_2\{\cdot\}$ are identical for all transmitter-receiver pairs. In contrast, when there is a temperature gradient across the structure, $\boldsymbol{T}_2\{\cdot\}$ is different for each transmitter-receiver pair. This poses a challenge to localization because each sensor observes different temperature effects.

If we apply the stretching based temperature compensation implemented in (5-10) with the spatial temperature gradient described in (5-8) we obtain,

$$\boldsymbol{x}_s'(t) = \boldsymbol{T}_{\alpha_2/\alpha_{est}}\{\boldsymbol{x}_s(t)\} + \left(\boldsymbol{T}_{\alpha_1/\alpha_{est}}\{\boldsymbol{x}_b(t)\} - \boldsymbol{x}_b(t)\right), \tag{5-11}$$

where α_1, α_2 are the stretching factors for the direct path and the scatter path, respectively, and α_{est} is the stretching factor estimated by the temperature compensation method. Equation (5-11) shows that when spatial temperature gradients are present, stretching-based temperature compensation is ineffective because the scatter-path stretching factor, the direct-path stretching factor, and the estimated stretching factor are all different. Consequently, the extracted scatter signal $\boldsymbol{x}_s'(t)$ is composed of a stretched scatter signal and an error term that comes from the imperfect recovery of the baseline.

We analyze the differences in the extracted scatter signal $\boldsymbol{x}_s'(t)$ and the true scatter signal $\boldsymbol{x_s}(t)$ in three separate scenarios: $\alpha_{est} = \alpha_2, \alpha_{est} = \alpha_1$, and $\min(\alpha_1, \alpha_2) < \alpha_{est} < \max(\alpha_1, \alpha_2)$.

If $\alpha_{est} = \alpha_2$, (5-11) becomes

$$\boldsymbol{x}_s'(t) = \boldsymbol{x}_s(t) + \left(\boldsymbol{T}_{\alpha_1/\alpha_2}\{\boldsymbol{x}_b(t)\} - \boldsymbol{x}_b(t)\right). \tag{5-12}$$

That is, the extracted scatter signal differs from the true scatter signal due to imperfect baseline removal. Considering that the baseline is usually an order of magnitude larger than the scatter signal, the error term could be sufficiently large to mask the scatter signal. As a result, the extracted signal contains little information that could be used to localize damage. Instead, it contains mostly information about the direct path, which, if used to localize damage, will generate false positives at transducer locations, where the direct path signal originates.

If $\alpha_{est} = \alpha_1$, then (5-11) becomes

$$\boldsymbol{x}_s'(t) = \boldsymbol{T}_{\alpha_2/\alpha_1}\{\boldsymbol{x}_s(t)\} + \left(\boldsymbol{x}_b(t) - \boldsymbol{x}_b(t)\right). \tag{5-13}$$

In this case, the baseline is perfectly removed. However, the extracted scatter signal is a stretched/compressed version of the true scatter signal, which corresponds to the earlier discussion of the baseline selection method. If we use the recovered scatter signals to localize damage, the arrivals will not focus well on the damage location and will cause noisy results.

If the estimation produces a value in between α_1 and α_2, $\min(\alpha_1, \alpha_2) < \alpha_{est} < \max(\alpha_1, \alpha_2)$, then both the baseline and scatter terms in (5-11) are inaccurate and produce error in the extracted signal. Based on our analysis of the extreme, we expect the localization result to be both noisy and negatively affected by false alarms near transducer locations.

In the next section, we demonstrate a robust baseline removal method, based on singular value decomposition, and show how it overcomes these challenges. In *Section 5.3*, we then integrate our approach with delay-and-sum localization (Michaels 2008) and illustrate that the resulting localization method is robust to environmental variations, including both uniform temperature variation and temperature variation with spatial gradients. We further show experimental localization results on an aluminum pipe in *Section 5.4* as a demonstration of the generality of our method on different applications.

5.2 Methodology

Singular value decomposition (SVD) seeks a linear decomposition of a data matrix that creates a set of orthonormal bases to represent the data. For a batch of pitch-catch records $\boldsymbol{X}^{N\times D} = [\boldsymbol{x}_1^T, \boldsymbol{x}_2^T, \dots, \boldsymbol{x}_N^T]$, where N is the dimension in slow time indicating the number of records, and D is the dimension in fast time indicating the number of voltage samples in one pitch-catch record, the SVD of $\boldsymbol{X}$ is denoted as

$$\boldsymbol{X}^{N\times D} = \boldsymbol{U}^{N\times N}\boldsymbol{S}^{N\times N}(\boldsymbol{V}^H)^{N\times D}, \tag{5-14}$$

where, superscript H is the conjugate transpose. $\boldsymbol{U}^{N\times N} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_N]$ is the left singular vector matrix, and each $\boldsymbol{u}_\mathrm{i}$ is a $N\times 1$ vector representing some variation in the slow-time domain – the trend of changes over many observations $\boldsymbol{S}^{\mathrm{N}\times\mathrm{N}}$ is the diagonal singular value matrix with diagonal terms $[\sigma_1, \sigma_2, \dots, \sigma_M]$ that are known as singular values of $\boldsymbol{X}$, typically in a descending order, and $\boldsymbol{V}^{D\times N} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_N]$ is the right singular vector matrix, where each $\boldsymbol{v}_\mathrm{i}$ is a $D\times 1$ vector representing a change in the fast-time domain that corresponds to the variation manifested in $\boldsymbol{u}_\mathrm{i}$.

In *Chapter 3* and *Chapter 4*, we showed that if we apply SVD on a set of ultrasonic signals collected under large environmental variations, we would observe the temperature variation and damage to be manifested in different left singular vectors $\boldsymbol{u}_i$. Particularly, if a batch of records is affected by cyclic temperature variations and by an abrupt change produced by damage, one of its left singular vector captures the abrupt change of damage as a step-like function, while other left singular vectors are visually associated with the cyclic temperature effect. We can then identify the presence of damage from the selected singular vector, even if the temperature variation is large. More interestingly, we showed that if $\boldsymbol{u}_i$ resembles the abrupt change of damage, then the i^{th} right singular vector $\boldsymbol{v}_i$ captures the scatter signal produced by the damage. This relationship holds true because the scatter signal is exactly the fast-time change associated with the abrupt variation in the slow-time domain. We can intuitively understand this relationship by multiplying both sides of (14) by $\boldsymbol{V}$, and use the fact that $\boldsymbol{V}$ is a unitary matrix, i.e. $\boldsymbol{V}^H\boldsymbol{V} = \boldsymbol{I}$.

$$\boldsymbol{XV} = \boldsymbol{USV}^H\boldsymbol{V}$$

$$\boldsymbol{XV} = \boldsymbol{US}$$

$$\boldsymbol{X}^{N\times D}\boldsymbol{v}_i^{D\times 1} = \boldsymbol{u}_i^{N\times 1}\sigma_i. \tag{5-15}$$

Equation (5-15) shows that the i^{th} left singular vector $\boldsymbol{u}_i$ is proportional to the projection of data matrix $\boldsymbol{X}$ on the corresponding i^{th} right singular vector $\boldsymbol{v}_i$. If $\boldsymbol{u_i}$ resembles the occurrence of damage across measurements, then in order for the projection of $\boldsymbol{X}$ on $\boldsymbol{v}_i$ to reflect that trend, $\boldsymbol{v_i}$ must resemble the changes in the waveform associated with the trend, i.e. the change produced by damage. Therefore, the selected right singular vector $\boldsymbol{v}_i$ reproduces the scatter signal. We will show in the next subsection that due to the orthogonality characteristics of singular vectors, $\boldsymbol{v}_i$ is not affected by the temperature variations that

manifest in other singular vectors. Therefore, we can use the extracted scatter signal as a robust baseline removed signal to facilitate damage localization under temperature variations.

The implementation of singular value decomposition robust baseline removal proceeds as follows. We collect N sets of records from M channels, where each record contains D voltage samples. For the m^{th} channel (m from 1 to M), we filter the records with a narrow band filter at some selected frequency that is effective for localization, and compute singular value decomposition on data matrix $\boldsymbol{X}_m^{N\times D}$, obtaining three matrices: $\boldsymbol{U}$, $\boldsymbol{S}$, and $\boldsymbol{V}$. From our previous discussion in *Chapter 3*, we know that left singular vectors $\boldsymbol{u}_i^{1\times N}$ represent slow-time trends over the N records. We also know that one of them best represents the change caused by the damage, which in our case is a step function. We then identify the damage-sensitive left singular vector $\boldsymbol{u}_k^{1\times N}$ with the method based on K-means that is developed in *Section 3.2*. From equation (5-15), we know that the corresponding right singular vector $\boldsymbol{v}_k^{1\times D}$ is the baseline-removed signal for the m^{th} channel. Note that because the temperature effect can be different across different channels, we do not expect the order of damage sensitive singular vector to be the same, i.e. k_i does not necessarily equal k_j for $i \neq j \in [1, M]$. Therefore, we repeat the above baseline removal and damage-sensitive feature selection process for each of the M channels, obtaining a scatter signal matrix $\boldsymbol{X}_S^{D\times M}$, which we then use to localize damage with any preferred localization algorithms.

5.3 Experimental Results On A Plate

We validate the proposed method with an experiment on an aluminum plate in a laboratory environment. We heated the plate non-uniformly to create both temporal and spatial temperature variations while we take pitch-catch records from 16 PZT transducers

that are adhered to the plate. We applied a grease coupled mass scatterer on the plate and sought to localize the scatterer in the presence of complex (temporal and spatial) temperature variations.

5.3.1 Experimental details

5.3.1.1 Experimental setup

Figure 5-2 shows the dimensions and the transducer placement of the plate specimen in a schematic drawing, and Figure 5-3 shows a photograph of the experimental setup. The dimension of the aluminum plate is $1.213 \times 1.213 \times 0.003$ m. We adhered 16 PZT wafer transducers to the plate at locations labeled with black crosses. We place the PZT wafers 'randomly' to make the distances between transducers non-uniform, such that dispersion curves of the plate can be extracted by the Sparse Wavenumber Analysis techniques developed in Harley and Moura 2013. SWA is not described or discussed here, but we note that our sensor placement is consistent with effective performance of SWA.

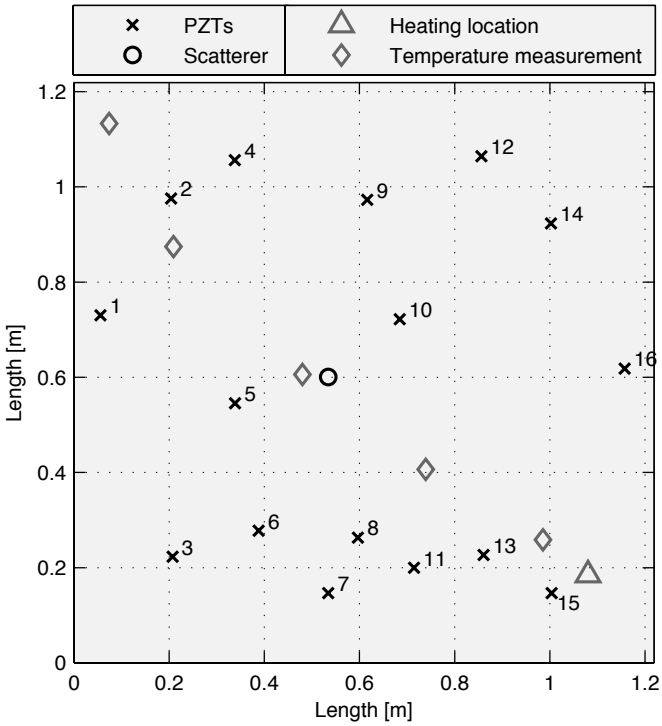


Figure 5-2. (a) Plate specimen dimensions, along with transducer locations (x), scatter location (o), heater location (Δ), and temperature measurement locations (◊); (b) Schematic chart of data acquisition system configuration.

Figure 5-3 shows a photograph of the instrumented aluminum plate specimen and the mass scatterer near the center of the plate. The mass scatterer is a bronze cylinder of 1.5 kg mass and 5 cm diameter, and is acoustically coupled to the plate specimen with vacuum grease such that it can scatter guided waves that propagate through the area. The scattering effect produces changes the received pitch-catch records, which we will use to localize the scatterer.

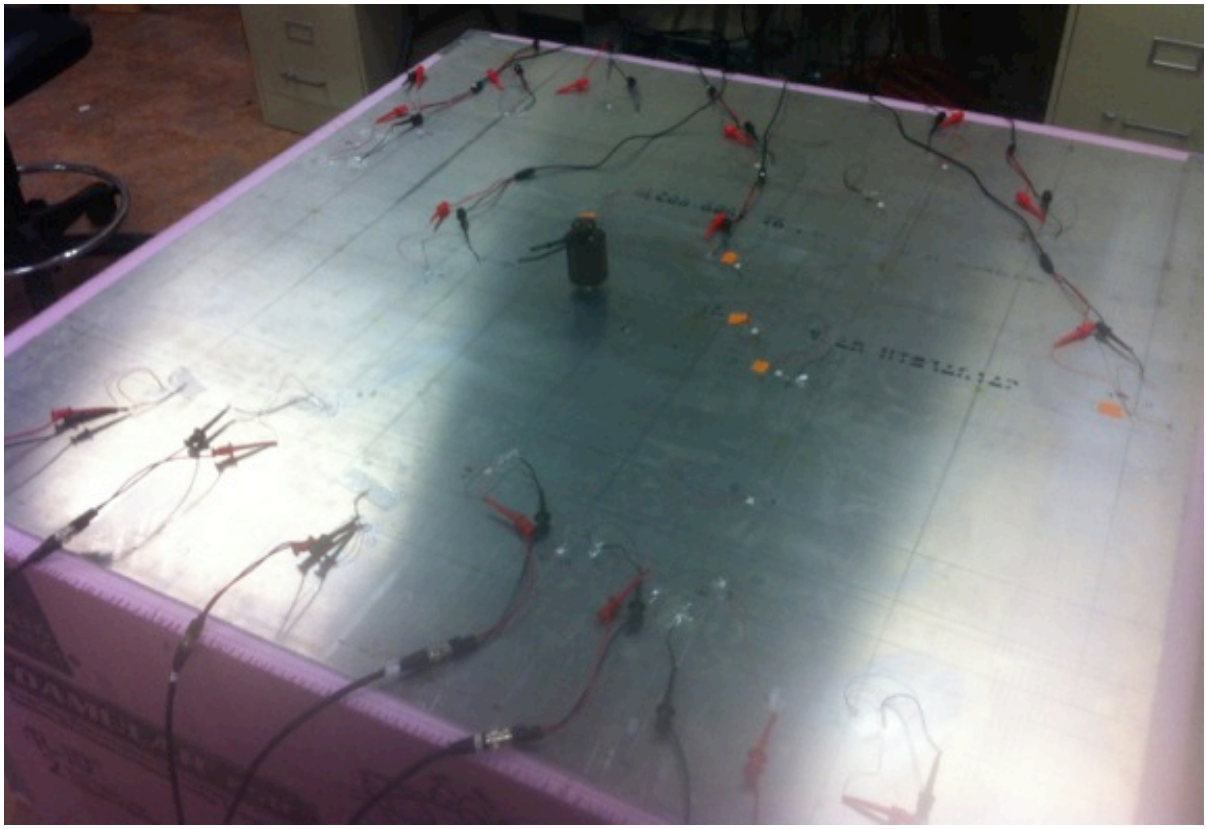


Figure 5-3 The plate specimen was put on foam insulation sheets that serve as a temperature chamber. Sixteen transducers are used to localize the mass scatterer located near the center of the plate.

5.3.1.2 **Data acquisition**

Figure 5-4 shows the NI data collection system in a 'handshake' configuration with 16 transducers to monitor the plate and detect damage. Each of the transducers in turn emits a short ultrasonic signal while 16 transducers record for 4 ms as the wave propagates across the plate. For excitation, we used a 10 V peak-to-peak, 0.1 ms-long windowed-chirp signal with a frequency range of 50 kHz to 500 kHz. The sampling frequency at the receiving channels is 1 MHz, so a 4 ms of record results in 4,000 samples per channel. In our current configuration, roughly five minutes is needed to record all 256 channels (transmitting from each of 16

transducers while receiving from each of 16 transducers). During our continuous monitoring, one set of handshaking data is immediately collected after the previous one is finished.

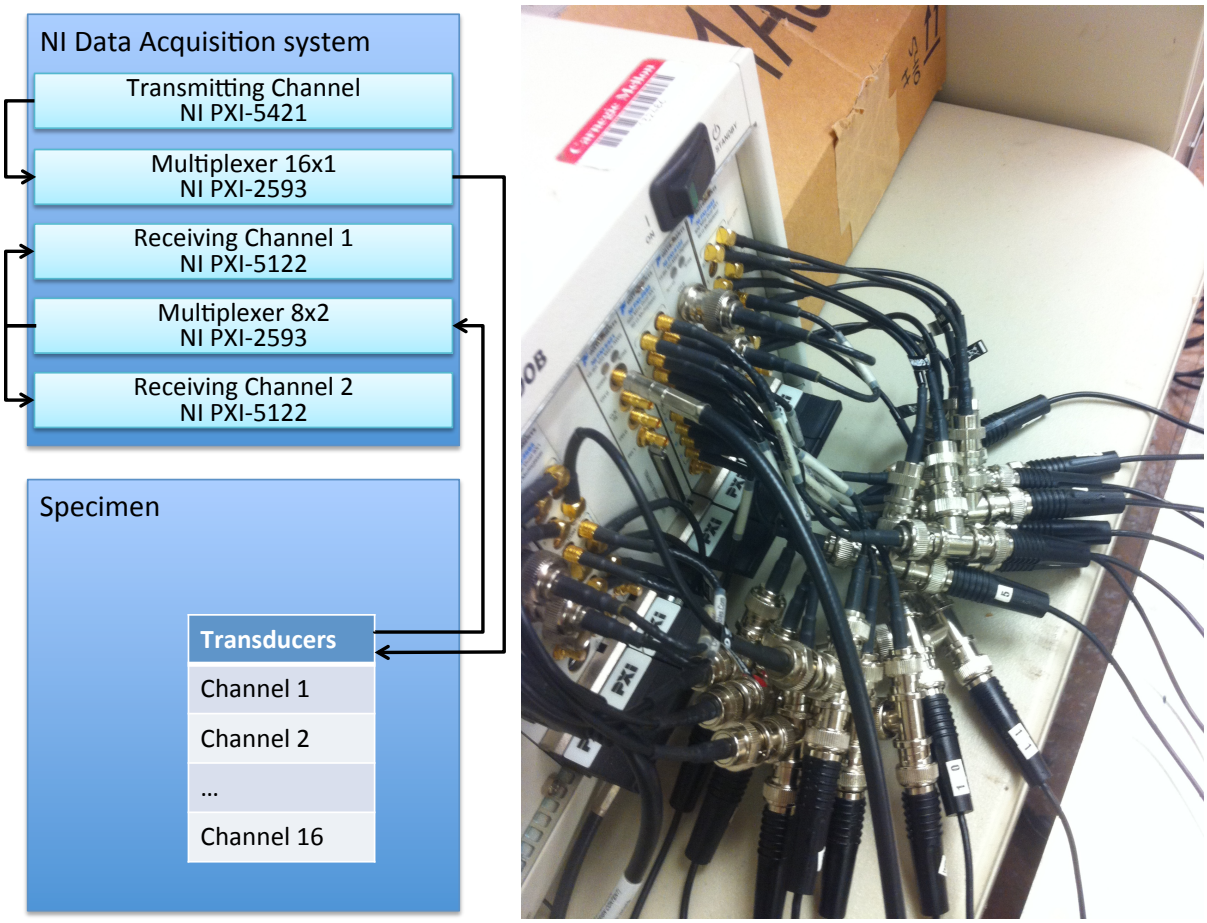


Figure 5-4 Schematic chart and photo of data acquisition system configuration

During the experiment, we periodically toggled a heating fan to create temperature variations in the plate. We placed the heating fan inside a temperature chamber and beneath the plate specimen, as shown in Figure 5-5(a). We pointed the heating fan at a location near the corner of the plate, shown by the triangle in Figure 5-2, to create a temperature gradient across the plate along with the overall temporal temperature variation. The temperature chamber was made from foam insulation sheets to accomplish three experimental purposes:

(1) to support the plate at its boundary, (2) to increase the overall temperature variation, and (3) to slow the cooling phase.

Figure 5-5 (a) A heater is placed inside the temperature chamber to heat the plate at the location indicated in Figure 5-2(a). (b) Temperature measured with hand-held laser IR thermometer at different locations during the heating. (c) A carbon black spray was applied at temperature measurement locations to reduce reflection of the aluminum surface.

Table 3 shows a time sheet summarizing the experimental procedure. We collected 76 sets of handshaking records in 6 hours, during which time we toggled the heating fan ten times, created five heating-cooling cycles. We also show in Table 3 that the first 36 sets of records were collected without the presence of a scatterer. We then attached a scatterer to the plate before we collected the 37th set of records at 14:25.

Table 3 Time log showing environmental parameters in an experiment on a plate specimen

Time (HH:MM)	Heating*	Scatterer	# Records
11:49 – 12:19	Low	Off	7
12:19 – 12:50	Off	Off	14
12:50 – 13:27	High	Off	22
13:27 – 14:16	Off	Off	34
14:16 – 14:25	Low	Off	36

14:25 – 14:53	Low	On	40
14:53 – 15:25	Off	On	43
15:25 – 16:00	High	On	50
16:00 – 16:30	Off	On	58
16:30 – 17:00	Low	On	65
17:00 – 18:00	Off	On	76

* High/Low/Off are the three power levels of the heating fan.

5.3.1.3 **Spatial and temporal temperature variations**

We measured the temperature at five different locations during one representative 45-minute 'high' level heating phase to quantify the non-uniformity of temperature variation across the plate. We measured the temperature with a hand-held laser IR thermometer, as shown in Figure 5-5(b). The five locations where temperature was measured are shown in Figure 5-2 with diamond symbols, and are selected such that they are of different distances from the heating point, which is shown by the triangle at the bottom-right corner of the plate. We applied carbon black spray (shown in Figure 5-5(c)) to the five measurement locations to reduce reflection of the aluminum surface to improve the accuracy of the IR thermometer measurements. Figure 5-6 plots temperature in degree Celsius against time and against the distance between the measurement point and the heating location. The variations along two axes are, respectively, temporal and spatial. Figure 5-6 shows that the temperature at every point started at 20 ℃, which is the room air temperature. During the 45-minute heating phase we observe the overall temperature to increase with time, and to increase more rapidly for locations closer to the heating point. The final temperature ranges from 24 ℃ to 39 ℃. Subtracting the room temperature at the beginning of the record, which is about 20 ℃, the temperature change on the plate due to the non-uniform heating ranges from 4 ℃ to 19 ℃.

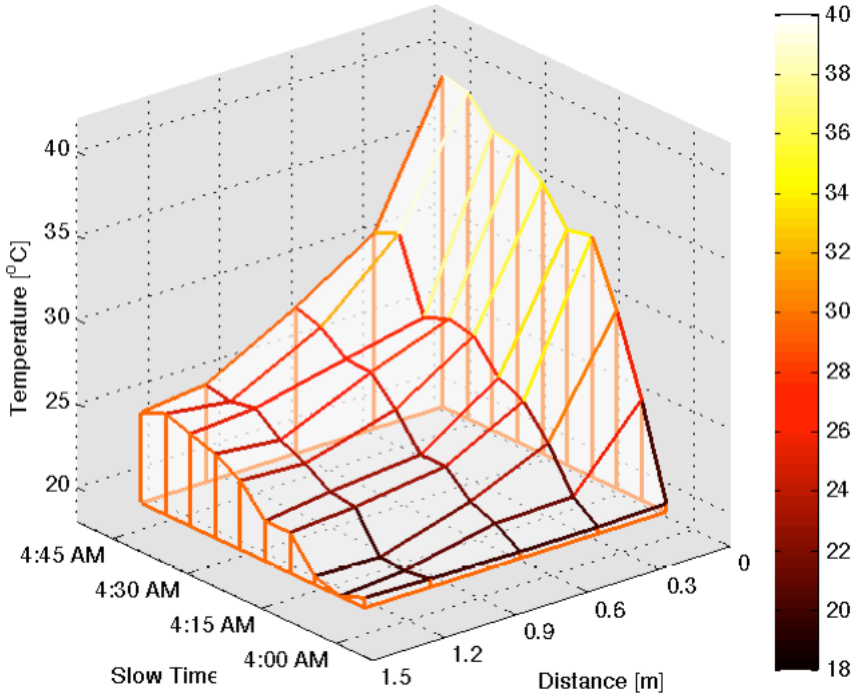


Figure 5-6. Temperature on the plate during a 45-min heating phase is a function of both time and distance from heating point.

We physically simulated a localized defect on the plate with a grease-coupled mass scatterer, indicated by a circle in Figure 5-2. We applied the mass scatterer after collecting 36 sets of records, after which we collected another 41 sets of records, as shown in Table 3. The mass scatterer creates only a small change that is difficult to observe even when environmental and operational variations are absent.

In order to quantify the change produced by the scatterer, we first collected two sets of records in an 'ideal' laboratory environment at room temperature: one set on the "intact" specimen and the second set in the presence of the mass scatterer. We performed baseline subtraction on these two sets of records to obtain the scatter signals from each channel. We then calculated the energy of the scatter signals and show them in Figure 5-7 in red circles.

The horizontal axis corresponds to 256 channels, and is separated into two plots for clarity. The energy values are normalized in decibels such that the baseline signal energy is 0 dB. Figure 5-7 shows that the scatter signals are generally 30 to 40 dB weaker than the received signals, which are mainly composed of the direct-path signal.

We also characterize the temperature effect from the experimental dataset we collected when spatial and temporal temperature variations are present, whose parameters were shown in Table 3. We calculate the energy of residual signals after subtracting the baseline signal directly from the temperature-affected records, and show them in blue boxes in Figure 5-7. Because we collected 76 sets of records at varying temperatures, we obtain 76 values of residual energy for each channel. The blue boxes plot the interquartile range (middle fifty) of the residual energy and the vertical-extended lines show the full range. In Figure 5-7 we show that the temperature-residual energy quantities are roughly 5 dB weaker than the received signal, and roughly 30 to 40 dB larger than the scatter signal of interest. The low signal-to-noise ratio for the scatter signal of interest poses a challenge to the localization task.

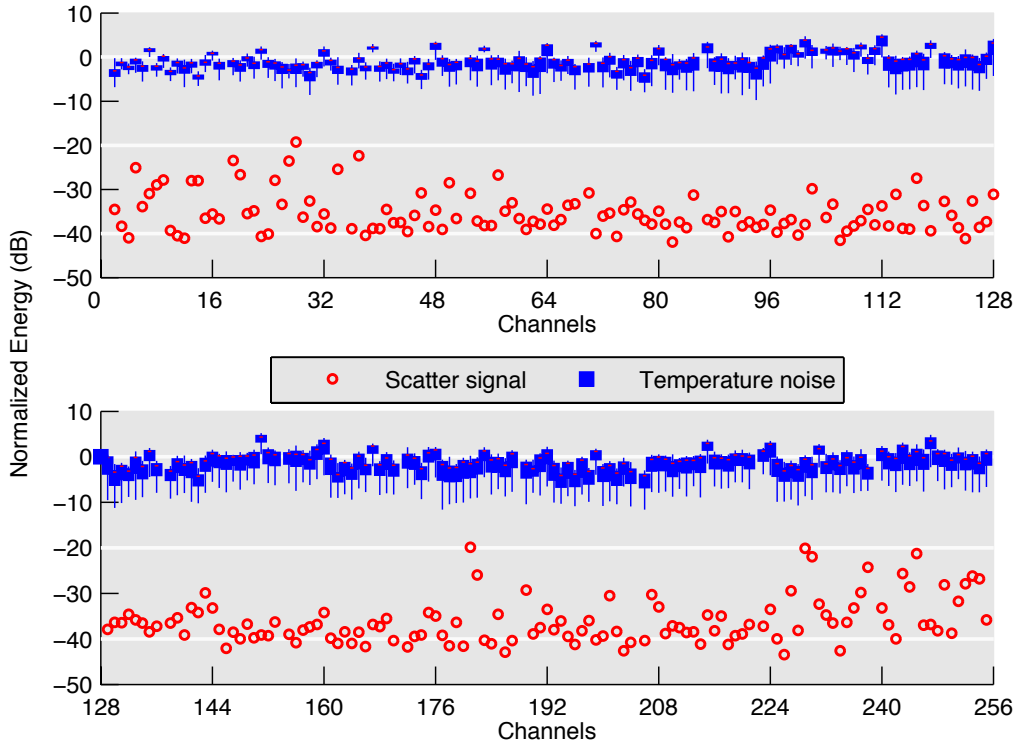


Figure 5-7. Normalized energy of scatter signal at each channel, shown in red circles, obtained from baseline subtraction of two sets of records collected at steady room temperature. Along with normalized energy of residual signal, shown in blue boxes, obtained by subtracting the baseline record without compensating for temperature changes.

Figure 5-8 plots the correlation coefficients of the experimental data and shows not only the effect of temperature variation over slow-time, but also the effect of spatial temperature gradients among different channels (transducer pairs). The horizontal axis denotes the index of records, and the y-axis the correlation coefficients. As shown in equation (2-1), the correlation coefficient of two waveforms measures the similarity between the two waveforms – the correlation coefficient is 1.0 if the two signals are identical. For a batch of N records collected from a particular channel (one transmitter-receiver pair), we obtain a vector of N correlation coefficients, where the i^{th} value is the correlation coefficient between the i^{th}

record and a baseline, indicating the similarity between the two records. In our experiment, we collected 76 pitch-catch records from 256 channels. Therefore, each of 256 curves in Figure 5-8 shows 76 correlation coefficients representing the variations in one channel. Most of the curves drop from 1.0 to around -0.1, indicating that the records are significantly affected by temperature variation. Also, the dissimilarity between the curves implies that temperature variation is non-uniform across the plate, i.e. there also exists spatial temperature gradients. We also see (16) straight (superimposed) lines plotting at 1.0; these correspond to the signals received by each transmitter during its own (self) 'handshake'.

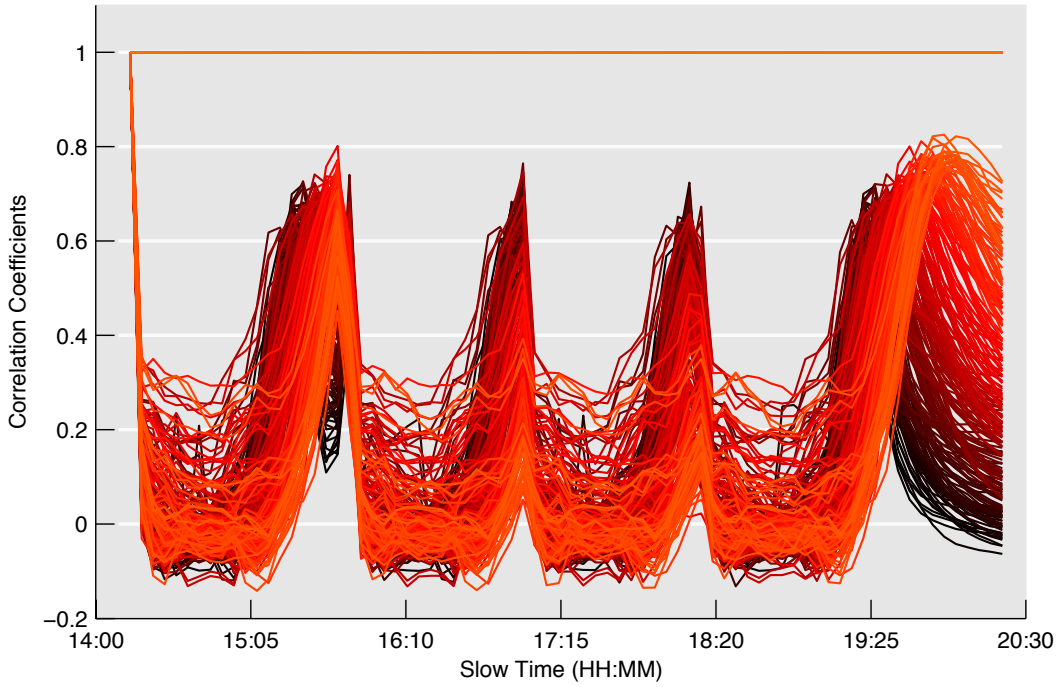


Figure 5-8. Correlation coefficients of pitch-catch records collected from plate specimen show the temperature varying cycle. Each individual line represents the variation in one of the 256 channels.

In the next section, we show that our SVD robust baseline removal method significantly improves the localization result, despite the challenges from complex temperature profiles and from low signal-to-noise ratios.

5.3.2 Experimental results

In this section we show delay-and-sum (Michaels 2008) localization results with scatter signals extracted from SVD robust baseline removal, and compare those results with those obtained using conventional temperature compensation. In a conventional implementation, difference signals are obtained by baseline subtraction. The delay-and-sum implementation then typically discretizes the targeting area into grid points and calculates the distance between the transducers and each of the grid points. For each grid point, the difference signals are enveloped and delayed according to the distance and an estimation of the wave velocity, and are then summed together. The results at the grid points are assembled into an image, where the value of each pixel indicates how likely the scatterer is located at the particular grid point. In our experiments, we collect pitch-catch records from 256 channels, one corresponding to each transmitter-receiver pair. For each channel, the scatter signals are extracted with our SVD robust baseline removal method and compared with a more conventional scale-transform temperature compensation approach (Harley and Moura 2012).

5.3.2.1 Extracted scatter signals

In this subsection, we show examples of extracted scatter signals both from SVD robust baseline removal and from (conventional) scale transform temperature compensation. We extract such scatter signals from each of the 256 channels (transducer-receiver pairs) but for readability we will only show ten example extracted scatter signals, randomly sampled from the results extracted from the 256 channels.

With our SVD robust baseline removal method, we extract scatter signals by first identifying the damage-sensitive left singular vectors. For each of the 256 channels, we process the data with SVD and the damage-sensitive singular vector selection, described in *Section 3.2*, to identify the left singular vector with a step change corresponding to the presence of a new scatterer (damage). Figure 5-9(a) shows the selected left singular vectors from ten of the 256 channels against slow time, which is represented by the index of measured records during monitoring. The grey dashed line in Figure 5-9(a) shows the step at which we introduced the mass scatterer. By comparing the left singular vectors with the ground truth shown in the dashed step line, we can visually identify the step-like behavior that corresponds to the introduction of the scatterer. We also show scores computed using the K-means damage-sensitive feature selection methods for each singular vectors; the scores describe how likely the left singular vector resembles an abrupt change produced by damage, and we use a threshold of 0.5 to determine whether damage exists in a batch of records.

Figure 5-9(b) shows the corresponding right singular vectors from those ten channels. The right singular vectors resemble the scatter signals and show dominant arrivals that we expect to have originated from the damage site. Figure 5-10(a) shows the baseline signals that we subtracted from the temperature-compensated records. Comparing the right singular vectors shown in Figure 5-9(b) and the baseline signals shown in Figure 5-10(a), we see that the dominant arrivals in the baseline signals are absent in the right singular vectors, indicating baseline information is removed in the singular vectors.

Each right singular vector represents the scatter signal extracted from one of the 256 channels. Combining all 256 right singular vectors, we obtain a matrix $\boldsymbol{X}_S$, whose element

$\boldsymbol{x}_{s_{ij}}$ represents the extracted scatter signal received by transducer j while we emit from transducer i.

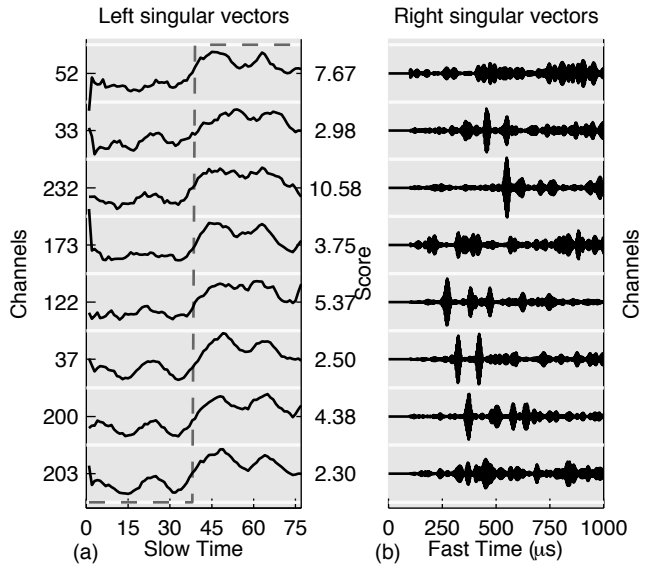


Figure 5-9 Selected left singular vectors and corresponding right singular vectors decomposed from experimental data. Left singular vectors shows steps corresponding to damage. Corresponding right singular vectors approximate scatter signals.

To compare the scatter signals extracted using different methods, we plot in Figure 5-10(b) scatter signals recovered using scale transform temperature compensation and baseline subtraction. In this example, we chose to use a pair of baseline and measurement records such that the temperature difference between them is minimal, and we then use scale transform temperature compensation to further compensate for the small temperature change. Therefore, we are essentially implementing both optimal baseline selection and the stretching based method, which is similar to the implementation of "continuous growth baseline" temperature compensation (Putkis and Croxford 2013b). Because the temperature variation is minimal, the scatter signals shown in Figure 5-10(b) should resemble the scatter signals from those channels. Comparing Figure 5-9 (b) and Figure 5-10(b), we observe good similarity

between the scatter signals recovered by temperature compensation and the right singular vectors, indicating that the right singular vectors well approximate the actual scatter signals. However, the temperature-compensated baseline subtracted signals contains more noise than the right singular vectors, and the arrivals at the expected locations are also less dominant than those in the right singular vectors. As we will show later, these two factors make the localization result using singular vectors better than results using temperature-compensated baseline-subtracted signals, even when the temperature variation is small.

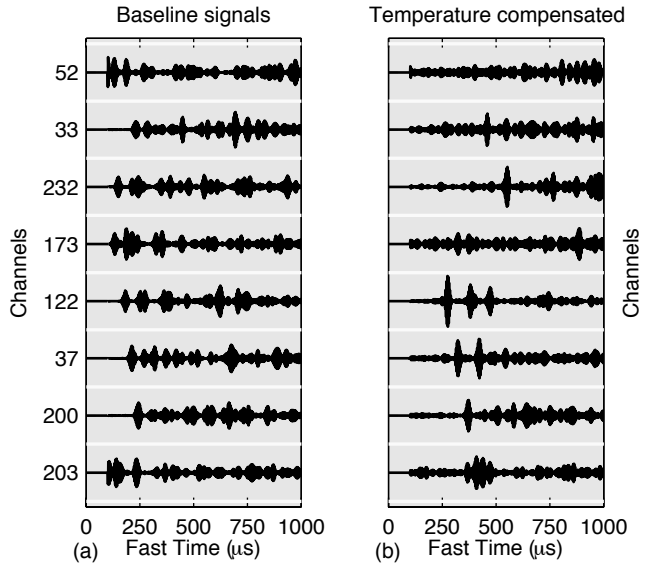


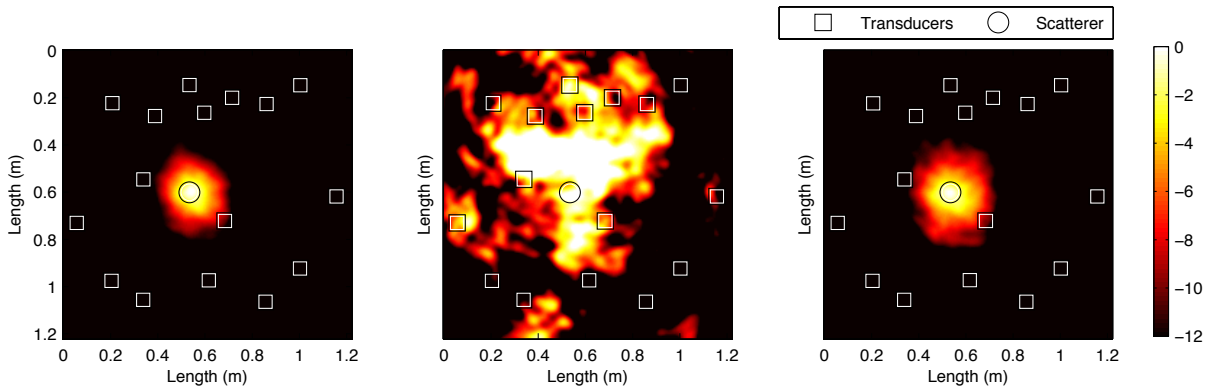
Figure 5-10 Baseline signals and temperature compensated baseline subtracted signals.

5.3.2.2 **Localization results with a spatial temperature gradient**

We now use the extracted scatter signal matrices to localize the scatterer in the presence of temporal temperature variation up to 19 ℃ and a spatial temperature gradient across the plate up to 15 ℃. By utilizing the scatter signal matrices, the location information of the transducers, and an estimation of the wave velocity, we localize the damage using delay-and-sum localization algorithm [9]. Figure 5-11 shows the localization results as color-coded

images with squares indicating the location of transducers, and a circle indicating the location of the mass scatterer. The brightness of each pixel indicates how likely the scatterer is located at that pixel. The values are normalized such that the closer it is to 1 (bright), the more likely the scatterer is located at that pixel. The image is then plotted in a decibel scale from 0 dB to -12 dB.

Figure 5-11(a) shows a localization result with scatter signals extracted by SVD robust baseline removal using the entire dataset containing 76 sets of handshaking records. Figure 5-11 (b) and (c) show a localization result in which signals are first processed with scale transform temperature compensation (Harley and Moura 2012), and then subtracted with the first set of records as a baseline. The difference is that for Figure 5-11(b), the temperature change between the record and the baseline is large (4~19 ℃); for Figure 5-11(c), the temperature between records and baseline is minimal. Note that a sample of scatter signals used to produce Figure 5-11(a) are shown in Figure 5-9(b), and a sample of scatter signals used to produce Figure 5-11(c) are shown in Figure 5-10(b). We did not show the scatter signals used to produce Figure 5-11(b), in which case the recovered scatter signals are visually dominant by residual error from the imperfect subtraction of baselines, because temperature compensation becomes less effective at the presence of large temperature variations and gradients.



(a) SVD Robust baseline removal on all monitoring records

(b) Temperature compensation in the presence of large temperature gradients

(c) Temperature compenstion in the presence of small temperature gradients

Figure 5-11 Localization results using scatter signal extracted using SVD baseline removal and scale transform temperature compensation

We observe visually that Figure 5-11(a), which is produced by scatter signals extracted using SVD robust baseline removal, shows the best result with a clear peak at the damage location indicated by the circle. This shows that the SVD robust baseline removal method is capable of extracting the scatter signal in the presence of large spatial and temporal temperature variations, without using a specific, optimal baseline.

We further compare the localization performance using SVD robust baseline removal and temperature compensation through the following five metrics:

A. **Localization accuracy**. The localization accuracy is defined as the distance between the maximum of the image at coordinates (x_s, y_s) and the true scatter coordinates $(\widehat{x_s}, \widehat{y_s})$. This error characterizes each method's capability of correctly estimating the target's position (Harley and Moura 2014).

$$\boldsymbol{r_{err} = \sqrt{(x_s - \widehat{x_s})^2 + (y_s - \widehat{y_s})^2}.} \tag{5-16}$$

B. **Energy concentration.** We calculated the accumulative energy that falls in the 5-cm diameter region around the scatterer. We then define the percentage of that focused energy to the total energy of the image as the energy concentration.

$$\boldsymbol{E_c} = \sqrt{\frac{\sum_{r<2.5cm} \boldsymbol{P(r)^2}}{\sum_{r\in R} \boldsymbol{P(r)^2}}}, \tag{5-17}$$

where, P(r) is the amplitude of pixels as a function of r, the distance to the true scatter location.

C. Exponential coefficient (Hall and Michaels 2010). If we fit the amplitude-distance function to an exponential function,

$$\boldsymbol{f(r) = e^{-cr}}, \tag{5-18}$$

the exponential parameter c can be estimated with a least-square fit on the logarithm of the amplitude (normalized such that the amplitude at true scatter localization is 1),

$$\hat{c} = arg\min_{c} \sum_{r\in R} \left(ln\left(\frac{P(r)}{P(0)}\right) + cr \right)^2. \tag{5-19}$$

This exponential parameter c is a nice single parameter that generally increases as the damage becoming more focused, and as the noise floor becoming lower. Therefore, the larger c is, the closer the image is to perfect.

D. **Peak-to-artifact ratio (PAR) (Harley and Moura 2014).** If the amplitude-distance function is instead fitted to a Gaussian function with a fixed maximum value of 1, a width controlled by the parameter a, and a minimum controlled by the parameter c,

$$\boldsymbol{f(r) = ce^{-ar^2} + (1 - c)}. \tag{5-20}$$

The peak-to-artifact ratio is then defined as a decibel value that increases as the noise amplitude becomes smaller compared to the peak amplitude, which is 1.

$$PAR = 20 * log_{10}(\frac{1}{1-c}), \tag{5-21}$$

E. Resolution (Harley and Moura 2014). The resolution of the image is defined in (Harley and Moura 2014) as the half-amplitude width of the main peak of the localization result, and is computed using the parameters from the Gaussian fit in equation (5-20),

$$\sigma = 2\sqrt{-\frac{1}{a} log\ (\frac{c-1/2}{c})}. \tag{5-22}$$

In general, a smaller σ indicates less ambiguity regarding the location of the target, and is more desirable.

The three metrics C, D, and E from (Harley and Moura 2014) and (Hall and Michaels 2010) require fitting the amplitude-distance function to different models. Figure 5-12 plots the experimental result in dots and the two fitted functions in dashed (exponential) and dash-dotted (Gaussian) curves. Both the exponential and the Gaussian models fit the experimental result well – the R square value of the two least-square fits are 0.94 and 0.89, respectively. We therefore evaluate metrics C, D, and E from the two models equally.

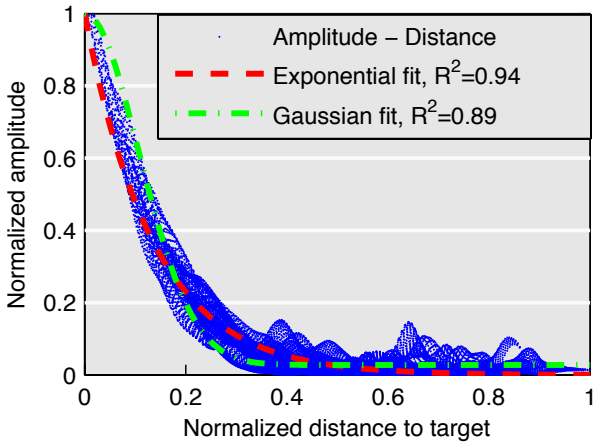


Figure 5-12 Normalized amplitude as a function of normalized distance from Figure 5-11, fitted with exponential function and Gaussian function.

Tables 2 summarized the performance metrics to compare the localization results using SVD robust baseline removal and that using temperature compensation. Our SVD robust baseline removal method performs best in four of the five metrics, confirming our previous observation that it produced best result in the presence of temperature variations and gradients. The accuracy (distance from maximum to true scatter location) of our method is slightly worse than the temperature compensation result, but the difference (0.66cm) is much smaller than the scatter diameter (5cm) and can be contributed to the difficulty in measuring the actual contact center of the scatterer to the specimen.

Table 4 Comparing localization results on a plate from SVD and temperature compensation using five performance metrics

Method/test	Better if the value is	**SVD robust baseline removal**	**Scale transform temperature compensation**	
		On sequential records	With large temp. variation	With small temp. variation
Corresponding localization image		Figure 5-11(a)	Figure 5-11(b)	Figure 5-11(c)
Accuracy [cm]	Smaller	0.78	14.58	0.12
Energy concentration [%]	Larger	24.58%	7%	20.09%

Gaussian	PAR [dB]	Larger	31.2	19.5	28.7
	Resolution [cm]	Smaller	23.8	66.1	37.5
Exponential	Coefficient [a.u.]	Larger	7.96	2.47	6.82

Figure 5-11(b) shows no indication of damage at the designated location, and expected poor performance in terms of the metrics, which demonstrates a limitation of stretch-based temperature compensation methods in the presence of large spatially and temporally temperature variations. Stretch-based temperature compensation fails because when a large spatial temperature gradient exists, the temperature effect is different for the direct path and scatter path. Therefore, it is impossible to accurately estimate the stretching. By stretching the records with inaccurate estimates of the stretching factors, the recovered scatter signals cannot resemble the actual scatter signal, and thus the localization fails.

On the other hand, when a small temperature change exists, as it is the case in Figure 5-11(c), the stretch-based localization also correctly indicates the true damage location, yet its performance is not as good as that from SVD baseline removal from the metrics. Note that the measurement records were collected under minimal temperature difference with the baseline and then processed with scale transform temperature compensation to further compensate for the small temperature change, similar to the process in (Putkis and Croxford 2013b). Therefore, the stretch-based baseline removal achieves its optimal performance, and the recovered scatter signals should be very close to the actual scatter signals. However, even with this optimal implementation, the energy concentration using is still 4% smaller than that of SVD robust baseline removal, and the performance is also worse than our SVD method using the three metrics from the literature.

SVD robust baseline removal produces better results because it uses the information of the entire dataset, in which the records were collected at different temperatures. Due to the

orthogonality of SVD, the extracted signal is orthogonal to the temperature variations exhibited in all other singular vectors that are associated with temperature stretching. Therefore the extracted scatter signals are minimally affected by the effects of temperature variation. Also, white noise and other transient variations are exhibited in singular vectors with higher indices, and are therefore excluded from the damage-sensitive singular vector. As a result, SVD also helps improve the localization result as it eliminates the white noise over the dataset.

5.4 An Additional Experiment On A Pipe Specimen

We conducted a similar experiment on an aluminum pipe specimen subjected to temperature variations with spatial gradients. The first goal of this pipe experiment is to further examine the robustness of our SVD baseline removal method. Guided waves in pipes are more complicated than Lamb waves in plates, because the wave can propagate through helical paths. Despite this complexity, our experimental data shows that the orthogonality assumption is still approximately sound. The second goal of this pipe experiment is to generalize our baseline removal method to other applications.

We created temperature variation and gradients by placing one end of the pipe segment in an ice water bath and heating the other end with a heater. We collect pitch-catch records from 24 transducers in a short time interval, and attached a grease-coupled mass scatterer to the pipe at some time during the monitoring period to simulate localized damage. We then compare the localization performance of our SVD robust baseline removal method to conventional temperature compensation methods.

5.4.1 Experimental setup

Figure 5-13 shows the schematic drawing of the experimental setup, using an aluminum pipe specimen with 89 mm diameter, 1320 mm length, and 3 mm wall thickness. We attached 24 PZT wafer transducers (7x6x1 mm) in three different arrays, each containing 8 transducers. Two arrays are circumferential, located 228 mm from each pipe end. The remaining 8 transducers form a linear array in the axial (longitudinal) direction.

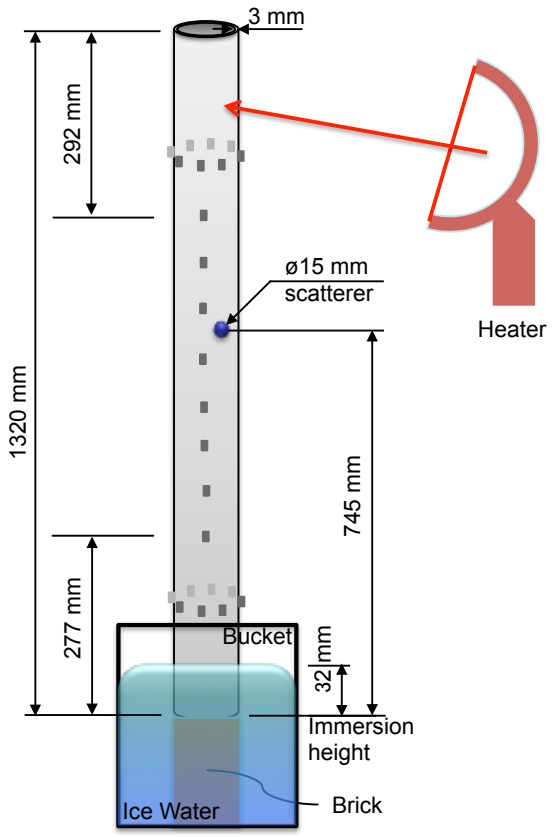


Figure 5-13 Experimental setup: placing a pipe segment in an ice water bath to create spatial temperature gradients to test our SVD robust baseline removal

During our experiments, we created temporal temperature variations and spatial temperature gradients in 30 min cycles. The time log in Table 5 shows the 5-hr experimental procedure. We first immersed the pipe in ice water for 30 minutes, allowing the pipe to cool

from the bottom up; then removed the pipe from the ice water and heated the top end of the pipe with a heater for another 30 minutes. We repeated this process several times while we collected pitch-catch records from all of the 24 transducers. We then exposed the full specimen to room temperature and collected records while the temperature of the pipe stabilized.

Table 5 Time log showing environmental parameters in an experiment on a pipe specimen

Phase #	**Condition**	**Time**
Phase 1	Cooling	2:30 – 3:00 PM
Phase 2	Heating	3:00 – 3:30 PM
Phase 3	Cooling	3:30 – 4:00 PM *Mass scatterer applied at 3:44*
Phase 4	Heating	4:00 – 4:30 PM
Phase 5	Cooling	4:30 – 5:00 PM
Phase 6	Room Temp.	5:00 – 7:30 PM

At the middle of phase 3 (03:44 PM), we placed the mass scatterer on the pipe. We also measured the temperature at five locations along the pipe for 2.5 hours (phase 1-5), and show the measurements in Figure 5-14.

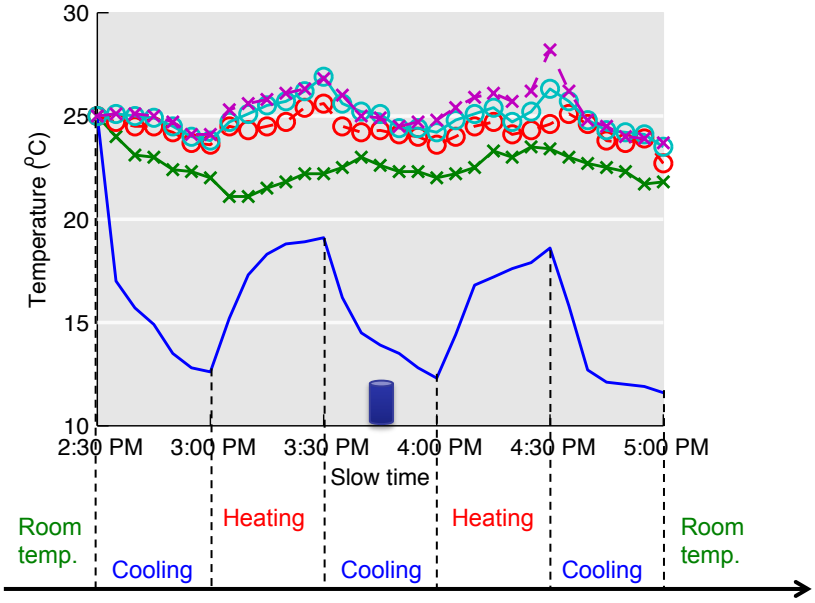


Figure 5-14 Temperature measurements at five locations along the pipe for 2.5 hours (5 cooling/hearing phases)

We collected pitch-catch records from the 24 transducers at 4.3-min intervals and obtained a total of 71 sets of records in the 5-hr monitoring period. Figure 5-15 plots the correlation coefficients (defined in equation (2-1)) of the experimental data and shows the effect of temperature variation on pitch-catch records. For clarity, the baseline records used to calculate the correlation coefficients are the set collected at 3:31 AM, at the conclusion of the first cooling-heating cycle. In this case, the variation of correlation coefficients corresponds well with the temperature measurements shown in Figure 5-14.

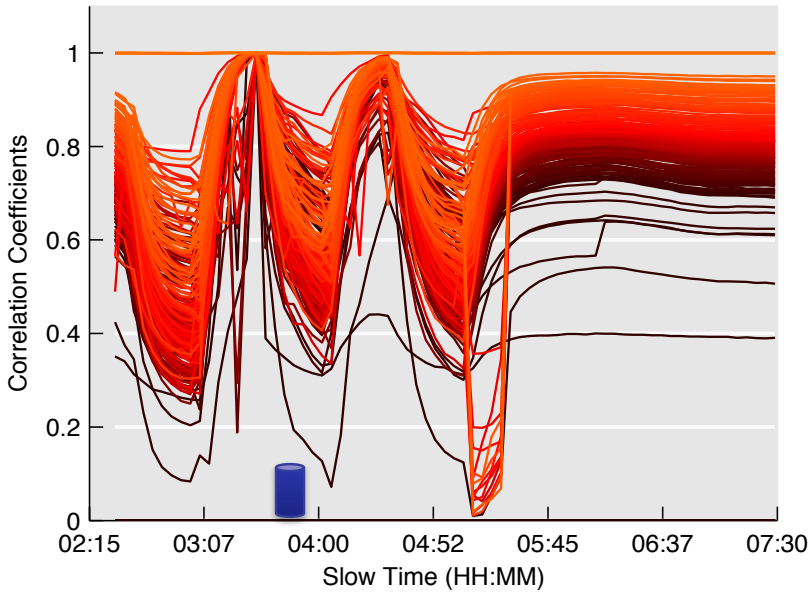


Figure 5-15 Correlation coefficients of pitch-catch records collected from plate specimen show the temperature varying cycle. Each individual line represents the variation in one of the 256 channels.

5.4.2 Experimental result

In this section, we demonstrate detecting and localizing a small scatterer on the pipe by processing the 5-hr monitoring records with SVD. We analyze the pitch-catch records from different transducer arrays separately. First, we process the pitch-catch records from the two circumferential arrays to detect the presence of the scatterer. We then analyze the handshaking records from the axial array, to accurately localize the scatterer. We also show that by using the helical-path information contained in the long received signal, we can effectively diminish the background noise and significantly improve the localization accuracy.

5.4.2.1 Damage detection using left singular vectors

We first show the detection results using pitch-catch records from the two circumferential arrays near the pipe end, and then show the localization result using handshaking records

from the axial array. Figure 5-16 shows left singular vectors decomposed from pitch-catch records from the circumferential array. The selected left singular vectors (solid lines) from different channels show clear jumps at the time the scatterer was placed on the pipe (black-dashed step line).

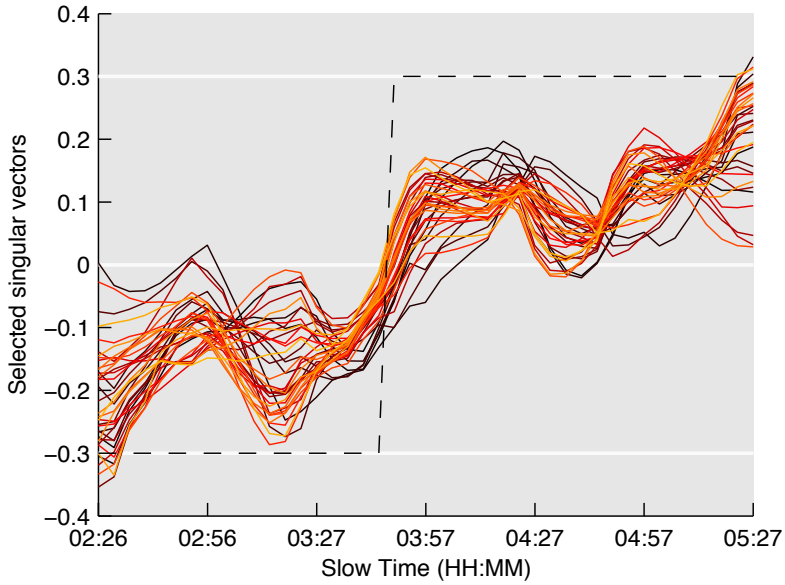


Figure 5-16 Select singular vectors (solid curves) corresponds well with the step change caused by scatterer (black dashed line)

We performed classification on the left singular vectors using a threshold classifier at zero. Figure 5-17 plots the same singular vectors shown in Figure 5-16 in two-color, scatter points. All points before the placement of the scatter (indicating intact records) are plotted in green crosses, and all points after the event (indicating damaged records) are plotted in red crosses. We then assigned label to the points using a threshold at 0: if the value of a point is smaller than 0, the point is labeled **intact**; if the value of a point is larger than 0, the point is labeled **damaged**, and indicated by a blue circle. Therefore, a green cross with a blue circle

indicates a false positive result; a red cross without a blue circle indicates a false negative result; the rest two combinations indicate correct classification results.

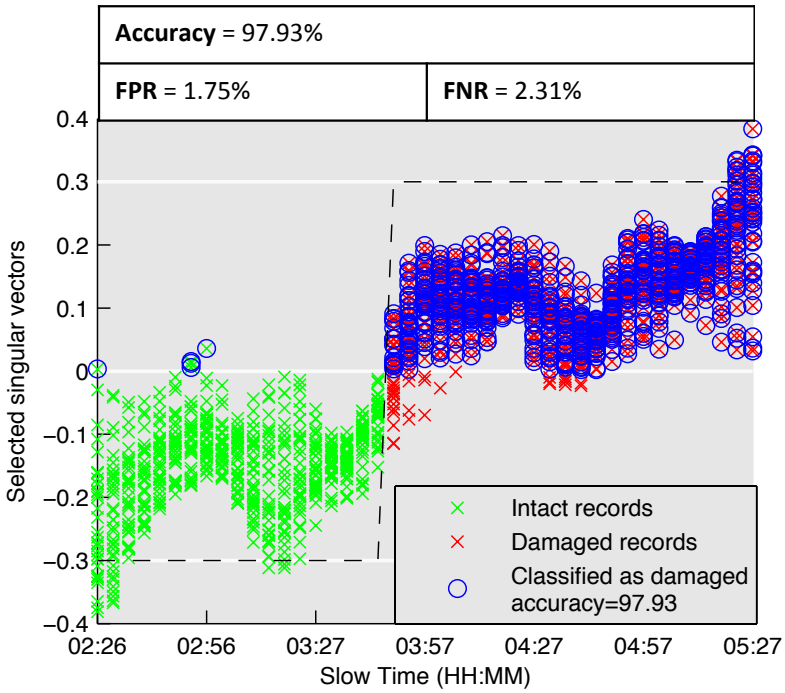


Figure 5-17 Classification results using left singular vectors with a simple threshold at zero.

By counting the different symbols in Figure 5-17, we observed excellent classification accuracy (97.93%), low false-positive rate (1.75%), and low false-negative rate (2.31%), confirming the effectiveness of the left singular vectors as damage indicators. Also, the left singular vectors correctly indicated the time point (03:44) when the mass scatterer is introduced, showing good accuracy in the slow-time domain.

5.4.2.2 **Damage localization using right singular vectors**

We then extracted the scatter signals (the right singular vectors) from the handshaking data from the axial array to localize the scatterer. The handshaking of the 8 transducers produced 64 channels of records, and 71 sets of handshaking records were collected from

each of the channels. We used SVD robust baseline removal to extract 64 scatter signals and used delay-and-sum localization algorithm to localize the scatterer. Figure 5-18 shows the delay-and-sum localization result as an image of an unwrapped pipe. The horizontal axis corresponds to the length of the pipe, and the vertical axis corresponds to the circumference of the pipe. The crosses indicate the transducers and the circle near the bottom center of the image shows the actual location of the scatterer. The image is normalized such that the maximum value is 1 (red) and the minimum value is 0 (blue), as also indicated by the color bar.

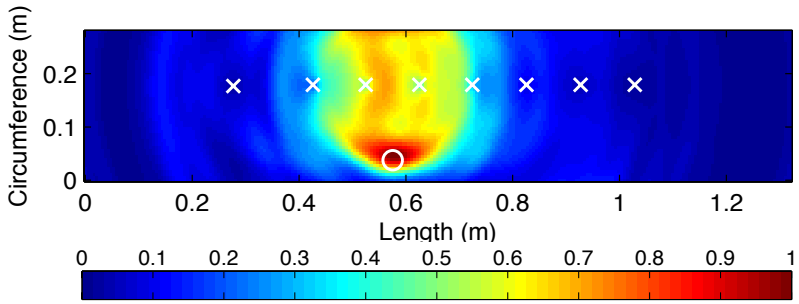


Figure 5-18 Delay-and-sum localization image using right singular vectors decomposed from monitoring records

We observed a major peak in Figure 5-18 at the actual scatter location, which indicates that we successfully extracted scatter signals using singular value decomposition. We will later compare our localization result using SVD with those using conventional temperature compensation methods. But before that, we first show that we can significantly improve the localization performance by utilizing the helical-path information contained in long received signals.

5.4.2.3 **Improved localization result using helical-path waves**

Because of the closed curved geometry of pipe structures, guided waves in pipes can propagate along helical paths. For example, a wave propagating in a pipe may travel directly between two points or may circle around the pipe's circumference multiple times before arriving at a destination, resulting in many wave arrivals in a long received record. Generally, many techniques focused on analyzing only the first arrival (Demma et al. 2004; Michaels 2008), which eliminated the effect of multipath while compromising circumferential resolution. For example, in our implementation shown in Figure 5-18, we unwrapped the pipe and approximate the hollowed cylinder-geometry to a plate-geometry. As a consequence, we broke the circumferential continuity of the pipe, and discarded the helical-path waves that may also interact with the scatterer.

To utilize the helical-path waves we introduce the concept of 'virtual pipes', i.e. we virtualize multiple unwrapped pipes and transducers, and shift them from their originate location by integer numbers of circumferences. The more circumferences we shift the virtual pipe, the higher order of helical-path we can calculate. Similar ideas that utilize virtualized path to obtain the multi-path information have been investigated in literature, such as the use of boundary reflections (referred to as the 'mirror path') to improve localization resolution, and the use of helical-paths to improve circumferential resolution of localizing active sources on a pipe (Leonard and Hinders 2003; Michaels et al. 2012; Harley et al. 2013).

In our implementation, we use helical-path waves with the ellipse method to enhance the performance of localizing a scatterer on a pipe. After virtualizing the meshed grids and transducers on the virtual pipes, we can obtain the propagation distance of the helical-path waves by computing the distance between virtualized transmitter, the meshed grids, and the

actual receivers. Figure 5-19 shows one example 'virtual' path as a solid arrow that is equivalent to the first-order helical path as a dashed arrow. The seven images in Figure 5-19 show localization results using different orders of helical-path waves, where each image shows indication of the scatterer at the correct location. We then combine the results from different virtual pipes to improve the localization performance, especially the circumferential resolution.

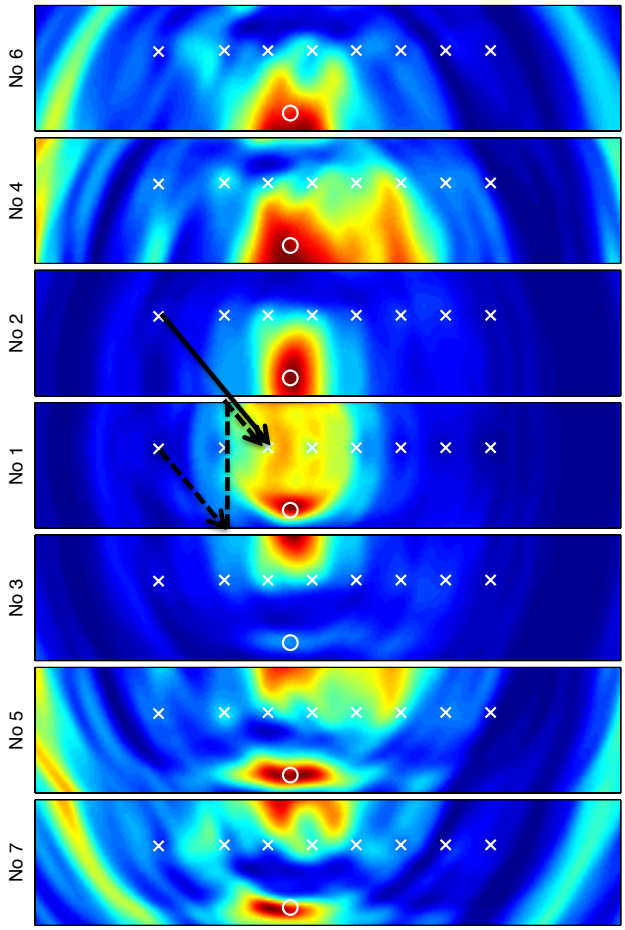


Figure 5-19 Localization results on seven virtual pipes. The helical path in dashed arrow is equivalent to the virtualized path in solid arrow.

To combine the results from the virtual pipes, we can either sum the images or multiply them. Figure 5-20 shows the results of summing or multiplying different number of virtual

pipe results. Multiplying localization images from seven virtual pipes, we obtain the best localization result with a single peak and a clean background. Comparing different rows of Figure 5-20, we observed that the localization performance increases with the number of virtual pipes used. When seven virtual pipe images were combined, we achieved a good circumferential resolution, that the peak in the combined image precisely represents the actual scatter location along the circumference.

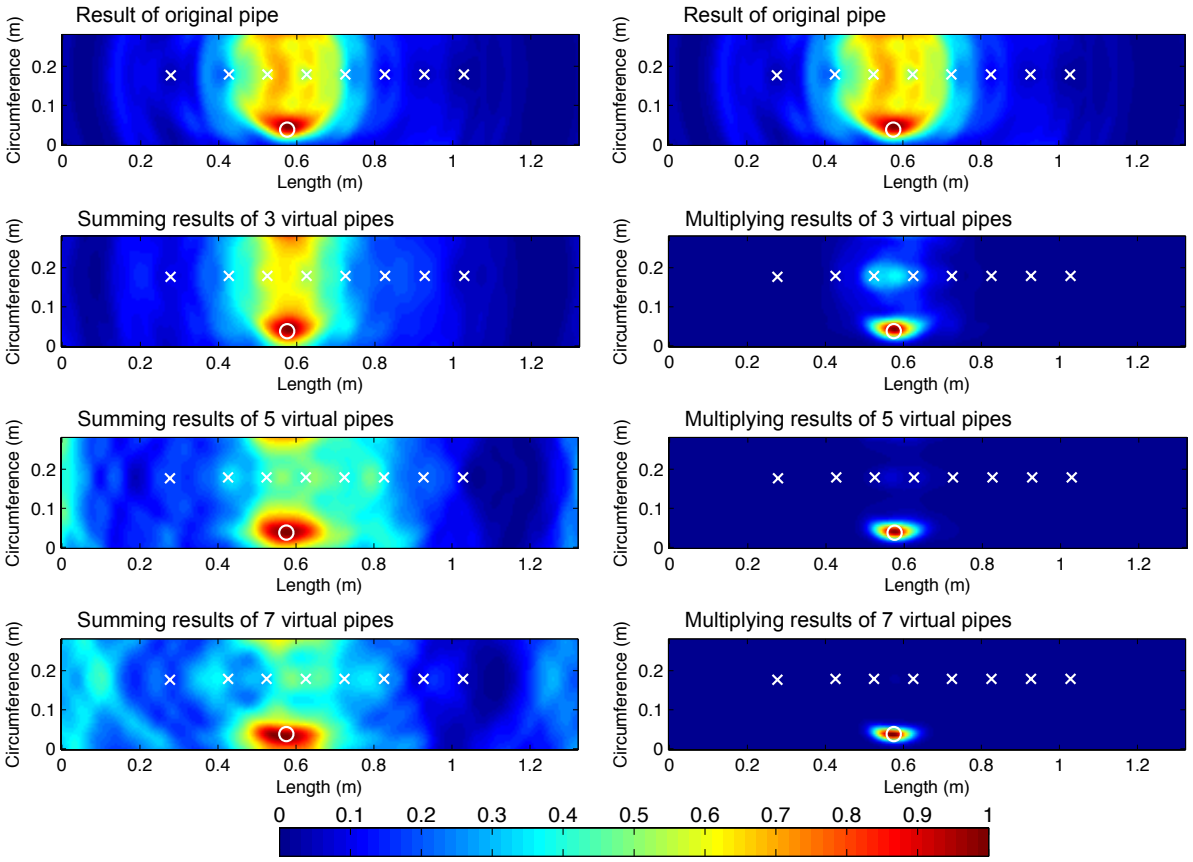


Figure 5-20 Combining localization results from virtual pipes by summing or multiplying different localization images.

Comparing the two columns of Figure 5-20, we observe that we obtain better result when we multiply the images from virtual pipes. Because the images are all strictly positive, the summing approach may emphasize large value that do not overlap across every environment,

and cannot effectively diminish the background noise. If we normalize the images such that the sum of the pixel values is 1, the images become two-dimensional probability distribution functions of damage location given windowed records. In that sense, multiplying the images is essentially same as computing the product of multiple PDFs, which represents the probability that each set of windowed signals correctly mapped the probability distribution of the damage location. This approach significantly helps in emphasizing the largest peak and diminishing the background noise, at the risk of losing secondary (or additional) targets.

In Figure 5-21, we compare localization results obtained using SVD robust baseline removal and scale transform temperature compensation, with the same number of orders of helical paths. We can clearly see that our SVD method produced the most accurate localization result, with a single peak at the scatter location, and a clean background. In the presence of a small uniform temperature variation, scale transform temperature compensation also produces a decent localization result (Figure 5-21 (b)). However, the compensation becomes less effective in the presence of temperature gradients (Figure 5-21 (c, d)). In Figure 5-21(c), we see the peak skewed to the right of the target, which may be due to imperfect compensation for the temperature gradient. The right end of the pipe was placed in the ice water and was therefore colder than the left side. Therefore, the waves propagated slower on the right side. Temperature compensation can effectively compensate for the overall temperature variation, but may likely fail to compensate for the gradients, thereby producing a peak that is skewed to the right. In Figure 5-21(d), where the measured signal is significantly distorted by the large temperature variation and gradients, the temperature compensation approach cannot effectively extract the scatter signal, and the localization fails.

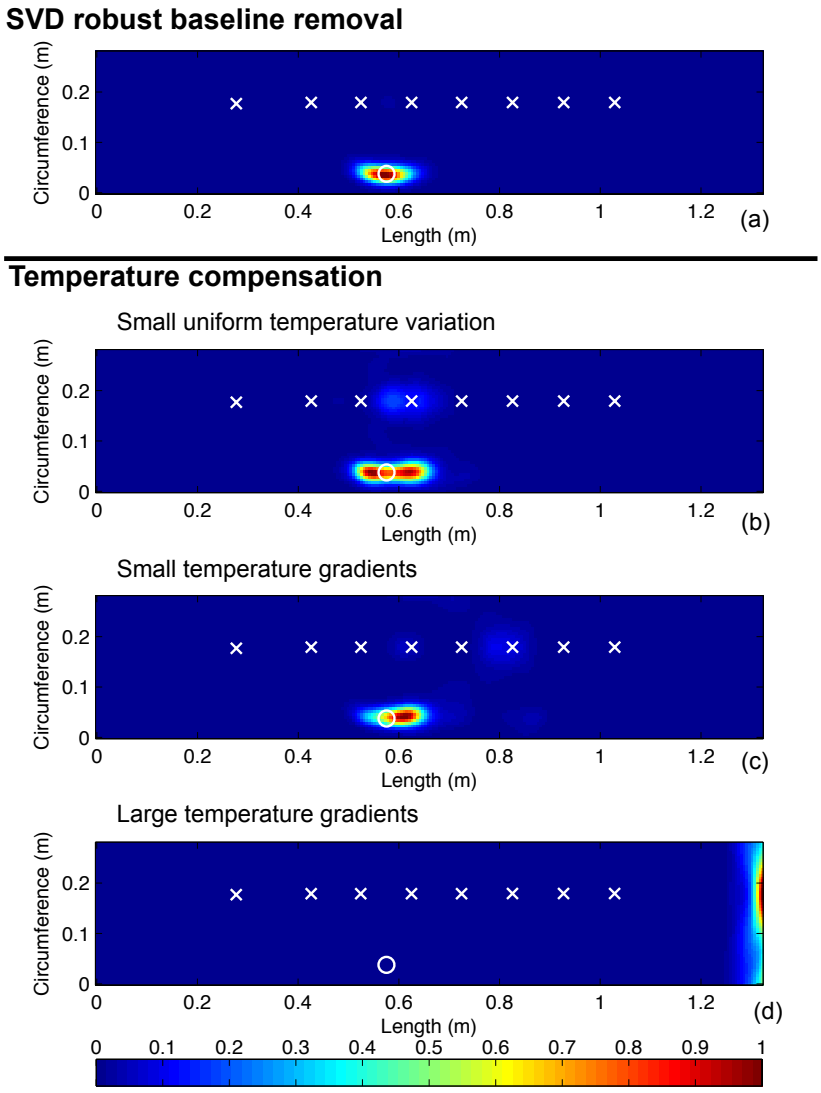


Figure 5-21 Images of pipe localization results using scatter signals extracted using: (a) SVD robust baseline removal; (b) temperature compensation in the presence of small uniform temperature variation; (c) temperature compensation in the presence of small temperature gradients; (d) temperature compensation in the presence of large temperature gradients

Table 6 summarizes the localization performance of the four images in Figure 5-21 using the five metrics previously adopted in *Section 5.3.2*. We observe that the SVD method produce the most accurate result in terms of distance from peak to scatter location, and the cleanest result in terms of the energy concentration near scatter location. Also note that we

obtained satisfactory resolution for the localization image because of the helical path information.

Table 6 Comparing localization results on a pipe from SVD and temperature compensation using five performance metrics

Method/test		SVD baseline removal	Scale transform temperature compensation		
		On sequential records	With small uniform temp. variations	With small temp. gradients	With large temp. gradients
Corresponding localization image		Figure 5-21(a)	Figure 5-21(b)	Figure 5-21(c)	Figure 5-21(d)
Accuracy	[cm]	0.34	3.46	2.92	75.8
Energy concentration	[%]	64.9%	33.8%	30.4%	0%
Gaussian PAR	[dB]	45.1	42.8	34.1	--
Resolution	[cm]	4.46	4.26	34.1	--
Exponential Coefficient	[a.u.]	55.1	42.8	42.1	--

5.5 Summary

Guided wave based damage detection and localization is attractive as a cost-efficient, long-term structural health monitoring technique. However, environmental and operational variations, including temporal temperature variations and spatial temperature gradients, remain challenges for practical implementation of guided wave damage diagnosis systems. Although various temperature compensation methods have been developed to target this problem, they do not successfully remove the effects of temperature gradients on measured data.

In this chapter, we demonstrated singular value decomposition (SVD) as a robust baseline removal method. We validated the method with laboratory experiments on aluminum plate and pipe specimens undergoing spatial and temporal temperature gradients.

We showed that the method could effectively extract a scatter signal from a batch of pitch-catch records. The localization result obtained using the extracted scatter signal is more robust than that obtained using temperature compensation technique.

In addition to the robustness to temperature variation and temperature gradient, our SVD method has other advantages. First, it helps reduce the noise contained in the ultrasonic records, thereby increasing the signal-to-noise ratio and improving the localization results. Second, SVD can be calculated iteratively as new records are taken, and is therefore suitable for long-term online monitoring. Third, the method does not require an explicit baseline(s) to compensate for temperature variation; instead, it identifies from a batch of records the common features of the baseline signal to perform temperature compensation and baseline removal (Liu et al. 2013b)(Liu et al. 2014).

In *Chapter 3*, we introduced our singular value decomposition framework, and in *Chapter 4* and *Chapter 5*, respectively, we showed successful applications of the framework in robust damage detection and robust baseline removal for damage localization. We have shown experimentally that our method is effective in damage detection and localization, and is robust to environmental and operational variations. However, real implementation of guided wave monitoring will likely continue for long periods of time and would produce huge volumes of raw data. It may be impractical to store all records and to process that growing volume of data whenever the structural status is to be determined. In the next chapter, we further expand our damage diagnosis framework by incorporating a sliding window adaptive singular value decomposition algorithm, which updates the singular vectors when each new record is collected. We will show that we can detect and localize damage

efficiently in an online manner, which would be more practical for real implementation of a guided wave monitoring system.

Chapter 6 Adaptive Update Of Singular Vectors For Online Monitoring

In this chapter, we extend our SVD framework for online monitoring and damage diagnosis. We integrate a sliding window adaptive SVD algorithm such that we can adaptively update the singular vectors when each new monitoring record is collected. At any particular time, we can use the updated singular vectors to detect and/or localize damage, as we have demonstrated in *Chapter 4* and *Chapter 5*.

In *Section 6.1* we first discuss the importance and necessity to adaptively update the singular vectors; by doing so, we provide a framework that extracts, stores and utilizes all useful historical information in an efficient and effective way. In *Section 6.2*, we illustrate the methodology of the sliding window adaptive algorithm. We then show two examples implementing online damage diagnosis on experimental data. In *Section 6.3*, we show online damage detection using the field data described earlier in *Chapter 2*. Finally, in *Section 6.4*,

we implement online damage detection using the experimental data collected from the plate experiments described earlier in *Section 5.3*.

6.1 Motivation

In principle, a major advantage of structural health monitoring is to be aware of a structure's condition by taking frequent measurements (Balageas 2006).

In the previous chapters, we introduced our singular value decomposition framework, which utilizes all the historical information contained in the monitoring records. Instead of comparing a new measurement record with a baseline record, our SVD method decomposes information contained in historical records, from which we extract damage information that is robust to other variations. Because singular value decomposition is computed with efficient calculation routines such as LAPACK (Anderson et al. 1999), extracting damage information from a modest number of records is usually very efficient. However, long-term structural health monitoring will generate a huge volume of data. In consequence, it is probably impractical to process all of those monitoring records. Similarly, it is probably prohibitive to store all historical information in terms of ultrasonic records.

We have shown in previous chapters that the singular vectors computed from singular value decomposition can be used to detect or localize damage and are robust to environmental and operational variations in real pipe systems. Moreover, the singular vectors and singular values together preserve all the information in records that are factorized, which means that we can store historical information compactly. For example, a batch of 5,000 ultrasonic records of 10,000 samples consists of a total of 50 million double-precision float-point numbers, totaling about 380 MB. Considering the necessary overhead for file format

and data structure, that amount of data in LabVIEW .lvm files would require about 1 GB of storage. In contrast, if we compute the singular value decomposition of that record matrix, we can efficiently represent the matrix with a few singular vectors with little loss of information. For example, using ultrasonic we collected from the field experiments described in *Chapter 2*, the reconstructed data matrix with 30 singular vector terms occupies less than 2.5 MB of storage. This singular vector representation not only saves storage space, but should also make it possible to implement the method in-situ with simple processors to determine the structural status.

However, in the implementation of the guided wave monitoring system, it would be impractical to recalculate the singular vector matrix of all records every time we seek to assess the structural status. Rather, we propose here to adaptively update the singular vectors as we add new records.

In this chapter, we further expand our damage diagnosis framework by integrating a sliding window adaptive singular value decomposition algorithm (Badeau et al. 2004), which adaptively updates the singular vectors when each new record is collected. We show that we can therefore detect and localize damage efficiently online.

6.2 Methodology

We have shown in previous sections that we can extract damage-sensitive features from ultrasonic pitch-catch records using SVD. However, performing SVD on all data collected in ultrasonic monitoring is not always practical because the computation time increases exponentially as the dataset grows. We now show that by implementing an adaptive method to compute the SVD of ultrasonic records during monitoring in an online manner, we can

process the incoming data at each time step, thereby reducing both the computational and the memory requirements.

The adaptive update of singular vectors can be achieved by subspace tracking methods, (Comon and Golub 1990). In order to monitor damage-sensitive features from ultrasonic records, we must track the full SVD of all received data while obtaining new records at each time step. Therefore, we implemented sliding window adaptive SVD (Badeau et al. 2004). Sliding window adaptive SVD calculates the r dominant singular values and vectors of a data matrix $\boldsymbol{X}$. The *r* dominant singular values are defined as the r largest singular values, and the r left/right singular vectors are the corresponding r columns of $\boldsymbol{U}$ and $\boldsymbol{V}$. Due to the similarity among ultrasonic records during monitoring, it is valid to assume that $r << min(N, D)$, where N and D are respectively the number of records and the number of samples in each record. Figure 6-1 shows the pseudo-code of the algorithm. Two auxiliary matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are generated and decomposed using QR decomposition. It has been shown (Badeau et al. 2004) that $\boldsymbol{Q}_B^{r\times r}$ and $\boldsymbol{Q}_A^{D\times r}$ in Figure 6-1 converge to $\boldsymbol{U}$ and $\boldsymbol{V}$, *respectively*, and both $\boldsymbol{R}_A^{D\times r}$ and $\boldsymbol{R}_B^{r\times r}$ in Figure 6-1 converge to the singular value matrix $\boldsymbol{S}$. In each time step, a new measurement $x(t)$ is obtained and $\boldsymbol{Q}$ and $\boldsymbol{R}$ matrices are updated according to the process in Figure 6-1. At an arbitrary time step $t = t_s$, $\boldsymbol{Q}_A^{D\times r}$ approximately equals the first *r* singular vectors in $\boldsymbol{V}$ for all the records before t_s. Using equation (3-5), we project the data matrix $\boldsymbol{X}$ onto the bases in $\boldsymbol{Q}_A^{D\times r}$ to obtain the approximate left singular vectors at time step t_s, which characterize the slow variation within the records. We intentionally choose to use the projection of $\boldsymbol{X}$ on $\boldsymbol{Q}_A^{D\times r}$ to approximate left singular vectors rather than to use $\boldsymbol{Q}_B^{r\times r}$ directly. The reason is that the dimension of $\boldsymbol{Q}_B^{r\times r}$ is limited by the parameter r; therefore $\boldsymbol{Q}_B^{r\times r}$ only approximates the slow time variations for the last r records in slow time. By

projecting $\boldsymbol{X}$ to $\boldsymbol{Q}_A^{D\times r}$, we have the flexibility of obtaining approximate left singular vectors of more than r records, which helps us to better understand the slow time variations over a longer period of time.

Initialize: $Q_A(0) = \begin{bmatrix} I_r \\ \cdots \\ 0 \end{bmatrix}$; $Q_B(0) = \begin{bmatrix} I_r \\ \cdots \\ 0 \end{bmatrix}$; $R_A(0) = I_r$;

For each time step do:

Input: $x(t)$

First Iteration:

$h(t) = Q_A(t-1)^H x(t)$

$$\begin{bmatrix} B(t) \\ --- \\ \text{x} \ldots . \text{x} \end{bmatrix} = \begin{bmatrix} h(t)^H \\ ------- \\ Q_B(t-1)R_A(t-1)^H \end{bmatrix}$$

$B(t) = Q_B(t)R_B(t)$

Second Iteration:

$x_\perp(t) = x(t) - Q_A(t-1)h(t)$

$A(t) = Q_A(t-1)R_B(t)^H + x_\perp(t)q_{B_1}(t)^H$

$A(t) = Q_A(t)R_A(t)$

Figure 6-1 Pseudo code of sliding window adaptive singular value decomposition (Badeau et al. 2004)

The performance of sliding window adaptive SVD is related to the selection of r, which controls the low rank approximation of $\boldsymbol{X}$ at each time step (Badeau et al. 2004). In principle, we want r to be small to improve the computational efficiency; we also want r to be sufficiently large such that sufficient information is retained to cover important variations in the data. In our field experiments, we found that $r = 200$ achieved a good balance between efficiency and accuracy, as shown in *Section 4.3*. In the next sections, we will revisit earlier experiments and assess the performance of our proposed online monitoring method on those data.

6.3 Online Detection Of A Scatterer On A Pipe In Operation

In *Chapter 4*, we have shown that our SVD method can extract damage-sensitive features from a collection of ultrasonic pitch-catch records obtained in a field experiment. In this section, we show that our SVD damage-sensitive feature extraction process can be implemented in an online manner. We apply the online framework on field experiments records (earlier described in *Section 2.3*) starting from 1:13 am to 1:44 pm on September 1st, 2011. More specifically, we used 750 monitoring records collected in one-minute interval for 12.5 hours. A scatterer was applied on the pipe at 11:32 am, before we collected the 618th record. We implemented the sliding window adaptive SVD with r equals 200. The larger r is, the more historical information we keep during the monitoring, and the less noise is the damage sensitive singular vector. On the other hand, we prefer a small r such that the computation can be done in an online manner. Due to the significant noise and variations in this environment, we use in this experiment $r = 200$ to balance the accuracy and efficiency. Note that even if we choose a value for r, the algorithm can successfully detect damage before we have collected r records. The r value determines the number of singular vectors we store and update, therefore controls the resolution after we have collected more than r records.

The dataset we use in this experiments consists of more than 600 intact records before damage is introduced, because in practice, most initial measurements will be taken when the structure is healthy, whereas damage usually develops long after the start of data collection. But the method also works if the damage occurs much earlier.

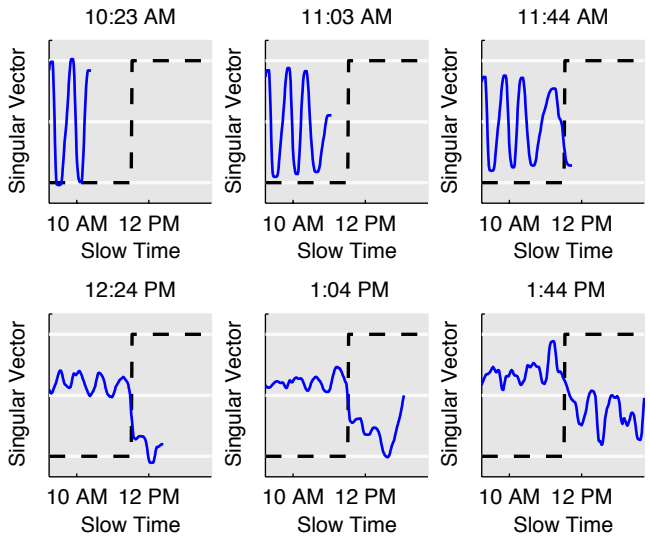


Figure 6-2 Snapshots of online updated damage-sensitive singular vector (blue), along with ground truth of mass toggling (black dashed). Scatterer was applied at 11:32AM.

Figure 6-2 shows the extracted damage-sensitive feature at six discrete time points during the experiment. The score η of the singular vector shown in each plot is labeled on top of the plot. The six snapshots were taken at 40-min intervals between 10:23 AM and 1:44 PM. In each snapshot, we plot the singular vector based on records that started at 9:13 AM, because in the preceding records (the first 8 hours of the test period) the structure remained intact. Note that the scatterer was applied at 11:32 AM (10.3 hours after the start of the experiment, or 2.3 hours after the 9:13 AM mark), and therefore the displayed time period covers about 4.5 hours, with roughly 2.3 hour before and 2.2 hour after the critical event – the occurrence of damage. Before the damage (the mass scatterer) is introduced, the singular vector is strongly correlated with temperature. Soon after the damage is introduced, the singular vector shows separation between the intact and the damaged records. Visually, we see when damage occurs in slow time, and we note that a simple linear threshold would perfectly separate the

damaged and undamaged records. The score η shown above each plot also confirms that damage was detected from the singular vector after the scatterer was introduced.

6.4 Online Baseline Removal For Scatterer Localization

We can apply the very same concept to the robust baseline removal process we introduced in *Chapter 5*. By adaptively updating the singular vectors, we localize the damage as each new record is collected, thereby monitoring the progress of the damage.

The process for the online SVD robust baseline removal is the same as that introduced in *Section 5.2 and 6.2*. For each receiving channel, we keep a right singular vector matrix of size $D \times R$, where D is the number of samples in each record and R is the reduced rank that controls how much information can be stored. As each new record arrives, the singular vector matrix is updated according to the process illustrate in Figure 6-1. We then identify the damage-sensitive singular vector from the singular vector matrix using the K-means based method demonstrated initially in *Section 3.2*. Combining the selected singular vectors from all channels, we obtain the scatter signal matrix that we will use to localize damage.

We implemented the above process on the data collected on the aluminum plate undergoing temperature variations, both spatially and temporally, illustrated earlier in Figure 5-6. We generate localization results roughly every 5 minutes for a duration of roughly 3 hours from 12:41 to 15:42, with damage (the mass scatterer) introduced at 14:14, obtaining a total of 37 images. The images are normalized such that the highest value is 1 (bright) and the lowest value is 0 (black).

Figure 6-3 shows 7 snapshots of the localization results during the 3-hour monitoring period. The three plots on the top row on the left show the localization results before damage

was introduced. At this stage, the images show no peak at the actual damage location. Instead, the peaks are near the transducer locations, because when no scatterer is present the baseline removal method returns the baseline signals that originate from the transducers.

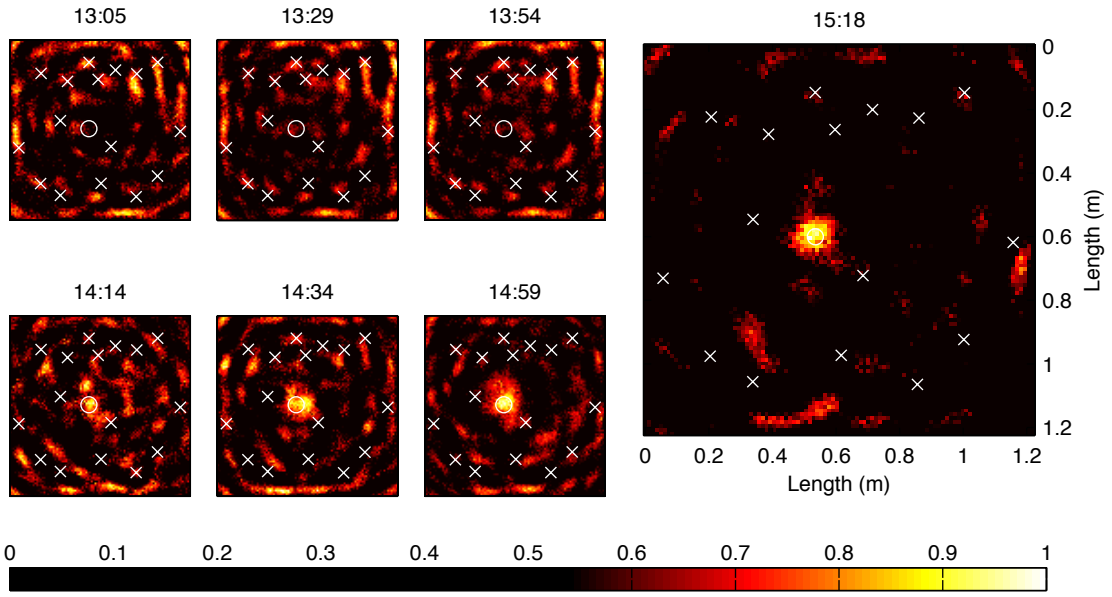


Figure 6-3 Snapshots of online localization results using SVD robust baseline removal

The three plots on the bottom row on the left in Figure 6-3 show results after the scatterer was applied, with dominant peaks at the scatterer location. We also observe that the localization performance improves as more records are added, because the scatter signal becomes a more dominant component in the data matrix and can be more accurately represented by a singular vector. The image on the right (which has been plotted at a different scale than the other six) shows the best performance localization result, which is obtained at 15:18, about an hour after the scatterer was introduced.

We note that in this experiment, localization performance decreased as we continued. The reason is that we maintain a singular vector matrix with 25 entries ($r = 25$). We obtained

best performance at 15:18, when 15 ‘damaged’ records had been collected after the scatterer was applied. Having sufficient information regarding the damage in the dataset, SVD effectively decomposes records, exposes the damage-sensitive scatter signals, and achieves satisfactory localization results. As we collect more ‘damaged’ records that contain the scatter signal, the average of the signals are increasingly affected by the presence of the scatter signals. Eventually, the $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$ vector from equation (3-11) no longer approximate the baseline signal, but rather a superimposition of baseline and scatter signals. Consequently, useful scatter information is removed (lost) along with baseline information when we compute the residual as in equation (3-16), thereby degrading the localization performance.

The online localization performance is satisfying in two ways. First, we can identify the indication of damage – the peak associated with the scatterer – shortly after the scatterer is introduced. Second, we can observe a clear improvement of performance as we collect more records, which confirms the detection and localization results and would help exclude false positives.

On the other hand, the scatter signal recovered by conventional temperature compensation methods fails to detect the presence of the scatterer. We performed temperature compensation to the records collected at the same times as shown above, and we show seven localization images in Figure 6-4 in the same layout used in Figure 6-3. We compensate the temperature using a reference record collected at 13:01, which should be a ‘near-optimal’ baseline for the record collected at 15:18 (the rightmost image in Figure 6-4)because the average temperature of the specimen at 13:01 and at 15:18 was almost the same. The correlation coefficient between the two records without temperature compensation is 0.9, which increases to 0.97 after conventional temperature compensation is performed;

this improvement largely corroborates the "near-optimal" baseline hypothesis. However, we see no indication of the scatterer location in any of these images, showing that temperature gradients across the plate defeat localization if conventional temperature compensation alone is used.

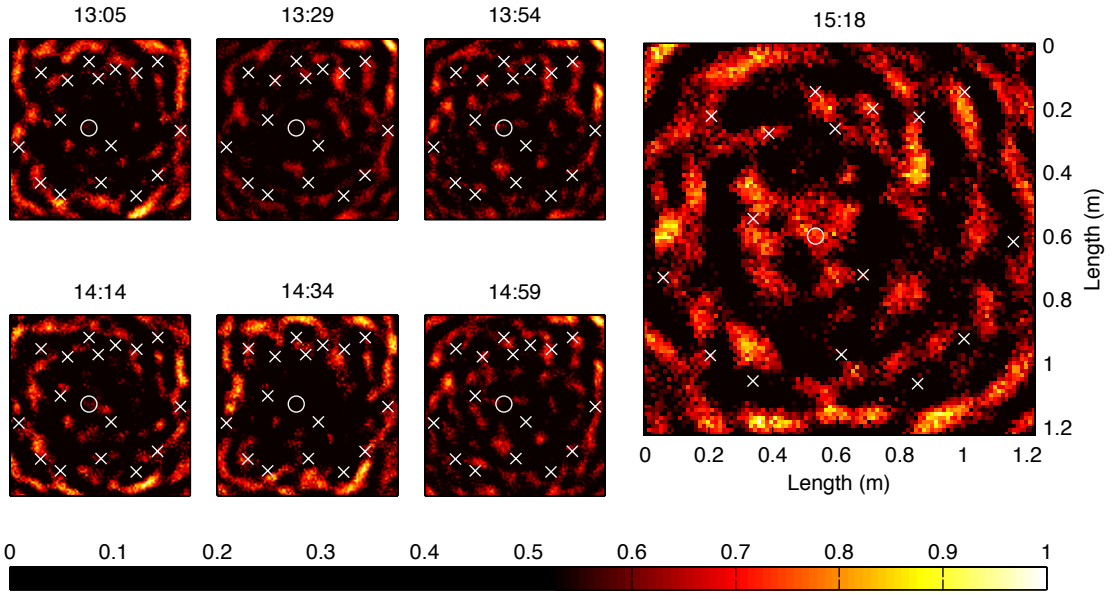


Figure 6-4 Localization results at the same time points using scatter signals recovered by temperature compensated

6.5 Summary

In this chapter, we extended our SVD framework to diagnose damage in an online manner. We integrate the sliding window adaptive framework to adaptively update the singular vectors with each new record. We demonstrated that by doing so, we efficiently and effectively utilize all historical information in the monitoring records, and conduct online damage detection and localization.

We also demonstrated the online framework using experimental data. We first showed online damage detection on field experimental data collected in the operating pipe system

introduced in *Chapter 2*. We showed snapshots of the updated singular vectors during the monitoring, and observed the presence of the mass scatterer after it is applied.

We then implemented SVD robust baseline removal in an online manner with the adaptive update framework, demonstrating its efficacy at detecting a scatterer on a plate specimen. We show that the updated singular vector was effective at recovering the scatter signals and localizing the scatterer. We show snapshots of the localization image during a 3-hr monitoring period; when we apply conventional temperature compensation methods we see no indication of damage.

Although the experimental results are promising, there are still several challenges/limitations associated with the practical implementation of this online damage detection and localization framework. The first is to choose a proper r value for the specific application. Ideally, we would like r to be as large as possible such that there is no limitation in the resolution in decomposing the variations. However, the value is often limited by the computational cost. The second challenge is that a threshold needs to be determined to classify a change as damage after the damage sensitive feature is extracted. As discussed before, we set a 0.5 empirical threshold on the η score to identify the presence of damage. Although the value works well in multiple experiments we took from different specimens, the threshold may vary in different experimental scenarios. Moreover, the threshold determines the delay-time for damage detection, which is crucial for some applications where damage needs to be determine right away. In our field implementation, the threshold of 0.5 results in a 20 minutes delay in identifying the presence of damage after the scatterer is applied. Using a smaller threshold shortens the delay of detection with the risk of generating false alarms. Thirdly, the current implementation of the framework is limited to detect abrupt

change. The damage sensitive feature selection method needs to be extended in order to implement the method to identify progressive damage.

Chapter 7 Conclusions And Future Work

In this dissertation, we presented a data-driven framework based on singular value decomposition to process guided wave monitoring records for robust and accurate damage detection and localization. We illustrated our field experimental investigation of guided wave monitoring of a pipe in operation, from which we observed large environmental and operational variations that significantly change the ultrasonic records and challenge conventional damage detection and localization methods. We then demonstrated our singular value decomposition methodology and show robust and accurate damage detection and localization using experimental monitoring records.

7.1 Contributions

Through this work, we have made several novel contributions to the current literature.

- **Field experimental investigation of guided wave monitoring of a pipe under environmental and operational variations.** In *Chapter 2*, we investigated the effects

of environmental and operational variations on ultrasonic records taken from a pressurized pipe in operation.

- **Singular value decomposition on ultrasonic records.** In *Chapter 3*, we analyzed the performance of singular value decomposition on ultrasonic records. Through simulations, we demonstrated that the decomposed singular vectors could represent different variations.

- **Damage-sensitive feature selection.** In *Chapter 3*, we developed a K-means based damage sensitive feature selection scheme identify the damage-sensitive singular vectors.

- **Robust scatterer detection using left singular vectors.** In *Chapter 4*, we applied our damage-sensitive feature extraction and selection process on field experimental data, and showed successful detection of a scatterer under large environmental and operational variations exhibited in a pipe in operation.

- **Robust baseline removal for scatterer localization.** In *Chapter 5*, we demonstrated successful localization of scatterers on plate and pipe structures under temperature variations and spatial temperature gradients, which defeats conventional temperature compensation methods.

- **Integrating adaptive updating of singular vectors for online monitoring**. In *0*, we integrated sliding window adaptive singular value decomposition algorithm with our damage detection and localization framework, and demonstrated successful online monitoring of the structural integrity.

The combination of these experimental investigations and data processing techniques presents a strong data-processing framework to extract and identify damage related information from guided wave structure health monitoring records. We evaluated the methodology with theoretic analysis, synthetic simulations, and laboratory/field experiments. The methodology showed satisfactory robustness and capability of processing long-term monitoring records, and successfully addressed two major challenges associated with practical implementation of guided wave structural health monitoring on large civil infrastructures. In the following section, we discuss several directions to improve the developed framework and to expend its applications.

7.2 Future Research Directions

7.2.1 Extending the study of degradation identification

In *Chapter 4-6*, we applied our SVD methodology experimentally to detect and localize a scatterer that produces an abrupt change in the ultrasonic monitoring records. However, structural damage does not form immediately. In *Chapter 4*, we briefly discussed and demonstrated using our SVD methodology to identify progressive degradation caused by pitting corrosion in a pipe. In laboratory experiments our method has shown to be effective in separating the trend of structural degradation from benign temperature variation. However, in real-world environments, structural degradation may form in a complex process that is usually much slower than the accelerated corrosion process in our laboratory experiment. Moreover, besides temperature variations that we recreated experimentally, there exist other slow-time variations that may mimic degradation. Therefore, more data is needed to robustly distinguish the degradation from other long-term trends in monitoring data. The solution to

this problem may also require efforts to understand the behavior of degradation, to investigate other possible benign variations, and to develop statistical metrics to identify the change of interest.

7.2.2 Integration with other damage localization/characterization methods

In *Chapter 5*, we demonstrate our SVD method as a robust baseline removal method under temperature variation and spatial gradients. We demonstrate successful localization of a scatterer integrating our method with widely used delay-and-sum localization algorithms. However, delay-and-sum localization has limited resolution because it discards the phase information contained in the scatter signal. Moreover, the method assumes a point scatterer and may be less effective localizing damage with complex shapes.

Our SVD methodology can robustly and accurately extract damage information from monitoring data, and therefore can be integrated with any damage detection and localization algorithms that require removing the baseline or background information. We have explored integrating other localization algorithms besides delay-and-sum localization, such as data-driven match-field processing, and minimum variance distortionless response (MVDR) localization algorithm. In a broader sense, the SVD-based methods introduced in this dissertation are potentially beneficial to fields and disciplines other than guided wave structural health monitoring.

7.2.3 Extension to new applications

For our future work, we plan to extend the structural health monitoring framework and data processing methodology developed to more complex structures under more complex environments. We plan to extend our investigation in two directions:

- Inhomogeneous and/or anisotropic materials.

The experimental works in this thesis were conducted on pipe and plate structures made of aluminum or steel, which are generally homogeneous and isotropic. Many other materials that are found in modern structures exhibit inhomogeneous and/or anisotropic properties that further complicate the analysis. Therefore, the robustness of our method needs to be reevaluated on such material and structures.

- Other environmental and operational variations

In our experimental investigation, we observed or synthesized environmental and operational variations that exhibited in operating pipe systems. However, real world implementation of guided wave monitoring system may be challenged by application-specific variations that were not investigated in our analysis. Therefore, more experimental data and modeling is needed to evaluate the robustness of our method.

References

Alleyne, D. N., and Cawley, P. (1992). "The interaction of Lamb waves with defects." *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 39(3), 381–397.

Alleyne, D. N., Pavlakovic, B., Lowe, M. J. S., and Cawley, P. (2001). "Rapid, long range inspection of chemical plant pipework using guided waves." *AIP conference proceedings*, 180.

Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammerling, S., McKenney, A., and others. (1999). *LAPACK Users' guide*. Siam.

Auld, B. A. (1973). *Acoustic fields and waves in solids*. Wiley New York.

Badeau, R., Richard, G., and David, B. (2004). "Sliding window adaptive SVD algorithms." *Signal Processing, IEEE Transactions on*, 52(1), 1–10.

Balageas, D. (2006). *Introduction to structural health monitoring*. Wiley Online Library.

Bishop, C. M. (1994). "Neural networks and their applications." *Review of Scientific Instruments*, 65(6), 1803–1832.

Bradley, A. P. (1997). "The use of the area under the ROC curve in the evaluation of machine learning algorithms." *Pattern Recognition*, 30(7), 1145–1159.

Cawley, P., Lowe, M. J. S., Alleyne, D. N., Pavlakovic, B., and Wilcox, P. (2003). "Practical long range guided wave inspection-applications to pipes and rail'." *Materials Evaluation*, 61(1), 66–74.

Cherkassky, V., and Muller, F. (1998). “Learning ffrom data.” *New York: John Wiey & Sons*.

Clarke, T., Cawley, P., Wilcox, P. D., and Croxford, A. J. (2009). “Evaluation of the damage detection capability of a sparse-array guided-wave SHM system applied to a complex structure under varying thermal conditions.” *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 56(12), 2666–2678.

Clarke, T., Simonetti, F., and Cawley, P. (2010). “Guided wave health monitoring of complex structures by sparse array systems: influence of temperature changes on performance.” *Journal of Sound and Vibration*, 329(12), 2306–2322.

Comon, P., and Golub, G. H. (1990). “Tracking a few extreme singular values and vectors in signal processing.” *Proceedings of the IEEE*, 78(8), 1327–1343.

Cortes, C., and Vapnik, V. (1995). “Support vector machine.” *Machine learning*, 20(3), 273–297.

Croxford, A. J., Moll, J., Wilcox, P. D., and Michaels, J. E. (2010). “Efficient temperature compensation strategies for guided wave structural health monitoring.” *Ultrasonics*, 50(4), 517–528.

Croxford, A. J., Wilcox, P. D., Konstantinidis, G., and Drinkwater, B. W. (2007). “Strategies for overcoming the effect of temperature on guided wave structural health monitoring.” *Proceedings of SPIE*, 65321T–1.

Davies, A. (1998). *Handbook of condition monitoring: techniques and methodology*. Springer.

Demma, A., Cawley, P., Lowe, M., Roosenbrand, A. G., and Pavlakovic, B. (2004). “The reflection of guided waves from notches in pipes: a guide for interpreting corrosion measurements.” *NDT & E International*, 37(3), 167–180.

Doebling, S. W., Farrar, C. R., and Prime, M. B. (1998). “A summary review of vibration-based damage identification methods.” *Shock and vibration digest*, 30(2), 91–105.

Edwards, G. R., and Gan, T.-H. (2007). “Detection of Corrosion in Offshore Risers Using Guided Ultrasonic Waves.” *ASME 2007 26th International Conference on Offshore Mechanics and Arctic Engineering*, American Society of Mechanical Engineers, 377–384.

Facility Management Services, C. M. U. (n.d.). “Carnegie Mellon Rooms Specification(s).” <https://fst-apps.andrew.cmu.edu/roomspec/roomspec>.

Farrar, C. R., and Worden, K. (2007). “An introduction to structural health monitoring.” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851), 303–315.

Fessler, R. R. (2008). *Pipeline Corrosion*. Report submitted to US Department of Transportation Pipeline and Hazardous Materials Safety Administration Office of Pipeline Safety.

Figueiredo, E., Park, G., Farrar, C. R., Worden, K., and Figueiras, J. (2011). "Machine learning algorithms for damage detection under operational and environmental variability." *Structural Health Monitoring*, 10(6), 559–572.

Fukuhara, M., and Yamauchi, I. (1993). "Temperature dependence of the elastic moduli, dilational and shear internal frictions and acoustic wave velocity for alumina,(Y) TZP and β′-sialon ceramics." *Journal of materials science*, 28(17), 4681–4688.

Gao, H., Shi, Y., and Rose, J. L. (2005). "Guided wave tomography on an aircraft wing with leave in place sensors." *Review of progress in quantitative nondestructive evaluation*, AIP Publishing, 1788–1794.

Golub, G. H., and Loan, C. F. V. (1996). *Matrix Computations*. JHU Press.

Golub, G., and Kahan, W. (1965). "Calculating the singular values and pseudo-inverse of a matrix." *Journal of the Society for Industrial & Applied Mathematics, Series B: Numerical Analysis*, 2(2), 205–224.

Graff, K. F. (1975). *Wave Motion in Elastic Solids*. Courier Dover Publications.

Grall, A., Bérenguer, C., and Dieulle, L. (2002). "A condition-based maintenance policy for stochastically deteriorating systems." *Reliability Engineering & System Safety*, 76(2), 167–180.

Hall, J. S., and Michaels, J. E. (2010). "Minimum variance ultrasonic imaging applied to an in situ sparse guided wave array." *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 57(10), 2311–2323.

Harley, J. B., and Moura, J. M. F. (2012). "Scale transform signal processing for optimal ultrasonic temperature compensation." *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 59(10), 2226–2236.

Harley, J. B., and Moura, J. M. F. (2013). "Sparse recovery of the multimodal and dispersive characteristics of Lamb wavesa)." *The Journal of the Acoustical Society of America*, 133(5), 2732–2745.

Harley, J. B., and Moura, J. M. F. (2014). "Data-driven matched field processing for Lamb wave structural health monitoring." *The Journal of the Acoustical Society of America*, 135(3), 1231–1244.

Harley, J. B., Thavornpitak, N., and Moura, J. M. (2013). "Delay-and-sum technique for localization of active sources in cylindrical objects." *REVIEW OF PROGRESS IN QUANTITATIVE NONDESTRUCTIVE EVALUATION: VOLUME 32*, AIP Publishing, 294–301.

Harley, J. B., Ying, Y., Moura, J. M., Oppenheim, I. J., Sobelman, L., and Garrett, J. H. (2012). "Application of Mellin transform features for robust ultrasonic guided wave structural health monitoring." *AIP Conference Proceedings*, 1551.

Haynes, C., Todd, M. D., Flynn, E., and Croxford, A. (2013). "Statistically-based damage detection in geometrically-complex structures using ultrasonic interrogation." *Structural Health Monitoring*, 12(2), 141–152.

Jardine, A. K., Lin, D., and Banjevic, D. (2006). "A review on machinery diagnostics and prognostics implementing condition-based maintenance." *Mechanical systems and signal processing*, 20(7), 1483–1510.

Konstantinides, K., and Yovanof, G. S. (1995). "Application of SVD based spatial filtering to video sequences." *, 1995 International Conference on Acoustics, Speech, and Signal Processing, 1995. ICASSP-95*, 2193–2196 vol.4.

Konstantinidis, G., Drinkwater, B. W., and Wilcox, P. D. (2006). "The temperature stability of guided wave structural health monitoring systems." *Smart Materials and Structures*, 15(4), 967.

Konstantinidis, G., Wilcox, P. D., and Drinkwater, B. W. (2007). "An investigation into the temperature stability of a guided wave structural health monitoring system using permanently attached sensors." *Sensors Journal, IEEE*, 7(5), 905–912.

Krautkrämer, J., and Krautkrämer, H. (1983). *Ultrasonic Testing of Materials*. Springer Berlin Heidelberg, Berlin, Heidelberg.

Laory, I., Trinh, T. N., and Smith, I. F. C. (2011). "Performance of Two Model-Free Data Interpretation Methods for Continuous Monitoring of Structures under Environmental Variations." ASCE, 4–4.

Leonard, K. R., and Hinders, M. K. (2003). "Guided wave helical ultrasonic tomography of pipes." *The Journal of the Acoustical Society of America*, 114(2), 767–774.

Li, J., and Rose, J. L. (2001). "Implementing guided wave mode control by use of a phased transducer array." *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 48(3), 761–768.

Liu, C., Harley, J. B., Berges, M., Greve, D. W., Junker, W., and Oppenheim, I. J. (2014). "A Robust Baseline Removal Method for Guided Wave Damage Localization." *SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring*.

Liu, C., Harley, J. B., O'Donoughue, N., Ying, Y., Altschul, M. H., Garrett Jr, J. H., Moura, J. M., Oppenheim, I. J., and Soibelman, L. (2012). "Ultrasonic monitoring of a pipe under operating conditions." *SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring*, 83450B–83450B.

Liu, C., Harley, J. B., Ying, Y., Altschul, M. H., Bergés, M., Garrett, Jr., J. H., Greve, D. W., Moura, J. M. F., Oppenheim, I. J., and Soibelman, L. (2013a). "Ultrasonic Monitoring of a Pressurized Pipe in Operation." *Structures Congress 2013*, American Society of Civil Engineers, 1903–1913.

Liu, C., Harley, J. B., Ying, Y., Oppenheim, I. J., Bergés, M., Greve, D. W., and Garrett, J. H. (2013b). "Singular value decomposition for novelty detection in ultrasonic pipe

monitoring." *SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring*, 86921R–86921R.

Lowe, M. J., Alleyne, D. N., and Cawley, P. (1998). "Defect detection in pipes using guided waves." *Ultrasonics*, 36(1), 147–154.

Lu, Y., and Michaels, J. E. (2009). "Feature extraction and sensor fusion for ultrasonic structural health monitoring under changing environmental conditions." *Sensors Journal, IEEE*, 9(11), 1462–1471.

McGregor, J., Nooteboom, B., Ivory, N., Lot, P.-I., Belait, T. P. P., and Di-Raja, J. S. (2004). "The application of long range guided ultrasonics for the inspection of riser pipes." *16th WCNDT, Montreal*.

Michaels, J. E. (2008). "Detection, localization and characterization of damage in plates with an in situ array of spatially distributed ultrasonic sensors." *Smart Materials and Structures*, 17(3), 035035.

Michaels, J. E., Jacobs, L. J., and Ruzzene, M. (2012). *A Multi-Scale Structural Health Monitoring Approach for Damage Detection, Diagnosis and Prognosis in Aerospace Structures*. DTIC Document.

Michaels, J. E., and Michaels, T. E. (2005). "Detection of structural damage from the local temporal coherence of diffuse ultrasonic signals." *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 52(10), 1769–1782.

Na, W.-B., and Kundu, T. (2002). "Underwater pipeline inspection using guided waves." *TRANSACTIONS-AMERICAN SOCIETY OF MECHANICAL ENGINEERS JOURNAL OF PRESSURE VESSEL TECHNOLOGY*, 124(2), 196–200.

National Instruments Corporation. (2010). "Specification for the LabVIEW Measurement File." National Instruments Corporation.

Nieuwenhuis, J. H., Neumann, J.J., J., Greve, D. W., and Oppenheim, I. J. (2005). "Generation and detection of guided waves using PZT wafer transducers." *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 52(11), 2103–2111.

Oh, C. K., and Sohn, H. (2009). "Damage diagnosis under environmental and operational variations using unsupervised support vector machine." *Journal of Sound and Vibration*, 325(1), 224–239.

Piezo systems, I. (n.d.). "PSI-5A4E piezoceramic sheets and their properties." <http://www.piezo.com/prodsheet1sq5A.html>.

Pipeline & Hazardous Materials Saftey Administration, U. S. D. of T. (n.d.). "Pipeline Incidents and Mileage Reports." <http://primis.phmsa.dot.gov/comm/reports/safety/PSI.html> (Jun. 26, 2013).

Posenato, D., Lanata, F., Inaudi, D., and Smith, I. F. (2008). “Model-free data interpretation for continuous monitoring of complex structures.” *Advanced Engineering Informatics*, 22(1), 135–144.

Putkis, O., and Croxford, A. J. (2013a). “Continuous baseline growth and monitoring for guided wave SHM.” *Smart Materials and Structures*, 22(5), 055029.

Putkis, O., and Croxford, A. J. (2013b). “Continuous baseline growth and monitoring for guided wave SHM.” *Smart Materials and Structures*, 22(5), 055029.

Raghavan, A., and Cesnik, C. E. (2007). “Review of guided-wave structural health monitoring.” *Shock and Vibration Digest*, 39(2), 91–116.

Raghavan, A., and Cesnik, C. E. (2008a). “Effects of elevated temperature on guided-wave structural health monitoring.” *Journal of Intelligent Material Systems and Structures*, 19(12), 1383–1398.

Raghavan, A., and Cesnik, C. E. (2008b). “Effects of elevated temperature on guided-wave structural health monitoring.” *Journal of Intelligent Material Systems and Structures*, 19(12), 1383–1398.

Rao, B. K. N. (1996). *Handbook of condition monitoring*. Elsevier.

Rose, J. L. (2000). “Guided wave nuances for ultrasonic nondestructive evaluation.” *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 47(3), 575–583.

Rose, J. L. (2004). *Ultrasonic Waves in Solid Media*. Cambridge University Press.

Salama, K., and Ling, C. K. (1980). “The effect of stress on the temperature dependence of ultrasonic velocity.” *Journal of Applied Physics*, 51(3), 1505–1509.

di Scalea, F. L., and Salamone, S. (2008). “Temperature effects in ultrasonic Lamb wave structural health monitoring systems.” *The Journal of the Acoustical Society of America*, 124, 161.

Scholkopft, B., and Mullert, K.-R. (1999). “Fisher discriminant analysis with kernels.”

Silva, M. Z., Gouyon, R., and Lepoutre, F. (2003). “Hidden corrosion detection in aircraft aluminum structures using laser ultrasonics and wavelet transform signal analysis.” *Ultrasonics*, 41(4), 301–305.

Sohn, H. (2007). “Effects of environmental and operational variability on structural health monitoring.” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851), 539–560.

Sohn, H., Czarnecki, J. A., and Farrar, C. R. (2000). “Structural health monitoring using statistical process control.” *Journal of Structural Engineering*, 126(11), 1356–1363.

Sohn, H., Farrar, C. R., Hemez, F. M., Shunk, D. D., Stinemates, D. W., Nadler, B. R., and Czarnecki, J. J. (2004). *A review of structural health monitoring literature: 1996-2001*. Los Alamos National Laboratory Los Alamos,, New Mexico.

Strang, G. (2009). *Introduction to linear algebra*. Wellesley Cambridge Press.

Su, Z., Ye, L., and Lu, Y. (2006). "Guided Lamb waves for identification of damage in composite structures: A review." *Journal of sound and vibration*, 295(3), 753–780.

Tiratsoo, J. N. H. (1992). *Pipeline Pigging Technology*. Gulf Professional Publishing.

Turk, M. A., and Pentland, A. P. (1991). "Face recognition using eigenfaces." *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on*, IEEE, 586–591.

Tuzzeo, D., and di Scalea, F. L. (2001). "Noncontact Air-Coupled Guided Wave Ultrasonics for Detection of Thinning Defects in Aluminum Plates." *Research in Nondestructive Evaluation*, 13(2), 61–77.

Vanlanduit, S., Parloo, E., Cauberghe, B., Guillaume, P., and Verboven, P. (2005). "A robust singular value decomposition for damage detection under changing operating conditions and structural uncertainties." *Journal of Sound and Vibration*, 284(3), 1033–1050.

Wei, J.-J., Chang, C.-J., Chou, N.-K., and Jan, G.-J. (2001). "ECG data compression using truncated singular value decomposition." *Information Technology in Biomedicine, IEEE Transactions on*, 5(4), 290–299.

Worden, K., and Manson, G. (2007). "The application of machine learning to structural health monitoring." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851), 515–537.

Yan, A.-M., Kerschen, G., De Boe, P., and Golinval, J.-C. (2005). "Structural damage diagnosis under varying environmental conditions—part I: a linear analysis." *Mechanical Systems and Signal Processing*, 19(4), 847–864.

Ying, Y. (2012). "A Data-Driven Framework for Ultrasonic Structural Health Monitoring of Pipes." *PhD Thesis. Carnegie Mellon University*.

Ying, Y., Garrett Jr, J. H., Harley, J. B., Oppenheim, I. J., Shi, J., and Soibelman, L. (2012). "Damage Detection in Pipes under Changing Environmental Conditions using Embedded Piezoelectric Transducers and Pattern Recognition Techniques." *Journal of Pipeline Systems Engineering and Practice*.

Ying, Y., Harley, J. B., Garrett Jr, J. H., Jin, Y., Oppenheim, I. J., Shi, J., and Soibelman, L. (2011). "Applications of machine learning in pipeline monitoring." *Proceedings of 2011 ASCE International Workshop on Computing in Civil Engineering*, 242–249.

Zhu, W., Rose, J. L., Barshinger, J. N., and Agarwala, V. S. (1998). "Ultrasonic guided wave NDT for hidden corrosion detection." *Journal of Research in Nondestructive Evaluation*, 10(4), 205–225.

Zolet, T. I., and Moore, S. A. (2003). "Pipeline Safety Improvement Act of 2002, The." *Energy Law Journal*, 24, 107.