

AD_____

Award Number: W81XWH-06-1-0389

TITLE: Contrast Enhancement for Thermal Acoustic Breast Cancer Imaging via Resonant Stimulation

PRINCIPAL INVESTIGATOR: Jian Li Ph.D.
Mark Sheplak Ph.D.
Lou Cattafesta Ph.D.
Henry Zmuda Ph.D.
Huabei Jiang Ph.D.
Manuel Arreola Ph.D.

CONTRACTING ORGANIZATION: University of Florida
Gainesville, FL 32603

REPORT DATE: March 2009

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) 01-03-2009		2. REPORT TYPE Annual		3. DATES COVERED (From - To) 27 Feb 2008 - 26 Feb 2009	
4. TITLE AND SUBTITLE Contrast Enhancement for Thermal Acoustic Breast Cancer Imaging via Resonant Stimulation				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-06-1-0389	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Jian Li Ph.D., Mark Sheplak Ph.D., Lou Cattafesta Ph.D., Henry Zmuda Ph.D., Huabei Jiang Ph.D., Manuel Arreola Ph.D. E-Mail: li@dsp.ufl.edu				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Florida Gainesville, FL 32603				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT This research plans to develop enhanced contrast thermal acoustic imaging (TAI) technology for the detection of breast cancer by combining amplitude-modulated (AM) electromagnetic (EM) field excitation, resonant acoustic scattering, and advanced signal processing techniques. While EM-induced TAI possesses great promise, the thermal acoustic signals tend to be weak. However, when the tumor is excited into resonance via EM stimulation, the effective acoustic scattering cross-section may increase by a factor in excess of 100 based on predictions for microsphere-based ultrasound contrast agents. Such an increase would truly be revolutionary, making the EM-induced TAI technology a very promising candidate for routine breast screening. The image formation methods in the existing TAI systems are data-independent and have poor resolution and high sidelobe problems. We will devise adaptive image formation algorithms to achieve high resolution and excellent interference and noise suppression capability.					
15. SUBJECT TERMS Thermal Acoustic Imaging (TAI), Resonance, Electromagnetic, Adaptive Imaging					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	74	19b. TELEPHONE NUMBER (include area code)

Table of Contents

	<u>Page</u>
Introduction.....	1
Body.....	1
Key Research Accomplishments	7
Reportable Outcomes.....	8
Conclusion.....	8
References.....	9
Appendices.....	10

I. Introduction

This research is focused on developing enhanced contrast thermal acoustic imaging (TAI) technology for the detection of breast cancer by combining amplitude-modulated (AM) electromagnetic (EM) field excitation, resonant acoustic scattering, and advanced signal processing techniques. EM-induced TAI combines the merits of both EM stimulation and ultrasound imaging, while overcoming their respective limitations. EM imaging provides excellent contrast between cancerous and normal breast tissue, but the long wavelengths provide poor spatial resolution. Conventional ultrasound imaging possesses very fine millimeter-range spatial resolution but poor soft tissue contrast. While EM-induced TAI possesses great promise, the thermal acoustic signals tend to be weak. However, when the tumor is excited into resonance via EM stimulation, the effective acoustic scattering cross-section may increase by a factor in excess of 100 based on predictions for microsphere-based ultrasound contrast agents. Such an increase would truly be revolutionary, making the EM-induced TAI technology a very promising candidate for routine breast screening. To induce the resonant response from the tumor, we are considering various approaches including, for example, AM continuous wave (CW) EM stimulation, where the modulation frequency range contains the predicted resonant frequencies for a distribution of tumor sizes and contrast ratios. The carrier frequency of the EM stimulation can be fixed and chosen for the best penetration and heat absorption. The image formation methods in the existing TAI systems are predominantly data-independent delay-and-sum (with or without weighting) type of approaches. These approaches tend to have poor resolution (relative to the best possible resolution a transducer array can offer) and high sidelobe problems, especially when the transducer array is not composed of uniformly and linearly spaced transducers, which is the case for the existing TAI systems. We are devising adaptive image formation algorithms to achieve high resolution and excellent interference and noise suppression capability.

II. Body

II.1 Experimental Progress

(a) Analytic Solutions

Previous work in the area of thermo-acoustic imaging all utilized high peak power, short pulse excitation [1-4]. In essence these approaches are time domain based, and capture the electro-acoustic impulse response of the phantom system to short, intense EM illumination with a prescribed (high) energy density. They generally require a prohibitively expensive power amplifier along with broadband microwave components and a high-speed data

acquisition system. The present study approaches the problem from the frequency domain, and seeks the same information as the time-domain approaches but with using lower power, narrow band excitation to obtain the steady-state response. It is anticipated that due to the thermo-acoustic resonance of phantom, similar imaging information can be obtained but with a significant reduction in excitation power.

In 1988 Diebold [5] provided a theoretical analysis for pressure wave generation by exciting droplets with a modulated laser pulse. Diebold's analysis closely parallels the situation studied in this effort since the excitation, be it microwave or laser, are both EM waves, sharing the same wave equations.

Using Diebold's approach, it can be shown that the steady-state pressure response of the phantom to an AM EM wave will be of the form:

$$p_f = \frac{j\beta\sigma E^2 c_s a}{C_p \hat{r}} \left(\frac{(\sin q - q \cos q) \frac{1}{q^2}}{\left(1 - \frac{\rho_s}{\rho_f}\right) \frac{\sin q}{q} - \cos q + j \frac{\rho_s c_s}{\rho_f c_f} \sin q} \right) e^{-jq\hat{r}} \quad (1)$$

Where the phantom has the properties:

β = thermal expansion coefficient [1/K]

σ = electrical conductivity [S/m]

C_p = specific heat [J/kg*K]

E = electric field intensity inside the phantom (assumed to be uniform over the volume of the phantom) [V/m]

c_s = speed of sound in phantom [m/s]

c_f = speed of sound in surrounding material [m/s]

ρ_s = density of phantom [kg / m³]

ρ_f = density of surrounding materials [kg / m³]

a = radius of phantom [m]

Additionally the terms q , $\hat{t}$, and $\hat{r}$ are the normalized modulation frequency, normalized retarded time, and normalized position of transducer, respectively, and are defined as:

$$\begin{aligned}\hat{t} &= \frac{c_s}{a} t && \text{dimensionless time} \\ \hat{\tau} &= \frac{c_s}{a} \left[t - \frac{(r-a)}{c_f} \right] && \text{retarded time} \\ q &= \frac{\omega a}{c_s} = k_s a && \text{wave vector} \\ \hat{r} &= \frac{r}{a} && \text{distance from sphere center}\end{aligned}$$

From this analysis, we can conclude that the frequency domain response characteristics of the phantom pressure signal are determined primarily by the density ratio and sound speed ratio between phantom and surrounding materials. The amplitude of pressure response primarily relies on the material property of phantom (thermal expansion coefficient, conductivity, specific heat), electric field intensity imposed on phantom, and relative position of transducer.

This analysis constitutes the foundation of our numerical simulation. Until very recently we have not had a high degree of confidence in two critical parameters, namely the density of phantom and the electric field intensity distribution which can be obtained from EM field measurements in the tank. For other material constants, specifically β , σ , C_p , and c_s , we have relied on intelligent estimates and are currently in the process of designing experiments to accurately measure and confirm the values used. Figure 1 shows the expected phantom pressure response as a function of modulation frequency based on the assumed material constants.

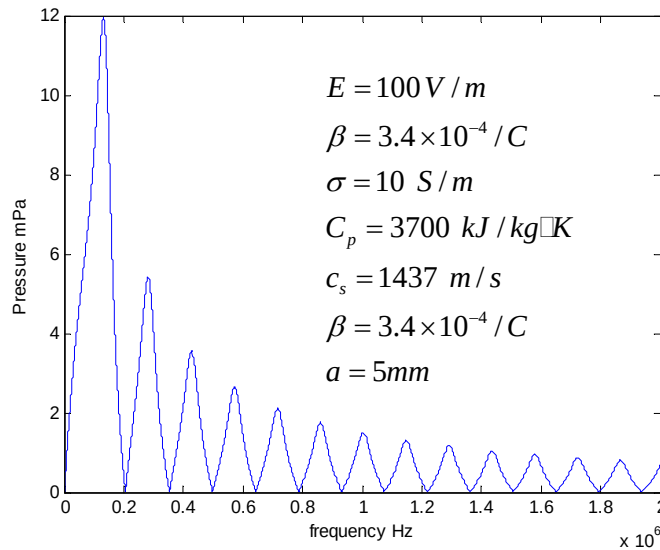


Figure 1: Phantom pressure response as a function of modulation frequency.

Note that the largest resonant peaks occur at frequencies which are outside the bandwidth of the transducer used (center frequency 1MHz, frequency band 0.6 MHz, Panametrics NDT,

Model: V303). Hence, in addition to better determining the material properties of the phantom, a smaller phantom with higher resonant frequencies is currently being designed and fabricated.

(b) Experimental Setup

The experimental setup used is shown in Figure 2. A National Instruments PXI system is used to generate the amplitude-modulated RF signal. Signal power out from PXI system is set to 0 dBm, with a RF carrier frequency of 542 MHz corresponding to the frequency at which the input reflection factor of the EM aperture is a minimum. This is the frequency at which maximum power transfer occurs. The modulating frequency can be swept from 10 kHz to 2 MHz with a user determined frequency step size. A 50 dB RF power amplifier is used to increase the power level into the aperture to 100 W. An acoustic transducer is connected to a spectrum analyzer via a 40 dB low noise amplifier (LNA). The frequency response of the overall system is displayed and recorded on the spectrum analyzer, centered around the modulation bandwidth.

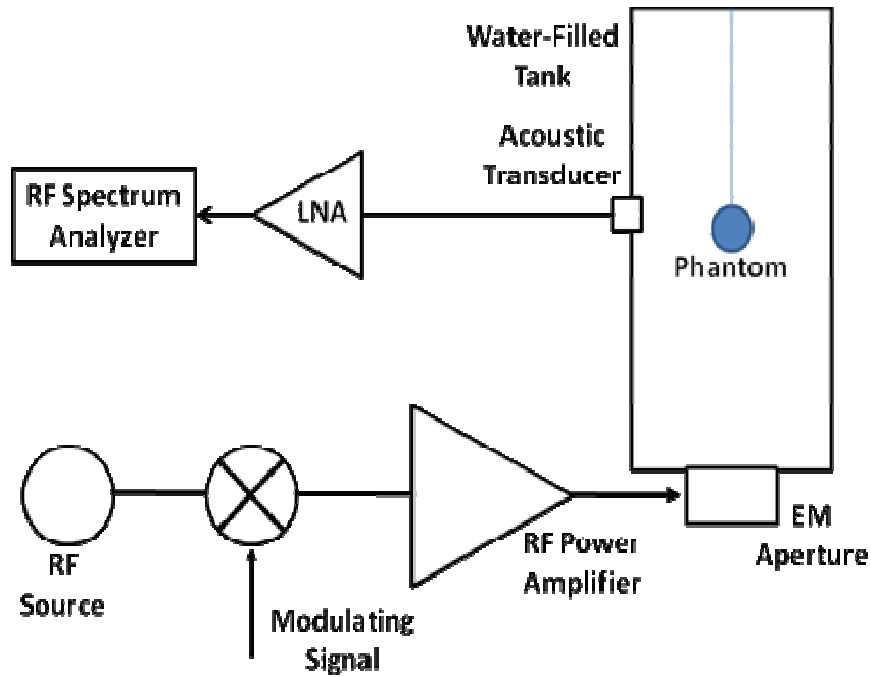


Figure 2: Experimental setup.

(c) Experimental Results and Analysis

The first step in measuring any thermoacoustic resonance is the determination of the noise floor of the overall measuring system, including the equivalent noise pressure present at the transducer. Experimental determination of the transducer sensitivity [6], in conjunction with the noise floor of the overall measurement system gives the equivalent noise pressure as a function of frequency shown in Figure 3.

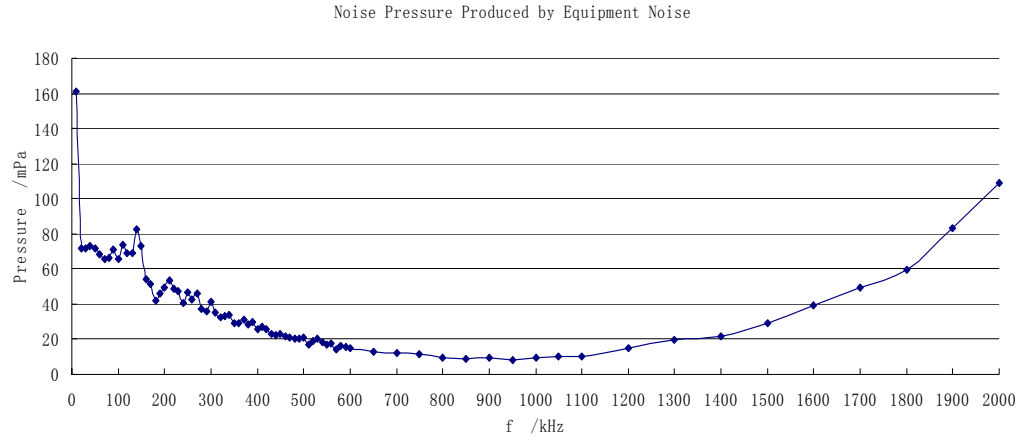


Figure 3: Equivalent noise pressure of overall measurement system.

Comparing Figures 2 and 3, it is observed that the equivalent noise pressure is greater than the estimated acoustic pressure signal obtained from simulations. Mitigation of this problem requires a lowering of the overall noise floor and/or increase in RF power level.

(d) Ongoing Work

The experimental evidence indicates that the thermo-acoustic signal is currently too weak to be detected. Comparison with related work shows that the energy levels used here are still roughly two orders of magnitude less than other approaches [1-4]. Four modifications are being made to the existing system to better extract the acoustic pressure wave out of the noise.

1. The RF radiating aperture, currently at the bottom of the tank (as shown in Figure 2), is being modified so as to place the aperture in close proximity to the phantom. This will deliver the maximum electric field intensity (maximum available energy) to the phantom. Note from Equation (1) that the pressure intensity varies as the square of the field amplitude, so this will result in a significant increase in acoustic pressure wave level.

2. Modified phantoms are being designed. Adding more salt to increase the conductivity of the phantom will also result in an increase in expected pressure. The material properties of the phantom as well as its size are being adjusted to place its resonance peak at a frequency corresponding to the location of maximum sensitivity of the transducer.

3. Improved data acquisition and signal processing, such as averaging and filtering, will serve to lower the noise floor of the overall system.

4. The tank is being modified so as to make the acoustic transducer (receiver) closer to the phantom, improving signal-to-noise-ratio (SNR).

The major reason for the delay in finding the acoustic pressure resonance is the lack of knowledge of its location in the frequency domain and its low intensity. Once a resonance is located using the mitigation strategy above, its acquisition can be optimized.

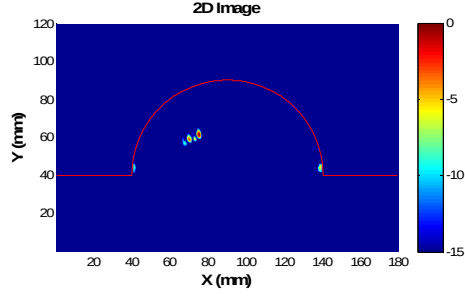
II. 2 Adaptive Image Formation Algorithms

Developing accurate and robust image reconstruction methods is one of the key challenges encountered in TAI. Delay-and-sum (DAS) beamforming is a widely used reconstruction algorithm for TAI. This method is data independent and may suffer from poorer resolution and worse interference suppression capability than its data adaptive counterpart, such as the standard Capon beamformer (SCB) [7]. However, the performance of SCB is sensitive to the errors in the sample covariance matrix and the signal steering vector. To deal with the aforementioned problems, many robust adaptive beamforming algorithms have been proposed [8]. However, most of the existing robust schemes are user parameter dependent and it may not be a simple task to determine the user parameters in practice. Therefore, user parameter-free robust adaptive approaches, including a shrinkage-based general linear combination (GLC) algorithm, are desirable [9, 10].

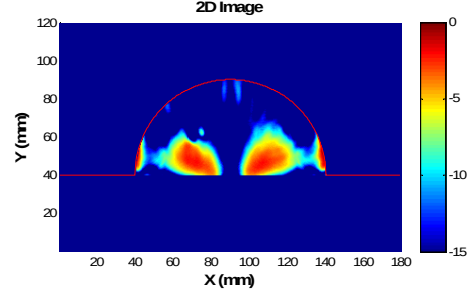
We propose an automatic (i.e., user parameter free) multifrequency adaptive and robust technique (AMART) based on GLC for TAI to achieve high resolution and good interference suppression capability. By using a multiple frequency source instead of a single frequency source, AMART can offer higher SNR and higher imaging contrast than its single frequency counterpart, which we refer to as the automatic single-frequency adaptive and robust technique (ASART), and also, the AMART algorithm can suppress the interference due to inhomogeneous breast tissue more effectively, since much more information about the human breast tissues can be harvested from the multiple frequencies. AMART is a three-stage imaging algorithm. Specifically, in the first stage of AMART, GLC is used to estimate the thermal acoustic responses from the grid points within the breast for each stimulating frequency. Based on these estimates, a scalar acoustic waveform at each grid point is estimated via GLC at the second stage. At the final stage, the energy of the estimated acoustic waveform at each grid point is computed and referred to as the image intensity.

To demonstrate the performance of AMART, we consider a 2-D inhomogeneous breast model developed in two steps. The electromagnetic field inside the breast model is simulated in the first step, and the second step is for the acoustic wave simulation. The finite-difference time-domain (FDTD) method is used in both steps for simulation [11-14]. The breast model is a 10 cm in diameter semicircle, which includes the skin, breast fatty tissue, glandular tissues, and the chest wall. In the following example, the thermal acoustic signals are simulated based on the aforementioned breast model for multiple stimulating frequencies from 500-800 MHz with frequency step 100 MHz. Two small 1.5-mm-diameter tumors are set inside the breast model. Their locations are at (X=70 mm, Y=60 mm) and (X=75 mm, Y=62.5 mm). The distance between the two tumors is 4 mm. For comparison purposes, the DAS method is applied to the same data set. We also present the imaging results obtained by ASART at different stimulating frequencies. Figures 4(a) and 4(b) show the imaging results for AMART and DAS, respectively. The two tumors are seen clearly in the AMART image, and the sizes and the locations of the two tumors are accurate. On the other hand, the DAS image contains much clutter and cannot show the two tumors clearly. Figures 4(c) and 4(d) show the imaging results for ASART at the stimulating frequencies $f=500$ and 800 MHz, respectively. The tumors can be seen in both ASART images, but with more clutter showing up in Figures 4(c)

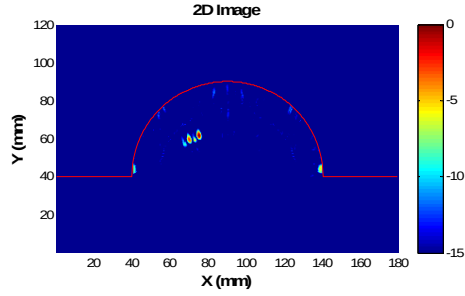
and 4(d) than in 4(a).



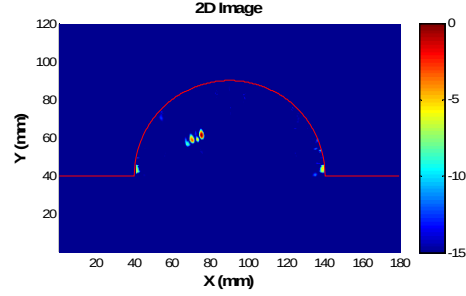
(a)



(b)



(c)



(d)

Figure 4: Reconstructed images obtained via (a) AMART, (b) DAS, (c) ASART at stimulating frequency $f=500$ MHz, (d) ASART at stimulating frequency $f=800$ MHz.

III. Key Research Accomplishments

- ✓ Numerical analysis about the electro-acoustic transduction in the phantom was conducted
- ✓ The entire excitation and detection system was set up and tested
- ✓ Noise floor was measured and compared to the signal level simulated. Some modifications of the whole system and phantoms are being implemented, aiming at higher SNR
- ✓ A 2-D inhomogeneous breast model was developed
- ✓ A user parameter-free robust adaptive image formation algorithm was developed

IV. Reportable Outcomes

Y. Xie, B. Guo, J. Li, G. Ku, and L. V. Wang, "Adaptive and Robust Methods of Reconstruction (ARMOR) for Thermoacoustic Tomography," IEEE Transactions on Biomedical Engineering, Vol. 55, No. 12, pp. 2741-2752, December 2008.

B. Guo, J. Li, H. Zmuda, and M. Sheplak, "Multi-Frequency Microwave Induced Thermal Acoustic Imaging for Breast Cancer Detection," IEEE Transactions on Biomedical Engineering, Vol. 54, No. 11, pp. 2000-2010, November 2007.

L. Du, J. Li, and P. Stoica, "User Parameter Free Approaches to Multistatic Adaptive Ultrasound Imaging," The 2008 IEEE International Symposium on Biomedical Imaging, Paris, France, May 2008.

L. Du, T. Yardibi, J. Li, and P. Stoica, "Review of User Parameter-Free Robust Adaptive Beamforming Algorithms," to appear in Digital Signal Processing – A Review Journal.

L. Du, T. Yardibi, J. Li, and P. Stoica, "Review of User Parameter-Free Robust Adaptive Beamforming Algorithms," 42nd Annual Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, October 2008.

L. Du, J. Li, and P. Stoica, "Fully Automatic Computation of Diagonal Loading Levels for Robust Adaptive Beamforming," to appear in IEEE Transactions on Aerospace and Electronic Systems.

J. Li, L. Du, and P. Stoica, "Fully Automatic Computation of Diagonal Loading Levels for Robust Adaptive Beamforming," The 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing, Las Vegas, NV, March 2008.

V. Conclusions

Numerical analysis about the electro-acoustic transduction of the phantom has been investigated, based on the Diebold's research [5], in order to get a more clear understanding about the entire picture. The major factors determining the frequency response of electro-acoustic transduction have been pointed out, and some material parameters are being measured based on specifically designed experiments.

The entire excitation and detection system has been set up, including modulated signal source, high power amplifier (100 W), low noise amplifier and data acquisition system. Relative programming has been completed and tested concerning the control of the entire system. The noise floor of the measurement system has been obtained and referred to the input of transducer, which provides the clear comparison between the signal level simulated and noise floor measured.

Phantoms are being redesigned to gain more conductivity and higher resonant frequency. Other modifications of the system are being conducted in order to enlarge the signal level and hence, get higher SNR.

A user parameter-free robust adaptive image formation algorithm, AMART, has been developed for a multifrequency TAI system. The new algorithm avoids the need to specify any user parameters. A numerical example based on a 2-D breast model has been provided to demonstrate the excellent performance of AMART: high resolution and good interference suppression capability.

VI. References

- [1] R. A. Kruger, *et. al.*, "Thermoacoustic CT with Radio Waves: A Medical Imaging Paradigm1," *Radiology*, 211, pp. 275–278, 1999.
- [2] L. V. Wang, *et. al.*, "Microwave-induced acoustic imaging of biological tissues," *Review of Scientific Instruments*, vol. 70, no. 60, 1999.
- [3] G. Ku and L. V. Wang, "Scanning thermoacoustic tomography in biological tissue," *Med. Phys.* 27 .5, May 2000.
- [4] R. A. Kruger, *et. al.*, "Breast Cancer in Vivo: Contrast Enhancement with Thermoacoustic CT at 434 MHz-- Feasibility Study," *Radiology*, 216, pp. 279-283, 2000.
- [5] G. J. Diebold and P. J. Westervelt, "The photoacoustic effect generated by a spherical droplet in a fluid," *J. Acoust. Soc. Am.* 84(6), December 1988.
- [6] R. R. Dagulo, *Ultrasonic Transducer Calibration*. Thesis of Bachelor of Science in Aerospace Engineering, Fall 2008.
- [7] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, pp. 1408-1418, August 1969.
- [8] J. Li and P. Stoica, eds., *Robust Adaptive Beamforming*. New York, NY: John Wiley & Sons, 2005.
- [9] L. Du, T. Yardibi, J. Li, and P. Stoica, "Review of user parameter-free robust adaptive beamforming algorithms," submitted to *Digital Signal Processing – A Review Journal*.
- [10] L. Du, J. Li, and P. Stoica, "Fully automatic computation of diagonal loading levels for robust adaptive beamforming," to appear in *IEEE Transactions on Aerospace and Electronic Systems*.
- [11] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Boston, MA: Artech House, second edition, 2000.
- [12] D. M. Sullivan, *Electromagnetic Simulation Using the FDTD Method*. New York, NY: IEEE Press, 2000.
- [13] J. G. Maloney and K. E. Cummings, "Adaptation of FDTD techniques to acoustic modeling," *11th Annual Review of Progress in Applied Computational Electromagnetics, Monterey, CA*, vol. 2, pp. 724-731, 1995.
- [14] X. Yuan, D. Borup, J. Wiskin, M. Berggren, and S. A. Johnson, "Simulation of acoustic wave propagation in dispersive media with relaxation losses by using FDTD method with PML absorbing boundary condition," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 46, pp. 14-23, January 1999.

VII. Appendices

“Adaptive and Robust Methods of Reconstruction (ARMOR) for Thermoacoustic Tomography”

“Multi-Frequency Microwave Induced Thermal Acoustic Imaging for Breast Cancer Detection”

“User Parameter Free Approaches to Multistatic Adaptive Ultrasound Imaging”

“Review of User Parameter-Free Robust Adaptive Beamforming Algorithms”

“Fully Automatic Computation of Diagonal Loading Levels for Robust Adaptive Beamforming”

Adaptive and Robust Methods of Reconstruction (ARMOR) for Thermoacoustic Tomography

Yao Xie*, *Student Member, IEEE*, Bin Guo, *Member, IEEE*, Jian Li, *Fellow, IEEE*,
Geng Ku, and Lihong V. Wang, *Fellow, IEEE*

Abstract—In this paper, we present new adaptive and robust methods of reconstruction (ARMOR) for thermoacoustic tomography (TAT), and study their performances for breast cancer detection. TAT is an emerging medical imaging technique that combines the merits of high contrast due to electromagnetic or laser stimulation and high resolution offered by thermal acoustic imaging. The current image reconstruction methods used for TAT, such as the delay-and-sum (DAS) approach, are data-independent and suffer from low-resolution, high sidelobe levels, and poor interference rejection capabilities. The data-adaptive ARMOR can have much better resolution and much better interference rejection capabilities than their data-independent counterparts. By allowing certain uncertainties, ARMOR can be used to mitigate the amplitude and phase distortion problems encountered in TAT. The excellent performance of ARMOR is demonstrated using both simulated and experimentally measured data.

Index Terms—Array signal processing, biomedical acoustic imaging, robustness.

I. INTRODUCTION

Thermoacoustic tomography (TAT), the earliest investigation of which dates back to the 1980s [1], has recently attracted much interest with its great promise in a wide span of biomedical applications (see, e.g., [2]–[4]). Its physical basis lies in the contrast of the radiation absorption rate among different biological tissues. Due to the thermoacoustic effect, when a short electromagnetic pulse (e.g., microwave or laser) is absorbed by the tissue, the heating results in expansion that generates acoustic signals. In TAT, an image of the tissue absorption properties is reconstructed from the recorded thermoacoustic signals. Such an image may reveal the physiological and pathological status of the tissue, which can be useful in many

applications including breast cancer detection [5]. Compared with microwave imaging and ultrasound imaging, TAT combines their merits and possesses both fine imaging resolution and good spatial contrast properties [4].

Developing accurate and robust image reconstruction methods is one of the key challenges encountered in TAT. Various image reconstruction algorithms have been developed for TAT. By using Radon transformation on the TAT data function, reflectivity tomography reconstruction algorithms can be used for TAT image reconstruction [6]. Exact inverse solutions have been found for different scanning geometries in both the frequency domain [7], [8] and the time domain [9], [10]. Approximate reconstruction algorithms, such as the time-domain delay-and-sum (DAS) beamforming method [11], [12] and the optimal statistical approach [13], have also been proposed. However, a common assumption of these existing methods is that the surrounding tissue is acoustically homogeneous. This approximation is inadequate in many medical imaging applications. According to previous studies, the sound speed in human female breast varies widely from 1430 to 1570 m/s around the commonly assumed speed of 1510 m/s [14], [15]. The heterogeneous acoustic properties of biological tissues cause amplitude and phase distortions in the recorded acoustic signals, which can result in significant degradation in imaging quality [16].

In ultrasound tomography (UT), wavefront distortion due to heterogeneity of biological tissue has been studied extensively. Various wavefront correction methods have been proposed [17]. However, they are not highly effective at correcting severe amplitude distortions [18], and they usually involve complicated procedures. The problem in TAT is somewhat different from that in UT. In the breast UT, the amplitude distortion caused by refraction is more problematic than the phase distortion induced by acoustic speed variation. In TAT, however, even for the biological tissue, such as the breast tissue, with a relatively weak heterogeneity, phase distortion dominates amplitude distortion [16]. These unique features suggest that new adaptive and robust imaging techniques should be designed especially for TAT.

Time-domain approximate reconstruction algorithms, such as the DAS (weighted or unweighted) type of data-independent approaches have various applications in medical imaging. They need little prior information on the tissue for image reconstruction and can be fast and simple to implement to process the wideband acoustic signals. Although not based on the exact solution, they provide similar image qualities to those of the exact reconstruction algorithms. However, these data-independent methods tend to suffer from poor resolution and high-sidelobe-level problems. Data-adaptive approaches, such as the recently

Manuscript received May 8, 2007; revised October 22, 2007. First published February 25, 2008; current version published December 17, 2008. This work was supported in part by the National Institutes of Health (NIH) under Grant 1R41CA107903-1, in part by the U.S. Army Medical Command under Contract W81XWH-06-1-0389, and in part by the National Natural Science Foundation of China under Grant 60428101. Asterisk indicates corresponding author.

*Y. Xie was with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6130 USA. She is now with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA.

B. Guo was with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA. He is now with the Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708 USA.

J. Li is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA.

G. Ku and L. V. Wang are with the Department of Biomedical Engineering, School of Engineering and Applied Science, Washington University in St. Louis, St. Louis, MO 63130-4899 USA.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TBME.2008.919112

introduced robust Capon beamforming (RCB) method [19], can have much better resolution and much better interference rejection capability than their data-independent counterparts.

We propose adaptive and robust methods of reconstruction (ARMOR) based on RCB for TAT. ARMOR can be used to mitigate the amplitude and phase distortion problems in TAT by allowing certain uncertainties. Specifically, in the first step of ARMOR, RCB is used for waveform estimation by treating the amplitude distortion with an uncertainty parameter. In the second step of ARMOR, a simple, yet effective, peak searching method is used for phase distortion correction. Compared with other energy- or amplitude-based response intensity estimation methods, peak searching can be used to improve image quality with little additional computational costs. Moreover, since the acoustic pulse is usually bipolar: a positive peak, corresponding to the compression pulse, and a negative peak, corresponding to the rarefaction pulse [11], we can further enhance the image contrast in TAT by using the peak-to-peak difference as the response intensity for a focal point. We will demonstrate the excellent performance of ARMOR by using both data simulated on a 2-D breast model and data experimentally measured from mastectomy specimens.

The remainder of this paper is organized as follows. In Section II, we formulate the problem of interest. Sections III–V describe the first, second, and third steps of ARMOR, respectively. Examples based on simulated and real-world experimental data are presented in Section VI. Finally, Section VII provides the conclusions.

II. PROBLEM FORMULATION

Consider a TAT imaging system, as shown in Fig. 1(a). A stimulating electromagnetic (laser or microwave) pulse is absorbed by the biological tissue under testing, which causes a sudden heat change (of the order of 10^{-4} °C [20]). Due to the thermoacoustic effect, an acoustic pulse is generated that can be recorded by an ultrasonic transducer array. The transducer array may be a real aperture array or a synthetic aperture array formed by rotating a sensor around the tissue and recording the acoustic waves at different locations. We assume that the number of transducers in the array (or in the synthetic aperture array case, the number of transducer data acquisition locations) is M . Each transducer is assumed to be omnidirectional; mutual couplings among the transducers are not considered in our model as they can be tolerated by our robust algorithms to a certain extent. The recorded acoustic signals are sufficiently sampled and digitized and a typical recorded pulse is shown in Fig. 1(b) (based on the data measured on the breast specimen II described in Section VI).

The data model for the sampled and digitized acoustic signal recorded by the m th transducer is given by:

$$x_m(n) = s_m(n) + \tilde{e}_m(n), \quad m = 1, \dots, M. \quad (1)$$

where n is the discrete time index, starting from t_0 after the excitation pulse. The scalar $s_m(n)$ denotes the signal component, which corresponds to the acoustic pulse generated at a focal point, and $\tilde{e}_m(n)$ is the residual term, which includes unmod-

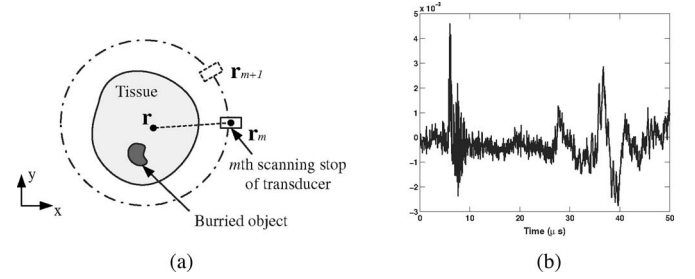


Fig. 1. (a) A schematic of a 2-D synthetic-aperture-based TAT scanning system. (b) A typical acoustic pulse recorded by a transducer (for data measured from breast specimen II).

eled noise and interference (caused by other sources within the tissue).

The goal of ARMOR is to reconstruct an image of thermoacoustic response intensity $I(\mathbf{r})$, which is directly related to the absorption property of the tissue, from the recorded data set $\{x_m(n)\}$. Herein, the (2-D or 3-D) vector $\mathbf{r}$ denotes the focal point location coordinate. To form an image, we scan the focal point location $\mathbf{r}$ to cover the entire cross section of the tissue (the transducers can acquire signals at different heights; for each height, a 2-D cross-sectional image can be reconstructed and a 3-D image can be formed from the 2-D images). We allow certain uncertainties in ARMOR to deal with amplitude and phase distortions caused by the background heterogeneity.

The discrete arrival time of the pulse (for the m th transducer) can be determined approximately as

$$t_m(\mathbf{r}) = \left\lfloor -\frac{t_0}{\Delta t} + \frac{\|\mathbf{r} - \mathbf{r}_m\|}{\Delta t v_0} \right\rfloor. \quad (2)$$

We will omit the dependence of the arrival time $t_m(\mathbf{r})$ on $\mathbf{r}$ hereafter for notational simplicity. Here, Δt is the sampling interval, and the 3-D vector $\mathbf{r}_m$ denotes the location of the m th transducer. The sound speed v_0 is chosen to be the average sound speed of the biological tissue under interrogation. The notation $\|\mathbf{x}\|$ denotes the Euclidean norm of $\mathbf{x}$, and $\lfloor y \rfloor$ stands for rounding to the greatest integer less than y . The second term in (2) represents the time-of-flight between the focal point and the m th transducer.

The signal components $\{s_m(n)\}_{m=1}^M$ are approximately scaled and shifted versions of a nominal waveform $s(t)$ at the source

$$s_m(n) \approx \frac{\exp(-\alpha \|\mathbf{r} - \mathbf{r}_m\|)}{\|\mathbf{r} - \mathbf{r}_m\|} \cdot s(n - t_m) \quad (3)$$

where α is the attenuation coefficient in Nepers/m. In TAT, the major frequency components of the acoustic signals take a relatively narrow band, and are usually lower than those in UT [16]. Hence, we can approximate α as a frequency-independent constant.

We preprocess the data to time delay all the signals from the focal point $\mathbf{r}$ and compensate for the loss in amplitude due to propagation decay. Let $y_m(n)$ denote the signal after preprocessing to backpropagate the detected signal to the source

$$y_m(n) = \exp(\alpha \|\mathbf{r} - \mathbf{r}_m\|) \cdot \|\mathbf{r} - \mathbf{r}_m\| \cdot x_m(n + t_m). \quad (4)$$

Then, the received vector data model can be written as

$$\mathbf{y}(n) = \mathbf{a}_0 s(n) + \mathbf{e}(n), \quad n = -N, \dots, N \quad (5)$$

where $\mathbf{a}_0$ is the corresponding steering vector, which is approximately equal to $\bar{\mathbf{a}} = [1, \dots, 1]^T$, $\mathbf{y}(n) = [y_1(n), \dots, y_M(n)]^T$, $\mathbf{e}(n)$ represents the noise and interference term after preprocessing, and $(\cdot)^T$ denotes the transpose. Here, we define the time interval of interests for the signal $\mathbf{y}(t)$ to be from $-N$ to N , which means that we only take N samples before and after the approximate arrival time given in (2) for the focal point at $\mathbf{r}$. The value of N should be chosen large enough so that the interval from $-N$ to N covers the expected signal duration in the region of interest.

In reality, both the amplitude and the phase (or pulse arrival time) of the acoustic pulse will be distorted. A major cause for these distortions is the acoustically heterogeneous background. Amplitude distortion is mainly due to the interferences caused by multipath, which is inevitable in the heterogeneous medium: refraction occurs due to acoustic speed mismatch across the tissue interface; consequently, acoustic pulses arrived at the transducer will be via different routes and interfere with each other. On the other hand, phase distortion is mainly caused by the nonuniform sound speed. For example, in human female breast, the sound speed can vary from 1430 to 1570 m/s; therefore, the actual arrival time will fluctuate around the approximately calculated time given in (2). Moreover, an inaccurate estimate of t_0 (t_0 is aligned with the focal point's signal arrival time) and the transducer calibration error may also contribute to the phase distortion. Amplitude and phase distortion will blur the image, raise the image background noise level, lower the values of the object of interest, and, consequently, decrease the image contrast [16].

We mitigate the effects of these distortions by allowing $\mathbf{a}_0$ to belong to an uncertainty set centered at $\bar{\mathbf{a}}$ and by considering the signal arriving within the interval from $-N$ to N .

III. STEP I OF ARMOR: WAVEFORM ESTIMATION

The first step of ARMOR is to estimate the waveform of the acoustic pulse generated by the focal point at location $\mathbf{r}$, based on the data model in (5). It will appear that we have neglected the presence of phase distortion by using this data model in the first step. However, by allowing $\mathbf{a}_0$ to be uncertain, we can tolerate some phase distortions as well. This approximation causes little performance degradation to our robust algorithm.

Covariance-fitting-based RCB [21] is used to first estimate the steering vector $\mathbf{a}_0$, and use the estimated $\mathbf{a}_0$ to obtain an optimal beamformer weight vector for pulse waveform estimation. By assuming that the true steering vector lies in the vicinity of the nominal steering vector $\bar{\mathbf{a}}$, we consider the following optimization problem [19]

$$\begin{aligned} \max_{\sigma^2, \mathbf{a}_0} \quad & \text{subject to} \quad \hat{\mathbf{R}} - \sigma^2 \mathbf{a}_0 \mathbf{a}_0^T \succeq 0, \\ & \|\mathbf{a}_0 - \bar{\mathbf{a}}\|^2 \leq \varepsilon, \end{aligned} \quad (6)$$

where $\mathbf{A} \succeq 0$ means that the matrix $\mathbf{A}$ is positive semidefinite, σ^2 is the power of the signal of interest, and

$$\hat{\mathbf{R}} = \frac{1}{2N+1} \sum_{n=-N}^N \mathbf{y}(n) \mathbf{y}^T(n) \quad (7)$$

is the sample covariance matrix. The second constraint in (6) is a spherical uncertainty set; an elliptical uncertainty set can be used instead, if a tighter constraint is desirable [21].

The parameter ε in (6) determines the size of the uncertainty set and is a user parameter. To avoid the trivial solution of $\mathbf{a}_0 = 0$, we require that

$$\varepsilon < \|\bar{\mathbf{a}}\|^2. \quad (8)$$

It can be verified that the smaller the ε , the higher the resolution and the stronger the ability of RCB to suppress an interference that is close to the signal of interest, and that the larger the ε , the more robust RCB will be to tolerate distortions and small-sample-size problems caused by calculating $\hat{\mathbf{R}}$ in (7) from a finite number of data vectors or snapshots. When ε is close to M , RCB will perform like DAS. To attain high resolution and to effectively suppress interference, ε should be made as small as possible. On the other hand, the smaller the sample size N or the larger the distortions, the larger should ε be chosen [19]. Since the performance of RCB does not depend very critically on the choice of ε (as long as it is set to be a "reasonable value") [21], such qualitative guidelines are usually sufficient for making a choice of ε . We will investigate the effect of ε in Section VI. In our examples in Section VI, we choose certain reasonable initial values for ε , and then make some adjustments empirically based on image quality: making it smaller when the resulting images have low resolution, or making it larger when the image is distorted by interferences.

By using the Lagrange multiplier method, the solution to (6) is given by [19]

$$\hat{\mathbf{a}}_0 = \bar{\mathbf{a}} - [\mathbf{I} + \mu \hat{\mathbf{R}}]^{-1} \bar{\mathbf{a}} \quad (9)$$

where $\mathbf{I}$ is the identity matrix, and $\mu \geq 0$ is the corresponding Lagrange multiplier that can be solved from the following equation

$$\|(\mathbf{I} + \mu \hat{\mathbf{R}})^{-1} \bar{\mathbf{a}}\|^2 = \varepsilon. \quad (10)$$

Consider the eigendecomposition on the sample covariance matrix $\hat{\mathbf{R}}$

$$\hat{\mathbf{R}} = \mathbf{U} \mathbf{\Gamma} \mathbf{U}^T \quad (11)$$

where the columns of $\mathbf{U}$ are the eigenvectors of $\hat{\mathbf{R}}$ and the diagonal matrix $\mathbf{\Gamma}$ consists of the corresponding eigenvalues $\gamma_1 \geq \gamma_2 \geq \dots \geq \gamma_M$. Let $\mathbf{b} = \mathbf{U}^T \bar{\mathbf{a}}$, where b_m denotes its m th element. Then, (10) can be rewritten as

$$\mathcal{L}(\mu) = \sum_{m=1}^M \frac{|b_m|^2}{(1 + \mu \gamma_m)^2} = \varepsilon. \quad (12)$$

Note that $\mathcal{L}(\mu)$ is a monotonically decreasing function of μ , with $\mathcal{L}(0) > \varepsilon$ by (8) and $\lim_{\mu \rightarrow \infty} \mathcal{L}(\mu) = 0 < \varepsilon$, which means that μ can be solved efficiently, say, by using the Newton's method

(see [19] for more details). After obtaining the value of μ , the estimate $\hat{\mathbf{a}}_0$ of the actual steering vector $\mathbf{a}_0$ is determined by (9).

Observe that there is a “scaling ambiguity” in (6) by treating both the signal power σ^2 and the steering vector $\mathbf{a}_0$ as unknowns (see [19] and [21]). The ambiguity exists in the sense that $(\sigma^2, \mathbf{a}_0)$ and $(\sigma^2/c, c^{1/2}\mathbf{a}_0)$ (for any constant $c > 0$) yield the same term $\sigma^2 \mathbf{a}_0 \mathbf{a}_0^T$. To eliminate this ambiguity, we scale the solution $\hat{\mathbf{a}}_0$ to make its norm satisfy the following condition

$$\|\hat{\mathbf{a}}_0\|^2 = M. \quad (13)$$

(Note that $M = \|\bar{\mathbf{a}}\|^2$.)

To obtain an estimate for the signal waveform $s(n)$, we apply a weight vector to the preprocessed signals $\{\mathbf{y}(n)\}_{n=-N}^N$. The weight vector is determined by using the estimated steering vector $\hat{\mathbf{a}}_0$ in the weight vector expression of the standard Capon beamformer (see, e.g., [19] and [21])

$$\hat{\mathbf{w}}_{\text{RCB}} = \frac{\|\hat{\mathbf{a}}_0\|}{M^{1/2}} \cdot \frac{\left[\hat{\mathbf{R}} + \frac{1}{\mu}\mathbf{I}\right]^{-1} \bar{\mathbf{a}}_0}{\bar{\mathbf{a}}_0^T \left[\hat{\mathbf{R}} + \frac{1}{\mu}\mathbf{I}\right]^{-1} \hat{\mathbf{R}} \left[\hat{\mathbf{R}} + \frac{1}{\mu}\mathbf{I}\right]^{-1} \bar{\mathbf{a}}_0}. \quad (14)$$

Note that (14) has a diagonal loading form, which allows the sample covariance matrix to be rank-deficient. The beamformer output can be written as

$$\hat{s}_{\text{RCB}}(n) = \hat{\mathbf{w}}_{\text{RCB}}^T \mathbf{y}(n), \quad n = -N, \dots, N \quad (15)$$

which is the waveform estimate for the acoustic pulse generated at the focal point at location $\mathbf{r}$.

RCB can provide a much better waveform estimate than the conventional DAS but at a higher computational cost. For a single focal point, RCB requires $O(M^3)$ flops, which mainly come from the eigendecomposition of the sample covariance matrix $\hat{\mathbf{R}}$ [19]; DAS needs only $O(M)$ flops. DAS can be used as a fast image reconstruction method to provide initial imaging results.

The weight vector used by DAS for waveform estimation is

$$\hat{\mathbf{w}}_{\text{DAS}} = \bar{\mathbf{a}} \quad (16)$$

and the estimated waveform is given by

$$\hat{s}_{\text{DAS}}(n) = \hat{\mathbf{w}}_{\text{DAS}}^T \mathbf{y}(n) = \sum_{m=1}^M y_m(n), \quad n = -N, \dots, N. \quad (17)$$

IV. STEP II OF ARMOR: PEAK SEARCHING

Based on the estimated waveform obtained in Step I for the focal point at location $\mathbf{r}$, in Step II of ARMOR, we will search for the two peaks of the bipolar acoustic pulse generated by the focal point. In a homogeneous background, where phase distortion is absent, we can accurately calculate the arrival time of the acoustic pulse generated by the focal point at location $\mathbf{r}$ by using (2). However, this is never true in heterogeneous biological tissues. It was reported in [16] that when the heterogeneity is weak, such as in the breast tissue, amplitude distortion caused by multipath is not severe. We can assume that the original peak remains a peak in the waveform estimated from Step I of ARMOR.

The bipolar acoustic pulse has one peak positive and another negative. We determine the positive and negative peak values as follows:

$$P^+ = \max \left\{ \max_{n \in [-\Delta, \Delta]} \hat{s}(n) \right\}, \quad (18)$$

$$P^- = \min \left\{ \min_{n \in [-\Delta, \Delta]} \hat{s}(n) \right\}, \quad (19)$$

where the searching range $[-\Delta, \Delta] \in [-N, N]$ is around the calculated arrival time given by (2). Here Δ is a user parameter. Since the peak searching is independent of the particular waveform estimation methods, we use $\hat{s}(n)$ to denote the waveform estimated by either DAS or ARMOR.

The search range is determined by the difference between the true arrival time t_m and the calculated arrival time $\bar{t}_m$, based on (2). This arrival time difference has been analyzed for breast tissue by taking into account its relatively weak heterogeneity acoustic property [16]. An expression for this difference is given in [16] by

$$\delta_m(\mathbf{r}') = \bar{t}_m - t_m \propto \frac{[v(\mathbf{r}') - v_0]}{v_0} \quad (20)$$

where $\mathbf{r}'$ is a point within the line connecting the focal point at location $\mathbf{r}$ and the m th transducer at location $\mathbf{r}_m$, and $v(\mathbf{r}')$ is the local sound speed. The higher order terms of $[v(\mathbf{r}') - v_0]/v_0$ in (20) have been ignored. It is reasonable to assume that $v(\mathbf{r}')$ is Gaussian-distributed with mean v_0 and variance σ_v^2 . Consequently, the arrival time difference is also Gaussian-distributed with zero-mean and variance $\sigma_\delta^2 \propto \sigma_v^2/v_0^2$. If we choose $\Delta = \sigma_\delta$, and the duration of the acoustic pulse is τ , we can find the two peaks of the pulse within the interval $(-\sigma_\delta, \sigma_\delta + \tau)$ on the recorded signals with a high probability of 0.6826. This analysis is consistent with the experimental measurements in [22]. From our examples, we found that a symmetric range $[-\Delta, \Delta]$ around the estimated arrival time performs similarly to the asymmetric range $[-\Delta, \Delta + \tau]$, and we use the former since it is easy to handle in practice. Also, we can use similar techniques as those in [22] to estimate σ_δ to find a good searching range for Step II of ARMOR, and to estimate τ for the energy-type methods, as shown in our examples later.

There is a tradeoff in choosing the searching range. The larger the searching range, the higher the probability we can find the peaks of the acoustic pulse within the range. However, if the range is chosen too large, the interferences may cause false peaks, and as a consequence, we are more likely to find a false peak. In our examples in Section VI, we choose the best searching range empirically based on the estimated variance of the arrival time difference $\hat{\sigma}_\delta$.

V. STEP III OF ARMOR: INTENSITY CALCULATION

After estimating the waveform generated by the focal point at location $\mathbf{r}$, we need to obtain the response intensity based on the estimated waveform. For the same estimated waveform, different approaches can be used to evaluate the focal point response intensity. These approaches extract different information from

the estimated waveform as the response intensity, and may be useful to physicians in different ways.

There are two major types of response intensity measurement approaches: amplitude-based and energy-based. The waveform peak values obtained in Step II of ARMOR can be used for both approaches.

Conventional DAS uses the amplitude-based measure for TAT imaging [11], [12], with the corresponding response intensity given by $\hat{s}(0)$, or equivalently

$$I_C = \hat{s}(0) = \sum_{m=1}^M y_m(0) \quad (21)$$

where the subscript “c” stands for “Conventional.”

The energy-based measure, such as the one used in [23], calculates the response intensity as follows

$$I_{E1} = \hat{s}^2(0) = \left[\sum_{m=1}^M y_m(0) \right]^2 \quad (22)$$

where the subscript “E1” means “Energy-type 1.”

The entire pulse energy has also been used as an intensity measure, such as in the monostatic and multistatic microwave imaging for breast cancer detection [24], [25], and the intensity is given by

$$I_{E2} = \sum_{n=0}^{\tau} \hat{s}^2(n) = \sum_{n=0}^{\tau} \left[\sum_{m=1}^M y_m(n) \right]^2, \quad (23)$$

where the subscript “E2” stands for “Energy-type 2.”

We can consider using the peak value as the response intensity measure due to the bipolar nature of the response at the focal point

$$I_P = \begin{cases} P^+, & \text{if } |P^+| \geq |P^-| \\ P^-, & \text{otherwise} \end{cases} \quad (24)$$

where the subscript “p” stands for “Peak,” with P^+ and P^- defined in (18) and (19), respectively. Herein, we keep the sign of the maximum amplitude since the sign of the peak may also contain some information about the focal point.

Peak searching maximizes the output signal-to-noise ratio. An intuitive explanation is that, given the fact that the acoustic pulse is bipolar [11], if we assume that the residual term $e(t)$ is stationary, or its power is uniform over time, then the signal-to-noise ratio (SNR) is maximized at the (positive or negative) peak of the acoustic pulse. As a comparison, the conventional DAS (21) fixes the samples to be summed up at the calculated arrival time. Due to phase distortions, the waveform at the calculated time may be far from the peak value.

We can also employ peak-to-peak difference as the response intensity for the focal point at location $\mathbf{r}$

$$I_{PP} = P^+ - P^- \geq 0 \quad (25)$$

where the subscript “PP” denotes the “peak-to-peak difference.” Peak-to-peak difference has higher imaging contrast than peak value measure: the peak-to-peak difference of the bipolar pulse is approximately twice the absolute peak value, which means that the output signal power of the former is four times of the

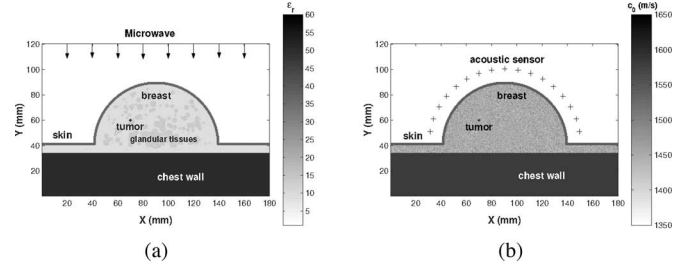


Fig. 2. 2-D breast model in an x - y coordinate system, with a 2-mm-diameter tumor present. (a) Model for electromagnetic simulation. (b) Model for acoustic simulation.

latter; yet, the noise power of the former may be only twice that of the latter. Therefore, the output SNR may be doubled by using the peak-to-peak difference rather than the peak value. Both peak-value and peak-to-peak difference measures belong to the amplitude-based measures.

VI. NUMERICAL AND EXPERIMENTAL EXAMPLES

We demonstrate the performance of ARMOR using both numerically simulated and experimentally measured TAT data. The ARMOR images are compared with the DAS images.

A. Numerical Examples

We consider a 2-D breast model, as shown in Fig. 2. The 2-D breast model includes 2-mm thick skin, chest wall, as well as randomly distributed fatty breast tissues and glandular tissues. The cross section of the breast model is a half-circle with a 10 cm diameter. In the first numerical example, a 2-mm-diameter tumor is located at 2.2 cm below the skin (at $x = 7.0$ cm, $y = 6.0$ cm). Fig. 2 shows the shape, dielectric properties, and sound speed variations of the breast model, as well as the tumor size and location for the first example. In the second numerical example, one large tumor (1 cm in diameter) is located at $x = 12$ cm, $y = 15$ cm. Other properties of the breast model for the second example are the same as those for the first example.

To reduce the reflections from the skin, the breast model is immersed in a lossless liquid with permittivity similar to that of the breast fatty tissue. Seventeen transducers (assumed omnidirectional) are located on a half-circle 10 mm away from the skin, with uniform spacing, to form a real aperture array.

The dielectric properties of the breast tissues are assumed to be Gaussian random variables with variations of $\pm 10\%$ around their nominal values. This variation represents the upper bound reported in the literature. The nominal values are chosen to be typical of those reported in the literature [5], [26], which is given in Table I [24]. The dielectric constants of glandular tissues are between $\epsilon_r = 11$ and $\epsilon_r = 15$. The dispersive properties of the fatty breast tissue and those of the tumor are also considered in the model. The randomly distributed breast fatty tissues and glandular tissues with variable dielectric properties are representative of the nonhomogeneity of the breast of an actual patient.

TABLE I
ACRONYMS

ART	Adaptive and Robust Methods Of Reconstruction
DAS	Delay-And-Sum
FDTD	Finite Difference Time Domain
PML	Perfectly matched layer
RCB	Robust Capon Beamforming
SNR	Signal-to-Noise Ratio
SAR	Specific Absorption
TAT	Thermoacoustic Tomography
UT	Ultra-sound Tomography
C	Conventional
E1	Energy-type 1
E2	Energy-type 2
P	Peak
PP	Peak-to-Peak difference

Following the report that the breast tissues have a weak acoustic heterogeneity [16], we model the sound speed within the breast as a Gaussian random variable with variation $\pm 5\%$ around the assumed average sound speed of 1500 m/s. Since the attenuation coefficient α in (3) is small for breast tissue (0.75 dB/(MHz·cm)) [15] and the acoustic signals are below 2 MHz, we neglect the exponential attenuation in acoustic wave propagation. Also, since the acoustic pressure field generated by the thermoacoustic effect is usually small [20], we do not consider the nonlinear acoustic effects. The probing microwave pulse used here is a modulated rectangular pulse with a modulating frequency of 800 MHz. The duration of the pulse is 1 μ s. More details about the thermal acoustic simulations are given in the Appendix. In the following, all the images are displayed on a linear scale, and we will name the imaging methods by their waveform estimation method followed by the intensity calculation approach, such as “DAS-C.”

Note that the skin also absorbs microwave energy and generates acoustic signals. The skin response is much stronger than that of the tumor, since the skin has a much larger area than the tumor and the skin is closer to the acoustic sensors. So, before applying the aforementioned preprocessing steps and ARMOR, we remove the strong skin response using techniques similar to those in [24]. A calibration signal is obtained as the average of the recorded signals containing similar skin response. Then, the calibration signal is subtracted out from all recorded signals to remove the skin response as much as possible.

The searching range is chosen by the guidelines presented in Section IV. To obtain a general profile of the arrival time difference caused by the phase distortion, we use a simple method similar to the one used in [27]. First, the cross-correlation functions for all the signals recorded by the two adjacent transducers are obtained. The peak value of the cross-correlation function is used to estimate the arrival time delay between the signals recorded by the adjacent transducers. Second, these arrival time delays are fitted using a fourth-order polynomial curve, which is dominated by the arrival time delays due to the path length differences in the absent of phase distortions. The fourth-order polynomial is used since the delay caused by the path length difference should vary smoothly [27]. Fig. 3(a) shows the estimated arrival time delay and the delay based on curve fitting.

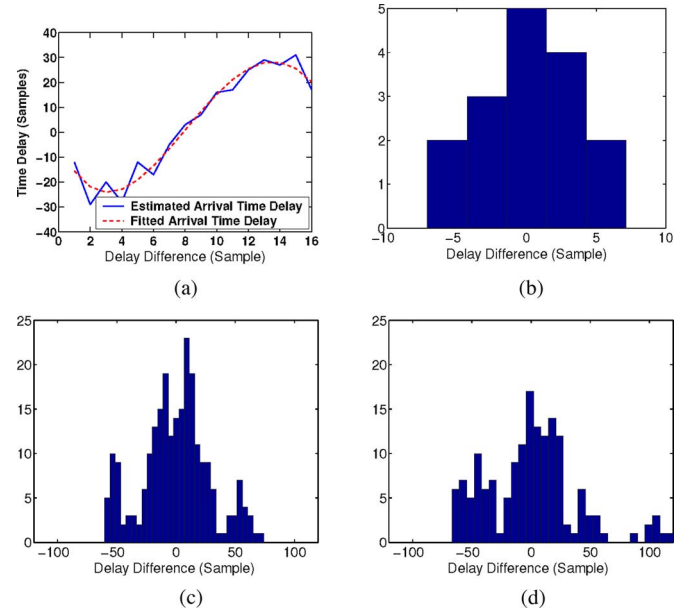


Fig. 3. (a) Comparison between the estimated and fitted arrival time delays, for the simulated breast model with one tumor (the curves for the two-tumor case are similar). Histograms of delay differences. (b) Simulated breast model with one tumor. (c) Breast specimen I. (d) Breast specimen II.

Third, the delay difference between the estimated arrival time delay and the fitted delay, or the fitting error, is treated as the arrival time distortion for the transducers. The standard deviation of the delay difference is used to estimate σ_δ . Although the accuracy of the cross-correlation method is limited due to false peaks and jitter problems, it is sufficient to obtain a qualitative profile for σ_δ .

Fig. 3 gives the histogram of the delay difference for all the cases that we considered herein. For the simulated example, the standard deviation of the delay difference is 4.5, which indicates a weak phase distortion in the breast model. We set an initial value for Δ , based on the estimated $\hat{\sigma}_\delta$, and adjust the length of the searching range to achieve the best imaging result.

To estimate the pulse duration $\hat{\tau}$ (used in DAS-E2 and RCB-E2), we select several typical signals (with clear peaks) and take the average of their pulse durations. In practice, the acoustic pulse duration is determined by the probing pulse duration, size and shape of the tumor, as well as the transducer response.

Fig. 4 shows the images for the simulated breast model with one 2-mm diameter tumor formed using ARMOR and DAS. The tumor response is weak for such a small tumor. In these images, we use $\varepsilon = 0.1M$ and the searching range $[-14, 14]$. Fig. 4(a) corresponds to DAS-C, where the tumor is buried by interference and noise. In Fig. 4(b), DAS-E1 fails to detect the tumor. In Fig. 4(c), for DAS-E2, a shadow of the tumor can be seen. In Fig. 4(d), for RCB-E2, most of the clutters are cleared up but a strong clutter shows up near the chest wall. Fig. 4(e)–4(h) shows the results of peak searching; none of them have false tumors, which may be attributed to proper corrections of phase aberrations. Images produced by ARMOR-P in Fig. 4(f)

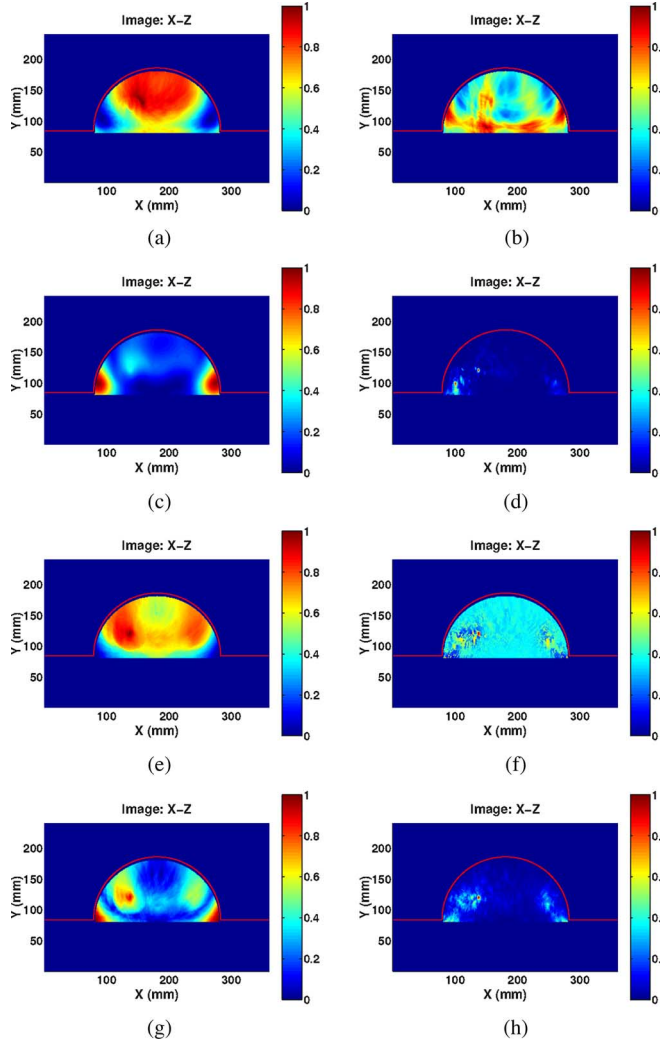


Fig. 4. Reconstructed images based on the 2-D simulated breast model with one 2-mm-diameter tumor. (a) DAS-C. (b) DAS-E1. (c) DAS-E2. (d) RCB-E2, with $\varepsilon = 0.1M$. (e) DAS-P. (f) ARMOR-P, with $\varepsilon = 0.1M$. (g) DAS-PP. (h) ARMOR-PP, with $\varepsilon = 0.1M$.

and by ARMOR-PP in Fig. 4(h) have lower sidelobe levels and higher resolutions, and the latter has a higher contrast than the former, due to the latter using the peak-to-peak difference as the intensity measure.

Fig. 5 shows the imaging results for the one large tumor (1 cm diameter) case. Here, we set $\varepsilon = 0.1M$ and the searching range $[-20, 20]$. (Note that different tumor sizes and locations will result in different sound speed variations in the breast model.) The white circle in the image corresponds to the actual contour of the tumor. Although all the methods can detect the tumor, only ARMOR can be used to form an image of the tumor with the best agreement with the actual tumor size and location.

By plotting a map (maps are not shown here due to limited space) of the values of μ used in ARMOR, for each focal point, we find that at the tumor locations, μ usually takes smaller values than that at other locations.

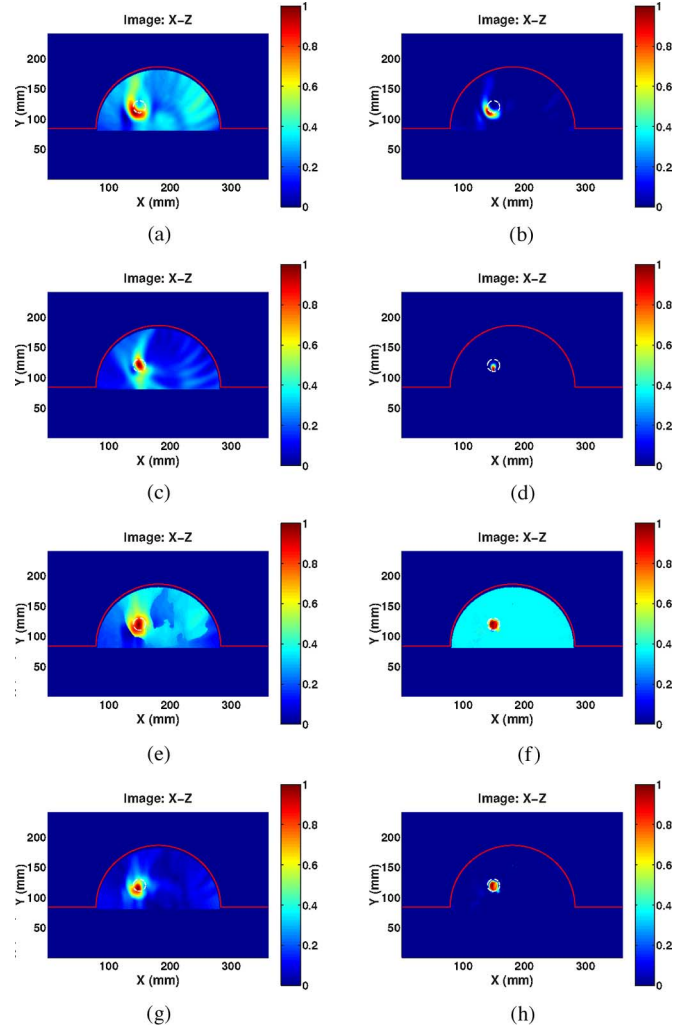


Fig. 5. Reconstructed images based on the 2-D simulated breast model with one large tumor (1 cm in diameter). The white circle in the image corresponds to the actual shape of the tumor. (a) DAS-C. (b) DAS-E1. (c) DAS-E2. (d) RCB-E2, with $\varepsilon = 0.1M$. (e) DAS-P. (f) ARMOR-P, with $\varepsilon = 0.1M$. (g) DAS-PP. (h) ARMOR-PP, with $\varepsilon = 0.1M$.

B. Experimental Results

We have also tested ARMOR and DAS on two sets of TAT experimental data from mastectomy specimens [4] obtained by the Optical Imaging Laboratory at the Texas A&M University.

The two data sets were acquired from mastectomy specimens using a TAT system. Microwave sources were used to heat the specimens transiently. In the experiment, the breast specimen was formed to a cylindrical shape inside a plastic bowl. The bowl was immersed in ultrasound coupling medium in a container. For breast specimen I, the acoustic signals were recorded at 240 equally spaced scanning stops on a circular track of radius 12.9 cm. The thickness of this specimen was about 4 cm in a round plastic bowl of 17 cm in diameter. This lesion was diagnosed as an invasive metaplastic carcinoma with chondroid and squamous metaplasia. The size of the tumor was measured to be 35 mm in diameter by TAT, and 36 mm in diameter by radiography (see [4] for details). For breast specimen II, the scanning radius was 9.7 cm, with 160 scanning stops. This specimen was

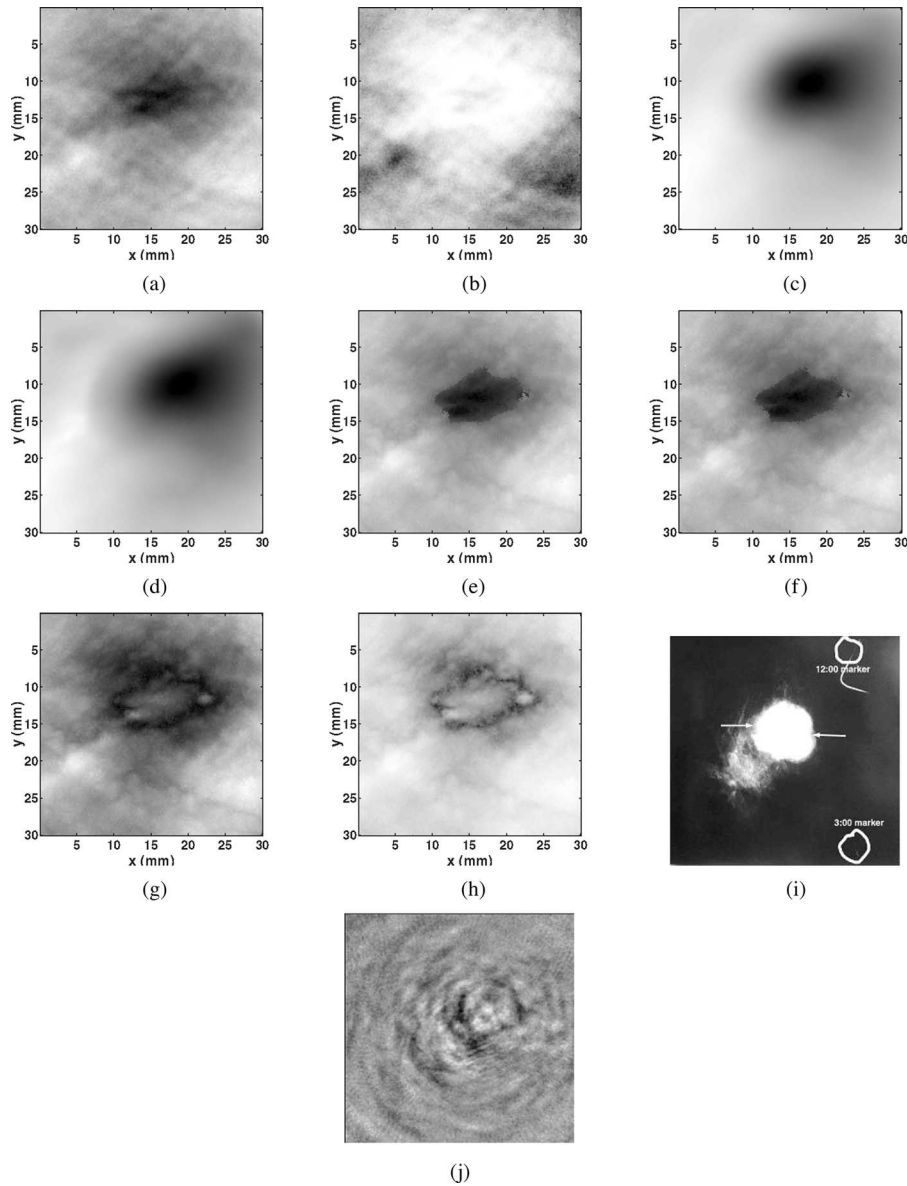


Fig. 6. Reconstructed images for breast specimen I. (a) DAS-C. (b) DAS-E1. (c) DAS-E2. (d) RCB-E2, with $\varepsilon = 0.5M$. (e) DAS-P. (f) ARMOR-P, with $\varepsilon = 0.5M$. (g) DAS-PP. (h) ARMOR-PP, with $\varepsilon = 0.5M$. (i) X-ray image. (j) Inverse solution.

9 cm thick in a round plastic bowl of 11 cm in diameter. The lesion in the specimen was diagnosed as infiltrating lobular carcinoma; the size of the tumor was about 20 mm $\times$ 12 mm on TAT image, and about 26 mm $\times$ 15 mm on the radiography (see [4] for more details).

First, we study the delay difference for both the breast specimens to get a qualitative guide for choosing the searching range in Step II of ARMOR. The results are shown in Fig. 3(c) and 3(d), respectively. Note that breast specimen II has a larger variance in delay differences than breast specimen I. In Fig. 3(c), 70% of the delay differences are roughly between -23 to 23 samples, whereas in Fig. 3(d), 70% of the delay differences are between -40 and 40 samples. Therefore we should set a larger searching range for breast specimen II than for breast specimen I.

Fig. 6 shows the reconstructed images for breast specimen I. In the following images, the searching range was set to $[-3, 3]$ after adjustment, and $\varepsilon = 0.5M$ for all the RCBs used herein. In Fig. 6(a), for DAS-C, the dark region shows a blurred object corresponding to the breast tumor. In Fig. 6(b), for DAS-E1, the light region shows a vague boundary of the tumor. Fig. 6(c), for DAS-E2, and 6(d), for RCB-E2, have similar performances. In Fig. 6(e), for DAS-P, and 6(f), for ARMOR-P, a dark region with a clear cut has a good correspondence with the location and shape of the tumor in the radiograph [4]. In Fig. 6(g), for DAS-PP, and 6(h), for ARMOR-PP, not only a clear image of the tumor is obtained, but also the detailed boundary is revealed. For comparison, the images from X-ray mammography, considered the “gold standard” of breast imaging, and the exact inverse solution of TAT (see [4] for more details) are shown in Fig. 6(i)

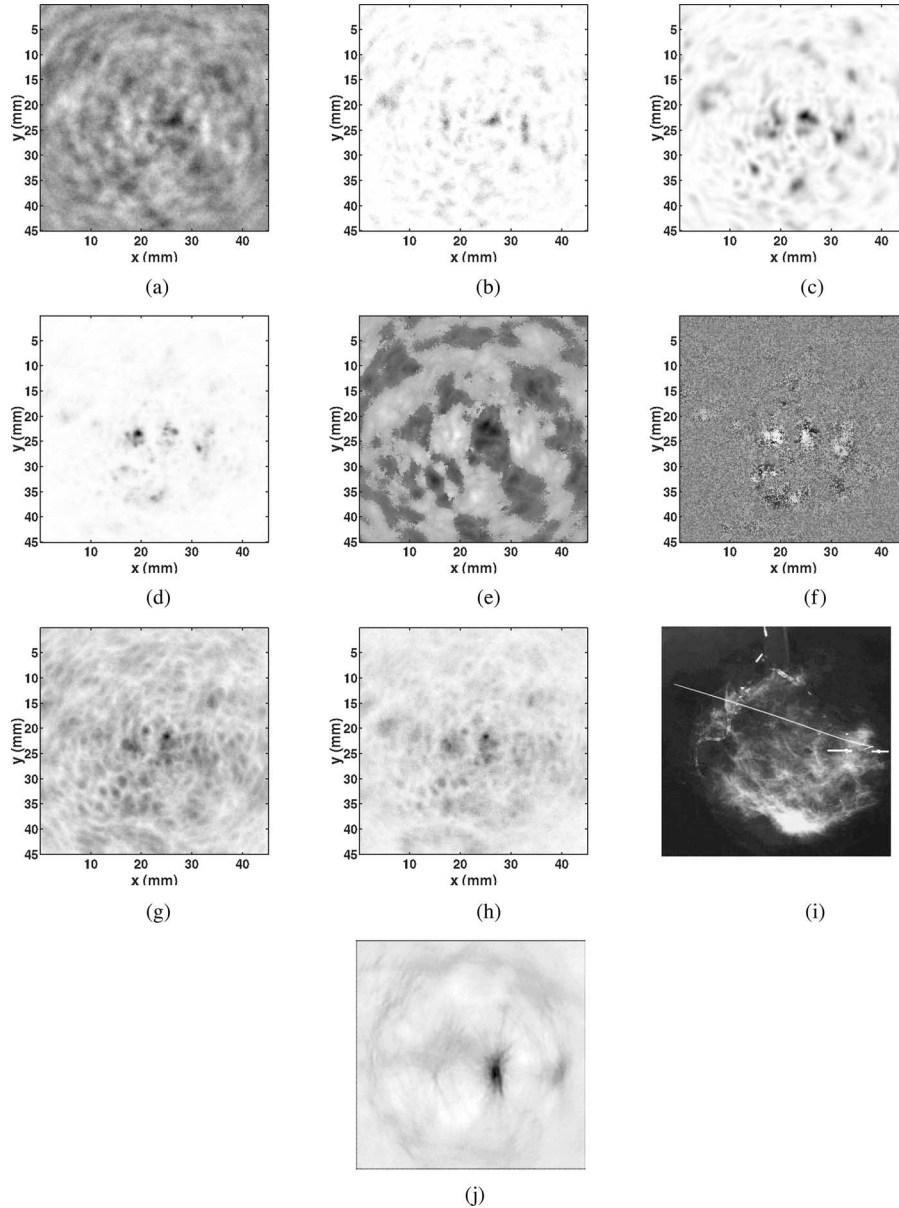


Fig. 7. Reconstructed images for breast specimen II. (a) DAS-C. (b) DAS-E1. (c) DAS-E2. (d) RCB-E2, with $\varepsilon = 0.5M$. (e) DAS-P. (f) ARMOR-P, with $\varepsilon = 0.5M$. (g) DAS-PP. (h) ARMOR-PP, with $\varepsilon = 0.5M$. (i) X-ray image. (j) Inverse solution.

and 6(j)), respectively. We give Fig. 6 and the following Fig. 7 in gray scale to have a better comparison with the X-ray images.

Fig. 7 shows the reconstructed images for breast specimen II. The tumor size here is smaller, and a high level of interference and noise is present in the recorded data. The searching interval is eventually adjusted to $[-120, 120]$ and RCB parameter $\varepsilon = 0.5M$. In Fig. 7(a), for DAS-C, the true tumor is barely identifiable from the surrounding clutters. The DAS-E1 shown in Fig. 7(b) and the DAS-E2 shown in Fig. 7(c) provide higher imaging contrast than DAS-C but show strong clutter. In Fig. 7(d), for RCB-E2, a false tumor shows up, which demonstrates the need for robustness in the presence of relatively strong phase distortion. DAS-P is shown in Fig. 7(e) and ARMOR-P is shown in Fig. 7(f). DAS-PP and ARMOR-PP produce the best images in Figs. 7(g) and 7(h), respectively, with the location and

shape of the tumor consistent with the radiograph in Fig. 7(i) [4]. If we define the signal-to-background ratio (SBR) (i.e., squaring the pixel values of the image, the ratio of the maximum to the total sum of the squared values) as an image quality measurement metric, ARMOR-PP has an SBR twice that of DAS-PP, which means a 3 dB gain for ARMOR-PP. For comparison, the image formed by the exact inverse solution of TAT (see [4] for more details) is shown in Fig. 7(j).

The effects of the uncertainty parameter ε in ARMOR is studied in our next example. We vary ε of RCB used in ARMOR. The imaging results for breast specimen I, shown in Fig. 8, are consistent with our previous analysis: when ε is large, the performance of RCB, in Fig. 8(a), is close to that of DAS in Fig. 6(g). When the parameter ε is small, as shown in Fig. 8(c), the resolution is improved at the cost of robustness.

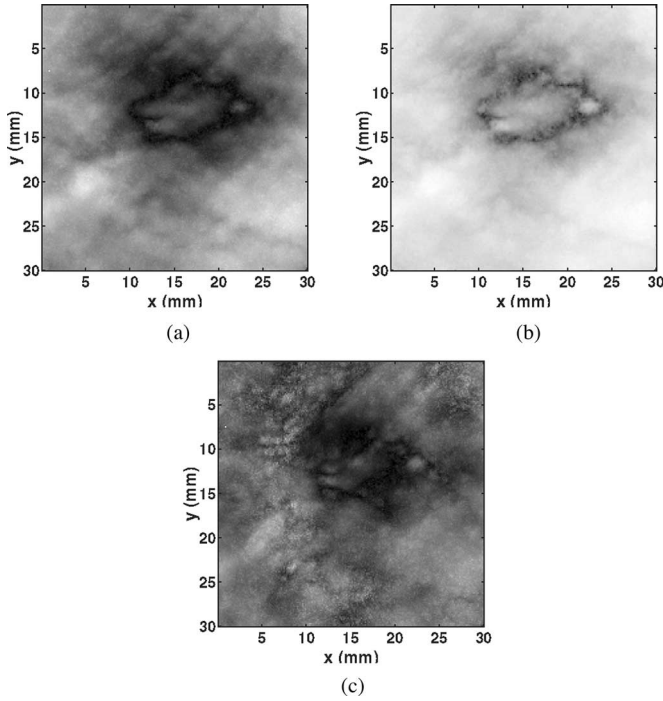


Fig. 8. Effects of uncertainty parameter ε on ARMOR-PP, with a searching range $[-3, 3]$. (a) $\varepsilon = 0.7M$. (b) $\varepsilon = 0.5M$. (c) $\varepsilon = 0.3M$.

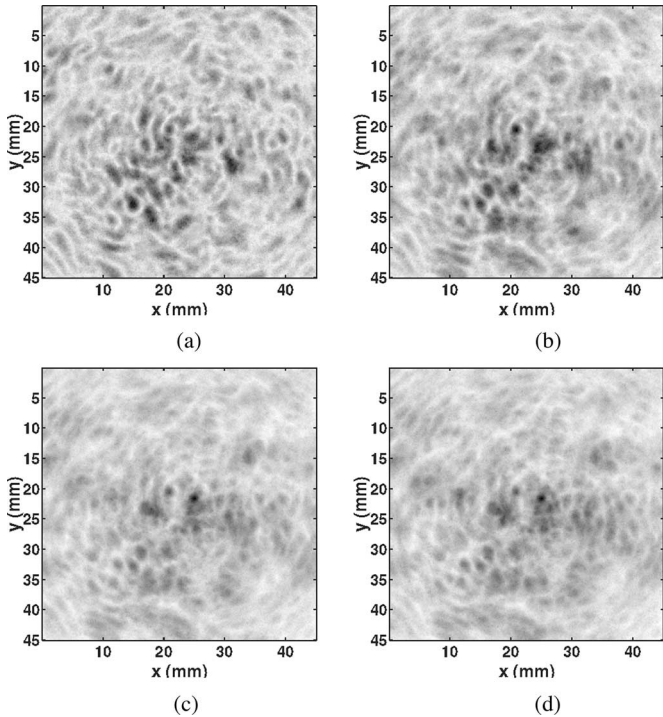


Fig. 9. Effects of the searching range on the DAS-PP images. (a) Searching range $[-20, 20]$. (b) Searching range: $[-40, 40]$. (c) Searching range $[-60, 60]$. (d) Searching range: $[-80, 80]$.

In our last example, the effect of the searching-range width on the imaging quality is considered. We use DAS-PP as an example since it shows more dependence on the searching range. The conclusion drawn for DAS applies to ARMOR. A symmetric searching range centered around the calculated arrival time is

used. From the discussions in Section IV, we know that there is a tradeoff in choosing the searching range. Clearly, when the searching range is too small, such as in Fig. 9(a), we miss the true peaks. With an increase in the searching range, the image quality becomes gradually better, as shown in Fig. 9(b) and 9(c). However, when the searching range passes a certain threshold, with too much interference coming into the searching range, the image quality degrades because of increased clutters, as shown in Fig. 9(d).

From our numerical examples, we conclude that ARMOR has higher resolution and better interference rejection capability and more robustness against wavefront distortion than DAS. Also, we find that the amplitude-based measures reveal more details of the tumor in the reconstructed images than their energy-based counterparts. The energy-based measures are not sensitive to phase distortions; however, they tend to blur the reconstructed images, causing loss of details with a low-pass filtering-like effect.

VII. CONCLUSION

ARMOR has been proposed for thermoacoustic tomography. ARMOR is robust to the amplitude and phase distortions in the recorded signals caused by the acoustic heterogeneity of biological tissues. ARMOR consists of three steps: in the first step, ARMOR uses the data-adaptive robust Capon beamforming (RCB) for waveform estimation; in the second step of ARMOR, a simple, yet effective, peak searching method is used to mitigate the phase distortion in the estimated waveform; in the third step, the response intensity is calculated for the focal point using various approaches, among which the peak-to-peak difference measure further enhances the image contrast. Examples based on a numerically simulated 2-D breast model and two sets of experimentally measured data from human mastectomy specimens demonstrate the excellent performance of ARMOR: high-resolution, low sidelobe level, and much improved interference suppression capability.

APPENDIX

THERMAL ACOUSTIC SIMULATIONS

We consider the microwave-induced thermal acoustic simulation in two steps. In the first step, the electromagnetic field inside the breast model is simulated and the specific absorption rate (SAR) distribution is calculated based on the simulated electromagnetic field. The second step is for the acoustic wave simulation, where the SAR distribution obtained in the first step is used as the acoustic pressure source through the thermal expansion coefficient. In both steps, the finite-difference time-domain (FDTD) method [28] is used for the simulation examples.

The 2-D electromagnetic breast model used is as shown in Fig. 2(a). A narrow electromagnetic pulse is used to irradiate the breast from the top of the model. The electromagnetic field is simulated using the FDTD method. The grid-cell size used by FDTD is $0.5 \text{ mm} \times 0.5 \text{ mm}$ and the computational region is terminated by perfectly matched layer (PML) absorbing boundary conditions [29].

TABLE II
NOMINAL DIELECTRIC PROPERTIES OF BREAST TISSUES [24]

Tissues	Dielectric Properties	
	Permittivity (F/m)	Conductivity (S/m)
Immersion Liquid	9	0
Chest Wall	50	7
Skin	36	4
Fatty Breast Tissue	9	0.4
Nipple	45	5
Glandular Tissue	11-15	0.4-0.5
Tumor	50	4

TABLE III
ACOUSTIC PARAMETERS FOR BIOLOGICAL TISSUES

Tissue	ρ (kg/m ³)	c (m/s)	α^* (dB/cm)	β (1/° C)	C_p (J/(° C · kg))
Breast	1020	1510	$0.75f^{1.5}$	3E-4	3550
Skin	1100	1537	3.5	3E-4	3500
Muscle	1041	1580	$0.57f$	3E-4	3510
Tumor	1041	1580	$0.57f$	3E-4	3510

* f is the acoustic frequency, and the unit is in megahertz.

The SAR distribution is given as [30]

$$\text{SAR}(\mathbf{r}) = \frac{\sigma(\mathbf{r})E^2(\mathbf{r})}{2\rho(\mathbf{r})} \quad (26)$$

where $\sigma(\mathbf{r})$ is the conductivity of the biological tissues at location $\mathbf{r}$, $E(\mathbf{r})$ is the electric field at location $\mathbf{r}$, and $\rho(\mathbf{r})$ is the mass density of the biological tissues at location $\mathbf{r}$.

In the microwave-induced TAI system, the microwave energy is small, and as a result, the acoustic pressure field induced by the microwave is also small. So, the nonlinear acoustic effect does not need to be considered in the TAI system. The two basic linear acoustic wave generation equations are [9]

$$\rho \frac{\partial}{\partial t} \mathbf{u}(\mathbf{r}, t) = -\nabla p(\mathbf{r}, t) \quad (27)$$

and

$$\nabla \cdot \mathbf{u}(\mathbf{r}, t) = -\frac{1}{\rho c^2} \frac{\partial}{\partial t} p(\mathbf{r}, t) + \alpha p(\mathbf{r}, t) + \beta \frac{\partial}{\partial t} T(\mathbf{r}, t) \quad (28)$$

where $\mathbf{u}(\mathbf{r}, t)$ is the acoustic velocity vector, $p(\mathbf{r}, t)$ is the acoustic pressure field, ρ is the mass density, α is the attenuation coefficient, β is the thermal expansion coefficient, and $T(\mathbf{r}, t)$ is the temperature. The values for these acoustic properties for different breast tissues are listed in Table III [25].

Because the duration of the microwave pulse is much shorter than the thermal diffusion time, thermal diffusion can be neglected [9], and the thermal equation is

$$C_p \frac{\partial}{\partial t} T(\mathbf{r}, t) = \text{SAR}(\mathbf{r}, t) \quad (29)$$

where C_p is the specific heat. Substituting (29) into (28) gives

$$\nabla \cdot \mathbf{u}(\mathbf{r}, t) = -\frac{1}{\rho c^2} \frac{\partial}{\partial t} p(\mathbf{r}, t) + \alpha p(\mathbf{r}, t) + \frac{\beta}{C_p} \text{SAR}(\mathbf{r}, t). \quad (30)$$

FDTD is used again to compute the thermal acoustic wave based on (27) and (30).

The breast model for the acoustic simulation is shown in Fig. 2(b), which is constructed similarly to the model for electromagnetic simulation. An acoustic sensor array deployed uni-

formly around the breast model is used to record the thermal acoustic signals. The grid-cell size used by the acoustic FDTD is 0.1 mm × 0.1 mm and the computational region is terminated by PML-absorbing boundary conditions. Note that the size of the FDTD cell for the acoustic simulation is much finer than that of the FDTD cell for the electromagnetic simulation because the wavelength of an acoustic wave is much smaller than that of a microwave. The SAR distribution data is interpolated to achieve a desired grid resolution for the acoustic breast model.

REFERENCES

- [1] R. G. Olsen and J. C. Lin, "Acoustic imaging of a model of a human hand using pulsed microwave irradiation," *Bioelectromagnetics*, vol. 4, pp. 397–400, 1983.
- [2] R. A. Kruger, P. Liu, Y. R. Fang, and C. R. Appledorn, "Photoacoustic ultrasound (PAUS)—reconstruction tomography," *Med. Phys.*, vol. 22, pp. 1605–1609, Oct. 1995.
- [3] R. A. Kruger, K. D. Miller, H. E. Reynolds, W. L. Kiser, D. R. Reinecke, and G. A. Kruger, "Breast cancer *in vivo*: Contrast enhancement with thermoacoustic CT at 434 MHz—feasibility study," *Radiology*, vol. 216, pp. 279–283, 2000.
- [4] G. Ku, B. D. Fornage, X. Jin, M. Xu, K. K. Hunt, and L. V. Wang, "Thermoacoustic and photoacoustic tomography of thick biological tissues toward breast imaging," *Technol. Cancer Res. Treat.*, vol. 4, pp. 1–7, Oct. 2005.
- [5] S. S. Chaudhary, R. K. Mishra, A. Swarup, and J. M. Thomas, "Dielectric properties of normal and malignant human breast tissues at radiowave and microwave frequencies," *Indian J. Biochem. Biophys.*, vol. 21, pp. 76–79, Feb. 1984.
- [6] M. A. Anastasio, J. Zhang, X. Pan, Y. Zou, G. Ku, and L. V. Wang, "Half-time image reconstruction in thermoacoustic tomography," *IEEE Trans. Med. Imag.*, vol. 24, no. 2, pp. 199–210, Feb. 2005.
- [7] Y. Xu, D. Feng, and L. V. Wang, "Exact frequency-domain reconstruction for thermoacoustic tomography: I. Planar geometry," *IEEE Trans. Med. Imag.*, vol. 21, no. 7, pp. 823–828, Jul. 2002.
- [8] Y. Xu, M. Xu, and L. V. Wang, "Exact frequency-domain reconstruction for thermoacoustic tomography: II. Cylindrical geometry," *IEEE Trans. Med. Imag.*, vol. 21, no. 7, pp. 829–833, Jul. 2002.
- [9] M. Xu and L. V. Wang, "Time-domain reconstruction for thermoacoustic tomography in a spherical geometry," *IEEE Trans. Med. Imag.*, vol. 21, no. 7, pp. 814–822, Jul. 2002.
- [10] M. Xu, Y. Xu, and L. V. Wang, "Time-domain reconstruction algorithms and numerical simulations for thermoacoustic tomography in various geometries," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 9, pp. 1086–1099, Sep. 2003.
- [11] C. G. A. Hoelen and F. F. M. de Mul, "Image reconstruction for photoacoustic scanning of tissue structures," *Appl. Opt.*, vol. 39, no. 31, pp. 5872–5883, 2000.
- [12] D. Feng, Y. Xu, G. Ku, and L. V. Wang, "Microwave-induced thermoacoustic tomography: Reconstruction by synthetic aperture," *Med. Phys.*, vol. 28, pp. 2427–2431, Dec. 2001.
- [13] Y. V. Zhulina, "Optimal statistical approach to optoacoustic image reconstruction," *Appl. Opt.*, vol. 39, no. 32, pp. 5971–5977, 2000.
- [14] G. Kossoff, E. K. Fry, and J. Jellins, "Average velocity of ultrasound in the human female breast," *J. Acoust. Soc. Amer.*, vol. 53, no. 6, pp. 1730–1736, 1973.
- [15] T. Szabo, *Diagnostic Ultrasound Imaging: Inside Out*. New York: Academic, Sep. 2004.
- [16] Y. Xu and L. V. Wang, "Effects of acoustic heterogeneity in breast thermoacoustic tomography," *IEEE Trans. Ultrason., Ferroelect. Freq. Control*, vol. 50, no. 9, pp. 1134–1146, Sep. 2003.
- [17] D. Liu and R. C. Waag, "Time-shift compensation of ultrasonic pulse focus degradation using least-mean-square error estimates of arrival time," *J. Acoust. Soc. Amer.*, vol. 95, pp. 542–555, Jan. 1994.
- [18] Q. Zhu and B. D. Steinberg, "Deaberration of incoherent wavefront distortion: An approach toward inverse filtering," *IEEE Trans. Ultrason., Ferroelect. Freq. Control*, vol. 44, no. 3, pp. 575–589, May 1997.
- [19] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.

- [20] M. Xu and L. V. Wang, "Pulsed-microwave-induced thermoacoustic tomography: Filtered backprojection in a circular measurement configuration," *Med. Phys.*, vol. 29, no. 8, pp. 1661–1669, 2002.
- [21] J. Li and P. Stoica, Eds., *Robust Adaptive Beamforming*. New York: Wiley, 2005.
- [22] Q. Zhu and B. D. Steinberg, "Wavefront amplitude distribution in the female breast," *J. Acoust. Soc. Amer.*, vol. 96, pp. 1–9, Jul. 1994.
- [23] X. Li and S. C. Hagness, "A confocal microwave imaging algorithm for breast cancer detection," *IEEE Microw. Wireless Compon. Lett.*, vol. 11, no. 3, pp. 130–132, Mar. 2001.
- [24] B. Guo, Y. Wang, J. Li, P. Stoica, and R. Wu, "Microwave imaging via adaptive beamforming methods for breast cancer detection," *J. Electromagn. Waves Appl.*, vol. 20, no. 1, pp. 53–63, 2006.
- [25] E. J. Bond, X. Li, S. C. Hagness, and B. D. Van Veen, "Microwave imaging via space-time beamforming for early detection of breast cancer," *IEEE Trans. Antennas Propag.*, vol. 51, no. 8, pp. 1690–1705, Aug. 2003.
- [26] C. Gabriel, R. W. Lau, and S. Gabriel, "The dielectric properties of biological tissues: II. Measured in the frequency range 10 Hz to 20 GHz," *Phys. Med. Biol.*, vol. 41, pp. 2251–2269, Nov. 1996.
- [27] Y. Sumino and R. Waag, "Measurements of ultrasonic pulse arrival time differences produced by abdominal wall specimens," *J. Acoust. Soc. Am.*, vol. 90, pp. 2924–2930, Dec. 1991.
- [28] A. Taflov and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 3rd ed. Boston, MA: Artech House, Jul. 2005.
- [29] S. D. Gedney, "An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD lattices," *IEEE Trans. Antennas Propag.*, vol. 44, no. 12, pp. 1630–1639, Dec. 1996.
- [30] P. Bernardi, M. Cavagnaro, S. Pisa, and E. Piuze, "SAR distribution and temperature increase in an anatomical model of the human eye exposed to the field radiated by the user antenna in a wireless LAN," *IEEE Trans. Microw. Theory Tech.*, vol. 46, no. 12, pp. 2074–2082, Dec. 1998.



Yao Xie (S'04) received the B.Sc. degree from the University of Science and Technology of China (USTC), Hefei, China, in 2004, and the M.Sc. degree from the University of Florida, Gainesville, in 2006, both in electrical engineering. She is currently working toward the Ph.D. degree at the Department of Electrical Engineering, Stanford University, Stanford, CA.

Her current research interests include optimization, information theory, and their applications in wireless communications, signal processing, and

medical imaging.

Ms. Xie is a member of Tau Beta Pi and Eta Kappa Nu. She was the first-place winner in the Student Best Paper Contest at the 2005 Annual Asilomar Conference on Signals, Systems, and Computers, for her work on breast cancer detection.



Bin Guo (S'05–M'07) received the B.E. and M.S. degrees in electrical engineering from Xian Jiaotong University, Xian, China, in 1997 and 2000, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Florida, Gainesville, in 2007.

Currently, he is a Research Associate in the Department of Electrical and computer Engineering, Duke University, Durham, NC. Prior to joining Duke University, he was a Research Assistant with the Department of Electrical and Computer Engineering,

University of Florida, from 2003 to 2007, and was an Associate Research Scientist with the Temasek Laboratories, National University of Singapore, Singapore, from 2002 to 2003. His current research interests include signal processing, microwave imaging, and computational electromagnetics.



Jian Li (S'87–M'91–SM'97–F'05) received the M.Sc. and Ph.D. degrees in electrical engineering from Ohio State University, Columbus, in 1987 and 1991, respectively.

From April 1991 to June 1991, she was an Adjunct Assistant Professor with the Department of Electrical Engineering, Ohio State University. From July 1991 to June 1993, she was an Assistant Professor with the Department of Electrical Engineering, University of Kentucky, Lexington. Since August 1993, she has been with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, where she is currently a Professor. Her current research interests include spectral estimation, statistical and array signal processing, and their applications.

Prof. Li is a Fellow of the Institution of Engineering and Technology (IET). She is a member of Sigma Xi and Phi Kappa Phi. She received the 1994 National Science Foundation Young Investigator Award and the 1996 Office of Naval Research Young Investigator Award. She is presently a member of two of the IEEE Signal Processing Society technical committees: the Signal Processing Theory and Methods (SPTM) Technical Committee and the Sensor Array and Multichannel (SAM) Technical Committee. She was an Executive Committee Member of the 2002 International Conference on Acoustics, Speech, and Signal Processing, Orlando, FL, in May 2002. She was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1999 to 2005, an Associate Editor of the IEEE SIGNAL PROCESSING MAGAZINE from 2003 to 2005, and a member of the Editorial Board of *Signal Processing*, a publication of the European Association for Signal Processing (EURASIP), from 2005 to 2007. She has been a Member of the Editorial Board of *Digital Signal Processing—A Review Journal*, a publication of Elsevier, since 2006. She is a coauthor of the papers that have received the First- and Second-Place Best Student Paper Award, respectively, at the 2005 and 2007 Annual Asilomar Conferences on Signals, Systems, and Computers in Pacific Grove, CA. She is also a coauthor of the paper that has received the M. Barry Carlton Award for the best paper published in IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS in 2005.

Geng Ku received the B.S. and M.S. degrees from Huazhong University of Science and Technology, Wuhan, China, and the Ph.D. degree from the Department of Biomedical Engineering, Texas A&M University, College Station.

He is a Research Faculty Member in the Department of Biomedical Engineering, Washington University at St. Louis, St. Louis. His current research interests include laser-based photoacoustic tomography and microwave-based thermoacoustic tomography.



Lihong V. Wang (M'96–SM'00–F'06) received the Ph.D. degree from Rice University, Houston, TX, under the tutelage of Dr. Robert Curl, Dr. Richard Smalley, and Dr. Frank Tittel.

He holds the Gene K. Beare Distinguished Professorship in the Department of Biomedical Engineering, Washington University at St. Louis, St. Louis, and directs the Optical Imaging Laboratory. He invented or discovered frequency-swept ultrasound-modulated optical tomography, dark-field confocal photoacoustic microscopy, photoacoustic Doppler sensing, photoacoustic reporter gene imaging, focused scanning microwave-induced thermoacoustic tomography, exact reconstruction algorithms for photoacoustic or thermoacoustic tomography, sonoluminescence tomography, Mueller-matrix optical coherence tomography, and oblique-incidence reflectometry. His Monte Carlo model of photon transport in scattering media is used worldwide as a standard tool. He authored more than 137 peer-reviewed journal articles, and one of the first textbooks in biophotonics entitled *Biomedical Optics Principles and Imaging* (Wiley, 2007).

Prof. Wang is a Fellow of the American Institute for Medical and Biomedical Engineering (AIMBE), The Optical Society of America (OSA), and The International Society for Optical Engineering (SPIE). He serves on the editorial boards of the *Journal of Biomedical Optics* and *Applied Optics*. He is the Chair of the International Biomedical Optics Society. He was the recipient of the National Institutes of Health (NIH) FIRST award, the National Science Foundation (NSF) CAREER award, and the Outstanding Young Scientist Award sponsored by Johnson & Johnson Medical and the Houston Society for Engineering in Medicine and Biology.

Multifrequency Microwave-Induced Thermal Acoustic Imaging for Breast Cancer Detection

Bin Guo, *Member, IEEE*, Jian Li*, *Fellow, IEEE*, Henry Zmuda, *Member, IEEE*, and Mark Sheplak

Abstract—Microwave-induced thermal acoustic imaging (TAI) is a promising early breast cancer detection technique, which combines the advantages of microwave stimulation and ultrasound imaging and offers a high imaging contrast, as well as high spatial resolution at the same time. A new multifrequency microwave-induced thermal acoustic imaging scheme for early breast cancer detection is proposed in this paper. Significantly more information about the human breast can be gathered using multiple frequency microwave stimulation. A multifrequency adaptive and robust technique (MART) is presented for image formation. Due to its data-adaptive nature, MART can achieve better resolution and better interference rejection capability than its data-independent counterparts, such as the delay-and-sum method. The effectiveness of this procedure is shown by several numerical examples based on 2-D breast models. The finite-difference time-domain method is used to simulate the electromagnetic field distribution, the absorbed microwave energy density, and the thermal acoustic field in the breast model.

Index Terms—Breast cancer detection, finite-difference time-domain (FDTD) methods, multifrequency adaptive and robust technology (MART), robust capon beamforming (RCB), thermal acoustic imaging (TAI).

I. INTRODUCTION

BREAST cancer is the most common nonskin malignancy in women and the second leading cause of female cancer mortality [1]. There are over 200 000 new cases of invasive breast cancer diagnosed each year in the U. S., and one out of every seven women in the U.S. will be diagnosed with breast cancer in their life time (the American Cancer Society, 2006) and early diagnosis is key to surviving breast cancer [2]. Microwave imaging is a method for early breast cancer detection [3]–[9], which exploits the significant contrast in dielectric properties between normal and cancerous tissues [10]–[12]. However, it is difficult for microwave imaging techniques to achieve good (submillimeter) spatial resolution because of

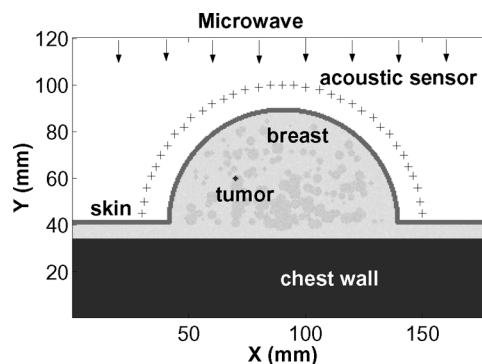


Fig. 1. Model of microwave-induced TAI for breast cancer detection.

its long wavelength [13]. Ultrasound is another option which offers a high spatial resolution because of its short acoustic wavelength [14]–[16]. However, the contrast in acoustic properties between normal and tumor tissues is very small due to both being soft tissues.

Microwave-induced thermal acoustic imaging (TAI) combines the advantages of microwave stimulation and ultrasound imaging [13], which offers a high imaging contrast (due to the significantly different dielectric properties of tumor and normal breast tissues), as well as high spatial resolution (due to the low propagation velocity or the short wavelength of acoustic waves in biological tissues) at the same time. To use microwave-induced TAI techniques for breast cancer imaging, a microwave source with a short duration time is used to irradiate the breast, as shown in Fig. 1. The normal breast tissues, as well as tumors, absorb microwave energy and emanate thermal acoustic waves by thermoelastic expansion. It is well-known that malignant breast tissue has a higher water content [1], [10], [12], [17], with a much higher conductivity than normal breast tissues (with low water content). As a result, the microwave energy absorbed by tumor and normal breast tissues will be significantly different and a stronger acoustic wave will be produced by the tumor. The acoustic waves generated in this manner carry the information about the microwave energy absorption properties of the tissues under irradiation. The thermal acoustic waves propagate out of the breast and are recorded by an acoustic sensor array placed around the breast. The tumor locations can be accurately determined since the received acoustic signals from the malignant tumors are at higher levels, hence aiding image construction.

During the last decade, several research groups have been working on the microwave-induced TAI of biological tissues [18]–[23]. The microwave frequency used ranges from 400 MHz [22] to 3 GHz [13]. Image reconstruction algorithms include the

Manuscript received August 22, 2006; revised February 7, 2007. This work was supported in part by the National Institutes of Health under Grant 1R41CA107903-1, in part by the U.S. Army Medical Command under Contract W81XWH-06-1-0389, and in part by the National Natural Science Foundation of China under Grant 60428101. Asterisk indicates corresponding author.

B. Guo and H. Zmuda are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: guobin@dsp.ufl.edu).

*J. Li is with the Department of Electrical and Computer Engineering, P.O. Box 116130, University of Florida, Gainesville, FL 32611 USA (e-mail: li@dsp.ufl.edu).

M. Sheplak is with the Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611 USA.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TBME.2007.895108

widely used delay-and-sum (DAS) method [20], [23], the frequency-domain inverse method [24], [25], and the time-domain inverse method [13], [19].

Microwave-induced TAI does, however, present several challenges. First, the human breast is large in size, usually has an irregular shape if not compressed, and is covered with a 2-mm-thick skin with dielectric properties significantly different from the normal breast tissues. Moreover the breast tissue is far from homogeneous because the dielectric properties of glandular tissue are different from that of breast fatty tissue. All these factors make it difficult to approximate the back propagation properties of thermal acoustic signals inside the breast. Due to the slow acoustic wave propagation speed or short wavelength in biological tissues, the errors on the order of millimeters in determining the acoustic signal propagation path lengths will severely degrade the image quality.

In this paper, a multifrequency microwave-induced TAI system is proposed which remedy the problems mentioned above. Instead of using a single frequency microwave source, as generally done by other research groups in this field, here a multiple frequency source is used, since the desired thermal acoustic signals can be induced by microwave sources operating at a wide range of frequencies. We show in this paper that the rich information collected from the multifrequency stimulation can help mitigate the challenges mentioned. The multifrequency microwave-induced thermal acoustic signals will offer higher signal-to-noise ratio (SNR) and higher imaging contrast than single-frequency microwave-induced thermal acoustic signals since much more information about the human breast can be harvested from the multiple stimulating frequencies within the microwave frequency band. Furthermore, the interference due to inhomogeneous breast tissue can be suppressed more effectively when multifrequency microwave-induced thermal acoustic signals are used for image reconstruction since more information about breast tissue can be used by the adaptive image reconstruction algorithms.

Another challenge encountered by microwave-induced TAI is the need to develop accurate and robust image reconstruction methods. DAS is a widely used reconstruction algorithm in medical imaging. This method is data-independent and tends to suffer from poor resolution and high sidelobe level problems. Data-adaptive approaches, such as the recently introduced robust Capon beamforming (RCB) [26], [27] method, can have much better resolution and much better interference rejection capabilities than their data-independent counterpart. Several medical imaging algorithms [4], [5], [28], [29] based on RCB have been developed and used for microwave imaging and thermal acoustic imaging. Good performances of these algorithms have been reported.

We present a multifrequency adaptive and robust technique (MART) based on RCB for multifrequency microwave-induced TAI. There are three stages in our MART. In Stage I, RCB is used to estimate the thermal acoustic responses from the grid locations within the breast for each stimulating microwave frequency. Then, in Stage II, a scalar acoustic waveform at each grid location is estimated based on the response estimates for all stimulating frequencies from Stage I. Finally, in Stage III, the positive peak and the negative peak of the estimated

acoustic waveform at each grid location are determined, and the peak-to-peak difference is computed and referred to as the image intensity.

To validate the effectiveness of the proposed algorithm, we develop a 2-D inhomogeneous breast model, which includes skin, breast fatty tissues, glandular tissues, and the chest wall. Small tumors are set below the skin. The finite-difference time-domain (FDTD) method is used to simulate the electromagnetic field inside the breast tissues [30], [31]. The specific absorption rate (SAR) distribution is calculated based on the simulated electromagnetic field [32], [33]. Then FDTD is used again to simulate the propagation of the microwave-induced thermal acoustic waves [34], [35].

The remainder of this paper is organized as follows. In Section II, the microwave frequency properties of human breast are described. A proper microwave frequency band for multifrequency microwave-induced TAI is also given in this section. MART is proposed for image formation in Section III. In Section IV, 2-D electromagnetic and acoustic breast models are developed. The electromagnetic and acoustic simulation methods are also presented in this section. Imaging results based on numerical examples are provided in Section V. Section VI concludes the paper.

II. MICROWAVE PROPERTIES OF HUMAN BREAST

A. Cutoff Frequency of Human Breast

In a microwave-induced TAI system, the biological tissues should be heated by microwave sources in a uniform manner, otherwise thermal acoustic signals will be induced by a nonuniform microwave energy distribution, resulting in images difficult to interpret. It is well-known that high-order electromagnetic field modes will be excited in a media if the microwave works at a frequency higher than a cutoff frequency of the media [36], and the microwave energy distribution is nonuniform at high-order modes [37]. To minimize the nonuniform microwave energy distribution inside the breast caused by the high-order electromagnetic modes, the microwave source should work at a frequency below a certain cutoff frequency.

To estimate the cutoff frequency for the human breast, we consider the simplified breast model shown in Fig. 2(a) consisting of a semicircular dielectric waveguide with a perfect magnetic conductor (PMC) at the bottom of the semicircle. Recall that the tangential components of the magnetic field are zero on the surface at the PMC. The PMC assumption is reasonable because the permittivity of the chest wall ($\epsilon_r = 50$) is much greater than that of the normal breast tissues ($\epsilon_r = 9$). In circular dielectric waveguide, if an electromagnetic mode has a field distribution whose tangential magnetic field components are zero at the center line of the circular waveguide, as shown in Fig. 2(b), the introduction of a PMC at the center line of the circular waveguide will not significantly change the boundary conditions and, hence, will not significantly alter the mode distribution. The modes in the semicircular dielectric waveguide can, thus, be estimated by determining the modes in a corresponding circular dielectric waveguide.

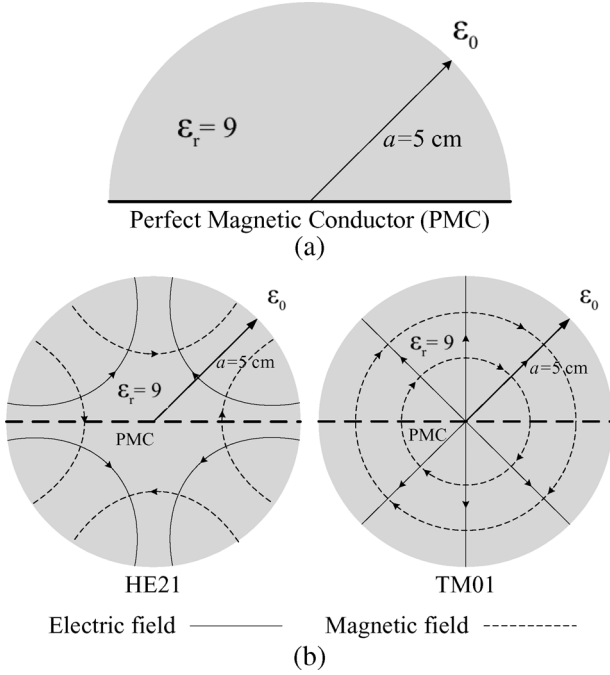


Fig. 2. Simplified breast model. (a) Semicircular dielectric waveguide with PEC, and (b) corresponding circular dielectric waveguide.

The dominant mode of a circular dielectric waveguide is the HE11 mode, the cutoff frequency of which is zero. The electromagnetic field distribution is near uniform at this mode. The dominant mode is followed by the TE01, TM01, and HE21 modes. These modes are degenerate, and have a cutoff frequency given by [36]

$$f_c = \frac{\chi_{01} C_0}{2\pi a \sqrt{\epsilon_r - 1}} \quad (1)$$

where C_0 is the speed of light in free space, $\chi_{01} = 2.405$ is the first root of the Bessel function of the first kind of order zero ($J_0(\chi_{01}) = 0$), a and ϵ_r are the radius and average permittivity of the circular dielectric waveguide, respectively. TM01 and HE21, as well as the interference between them, satisfy the zero tangential magnetic field component condition at the center line of the circular waveguide. These modes can also exist in the semicircular dielectric waveguide. By substituting the parameters of the breast model into (1), we obtain the cutoff frequency of the semicircular breast model to be

$$f_c = \frac{2.405 \cdot 3 \times 10^8}{2\pi \cdot 0.05 \cdot \sqrt{9 - 1}} = 812 \text{ MHz} \quad (2)$$

where we have used $a = 5 \text{ cm}$ and $\epsilon_r = 9$ as typical values for human breast. Consequently, the stimulating microwave frequency for the TAI system should be below 812 MHz.

B. Microwave Energy Absorption Properties of Breast Tissues and Tumor

It is well-known that the complex relative dielectric properties of a medium can be expressed as

$$\epsilon_r = \epsilon'_r - j\epsilon''_r \quad (3)$$

TABLE I
COLE-COLE PARAMETERS FOR BIOLOGICAL TISSUES

Tissue	Breast	Skin	Muscle	Tumor
ϵ_∞	2.5	4.0	4.0	4.0
σ	0.01	0.0002	0.2	0.7
$\Delta\epsilon_1$	3.0	32.0	50.0	50.0
τ_1 (ps)	17.68	7.23	7.23	7.0
α_1	0.1	0.0	0.1	0.0
$\Delta\epsilon_2$	15	1100	7000	0
τ_2 (ns)	63.66	32.48	353.68	N/A
α_2	0.1	0.2	0.1	N/A
$\Delta\epsilon_3$	5.0E4	0	1.2E6	0
τ_3 (μ s)	454.7	N/A	318.31	N/A
α_3	0.1	N/A	0.1	N/A
$\Delta\epsilon_4$	2.0E7	N/A	2.5E7	0
τ_4 (ms)	13.26	N/A	2.274	N/A
α_4	0.0	N/A	0.0	N/A

where ϵ'_r is the relative permittivity and ϵ''_r is the out-of-phase loss factor which can be written as

$$\epsilon''_r = \frac{\sigma}{\epsilon_0 \omega} \quad (4)$$

with σ being the total conductivity, ϵ_0 the free space permittivity, and ω the electromagnetic frequency. The tissue absorption property of the electromagnetic wave energy is

$$Q(\mathbf{r}) = \frac{1}{2} \sigma |E(\mathbf{r})|^2 \quad (5)$$

which is a function of the total conductivity and the electric field inside the tissue. If we assume that the microwave energy distribution is uniform inside the breast in a TAI system, the absorption of the microwave energy by the breast is characterized by the total conductivity of the breast tissues

$$\sigma(\omega) = \epsilon''_r \epsilon_0 \omega. \quad (6)$$

Hence, instead of using the attenuation coefficient α , as used in [23], in this paper, we study the absorption properties of breast tissues using the total conductivity σ .

The dielectric properties of biological tissues can be accurately modeled by the Cole-Cole equation [38]

$$\epsilon_r(\omega) = \epsilon_\infty + \sum_{i=1}^K \frac{\Delta\epsilon_i}{1 + (j\omega\tau_i)^{1-\alpha_i}} + \frac{\sigma_0}{j\omega\epsilon_0} \quad (7)$$

where K is the order of the Cole-Cole model, ϵ_∞ is the high-frequency permittivity, τ_i is the relaxation time, $\Delta\epsilon_i$ is the pole amplitude, α_i ($0 \leq \alpha_i \leq 1$) is a measure of the broadening of dispersion, and σ_0 is the static ionic conductivity. The Cole-Cole parameters for skin, breast fatty tissue, chest wall (mainly consisting of muscle), as well as tumor are listed in Table I [39], [40]. Because we cannot find the values specific to the tumor, the dielectric properties of the tumor is approximated using a Debye model [3], [41], which is a special case of the Cole-Cole model.

Substituting (7) into (6), we obtain the total conductivity of the breast tissue as follows:

$$\sigma(\omega) = -\text{imag} \left(\epsilon_\infty + \sum_{i=1}^K \frac{\Delta\epsilon_i}{1 + (j\omega\tau_i)^{1-\alpha_i}} + \frac{\sigma}{j\omega\epsilon_0} \right) \epsilon_0 \omega \quad (8)$$

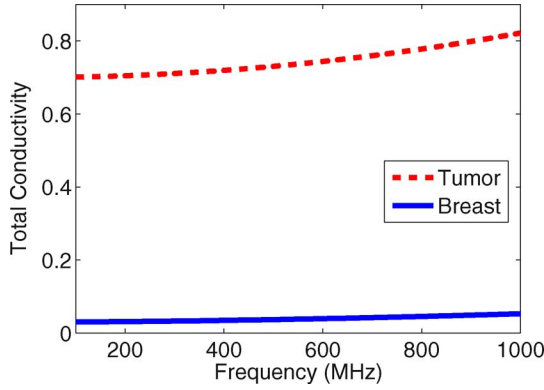


Fig. 3. Total conductivity of normal breast tissues and tumor as a function of frequency.

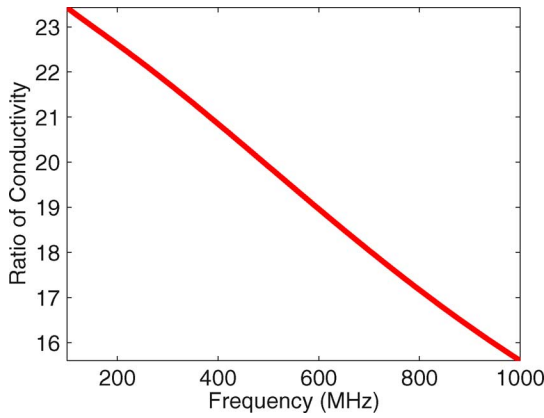


Fig. 4. Ratio of conductivity between tumor and normal breast tissue as a function of frequency.

which is a function of the stimulating microwave frequency, where $\text{imag}(\cdot)$ denotes the imaginary part of the complex relative permittivity. Fig. 3 gives the total conductivity of breast fatty tissue and tumor over a frequency band from 100–1000 MHz. Note that the total conductivity increases with the microwave frequency, which means that more microwave energy is absorbed and converted to heat by tissues at higher microwave frequency region, or in other words, the SNR is higher in the received thermal acoustic signals at higher stimulating microwave frequency region. On the other hand, the penetration at higher microwave frequencies is smaller because the tissues are lossy. We define the conductivity ratio between the tumor and the normal breast tissue as

$$r_{\sigma}(\omega) = \frac{\sigma_{\text{tumor}}(\omega)}{\sigma_{\text{breast}}(\omega)} \quad (9)$$

and plot it as a function of frequency in Fig. 4. A high conductivity ratio means that more microwave energy is absorbed and converted to heat by tumor than by normal breast tissues. In other words, the higher the conductivity ratio, the higher the imaging contrast. Fig. 4 shows that the imaging contrast is higher at the lower microwave frequency region because the conductivity ratio decreases with the frequency.

These microwave energy absorption properties of breast tissues and tumor motivate us to consider inducing thermal acoustic signals with different microwave frequencies. By

taking into account the aforementioned cutoff frequency given in (2), we choose a frequency range from 200–800 MHz. The frequency step is 100 MHz, with a total of seven frequencies. Note that wideband antenna techniques should be used for the practical implementation because the frequency range is wide. However, since the exciting microwave frequency is stepped, an antenna with a broad instantaneous bandwidth is not required. Another advantage of using multiple frequencies for stimulation is that more information about the inhomogeneous breast tissues will be harvested from the multifrequency microwave-induced thermal acoustic signals. The microwave energy distribution inside the breast model is not uniform because the human breast is inhomogeneous media, and thermal acoustic signals will be induced by the inhomogeneous energy distribution. These thermal acoustic signals will appear as clutter in the resulting images. However, the inhomogeneous microwave energy distributions are different at different stimulating frequencies because of the different microwave wavelengths in breast tissues. When a multifrequency microwave source is used for TAI, the thermal acoustic clutter induced by the inhomogeneous breast tissues can be suppressed by our adaptive and robust imaging algorithm.

III. MULTIFREQUENCY ADAPTIVE AND ROBUST TECHNOLOGY (MART) FOR BREAST CANCER IMAGING

We consider a multifrequency microwave-induced TAI system as shown in Fig. 1, where an acoustic sensor array is arranged on a semicircle relatively close to the breast skin. The location of each acoustic sensor is $\mathbf{r}_j$ ($j = 1, \dots, N$), where N is the number of acoustic sensors. Assume that $M = 7$ microwave sources with different frequencies are used to irradiate the breast model. Let $p_{i,j}(t)$ ($i = 1, \dots, M$; $j = 1, \dots, N$; $t = 0, \dots, T - 1$) denote the thermal acoustic signal induced by the i^{th} frequency and received by the j^{th} acoustic sensor, where T is the recording time which is sufficiently long to allow acoustic sensors to record all responses from the breast. Our goal is to detect the tumor by reconstructing an image of the thermal acoustic response intensity $I(\mathbf{r})$ as a function of scan location $\mathbf{r}$ within the breast.

A. Data Preprocessing

Because the breast skin, breast tissues, chest wall, and tumor absorb the microwave energy and convert the energy to heat, all of them produce thermal acoustic signals. The received thermal acoustic waveforms include the responses from the tumor, as well as from other healthy breast tissues. The thermal acoustic signals generated by the skin are much stronger than those by a small tumor because of the high conductivity of the skin and the acoustic sensors being very close to the skin. We must remove the skin responses to enhance the tumor responses. Because the distances between the acoustic sensors and the nearest breast skin are similar to one another, the signals recorded by various sensors have similar skin responses. Hence, we can remove the skin response by subtracting out a fixed calibration signal from all received signals. This calibration signal can be obtained simply by averaging the recorded signals from all channels.

Let $x_{i,j}(t)$ denote the signals after subtracting out the calibration signal. To process the signals coherently for a focal point

at $\mathbf{r}$, we align the signals $x_{i,j}(t)$ by time shifting each signal a number of samples $n_j(\mathbf{r})$. The discrete time delay between $\mathbf{r}$ and the j^{th} acoustic sensor can be calculated as

$$n_j(\mathbf{r}) = \left\lfloor \frac{\|\mathbf{r}_j - \mathbf{r}\|}{\Delta t c} \right\rfloor \quad (10)$$

where $\lfloor \gamma \rfloor$ stands for rounding to the greatest integer less than γ , $\|\mathbf{r}_j - \mathbf{r}\|$ is the distance between $\mathbf{r}_j$ and $\mathbf{r}$, c is the velocity of the acoustic wave propagating in breast tissues, and Δt is the sampling interval, which is assumed to be sufficiently small. The time-shifted signals are denoted as

$$\tilde{x}_{i,j}(t, \mathbf{r}) = x_{i,j}(t + n_j(\mathbf{r})), \quad t = -n_j(\mathbf{r}), \dots, T - n_j(\mathbf{r}). \quad (11)$$

After time shifting, the acoustic signals from the imaging location $\mathbf{r}$ are aligned so that they all start approximately from time $t = 0$ for all channels. Now the aligned signals are windowed by

$$\text{Window}(l) = \begin{cases} 1, & 0 \leq l \leq L - 1 \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

to isolate the signals from the focal point at $\mathbf{r}$. The windowed signals are denoted as $\tilde{x}_{i,j}(l, \mathbf{r})$, $l = 0, \dots, L - 1$, where $L\Delta t$ is the approximate duration of the thermal acoustic pulse, which can be determined from the pulse duration of the pulsed microwave source.

Attenuation exists when acoustic waves propagate within the breast. This attenuation has two parts: the attenuation due to the lossy media and the propagation attenuation. Thus, the attenuation of the tumor responses at various channels are different because of the different distances between the imaging position $\mathbf{r}$ and the acoustic sensors. For the 2-D case considered here, the compensation factor for the j^{th} channel is given by

$$K_j(\mathbf{r}) = \exp(\alpha \|\mathbf{r}_j - \mathbf{r}\|) \cdot \|\mathbf{r}_j - \mathbf{r}\|^{1/2} \quad (13)$$

where the first term of the right side of (13) compensates for the attenuation due to the lossy media, and the second term compensates for the geometric attenuation. The compensated signal can be calculated as

$$y_{i,j}(l, \mathbf{r}) = K_j(\mathbf{r}) \cdot \tilde{x}_{i,j}(l, \mathbf{r}), \quad l = 0, \dots, L - 1. \quad (14)$$

B. Multifrequency Adaptive and Robust Technology (MART)

Without loss of generality, we consider imaging at a generic location $\mathbf{r}$ only. For notational convenience, we drop the dependence of $y_{i,j}(l, \mathbf{r})$ on $\mathbf{r}$, and simply denote it as $y_{i,j}(l)$. Now we consider the data model

$$y_{i,j}(l) = s_{i,j}(l) + e_{i,j}(l) \quad (15)$$

where $s_{i,j}(l)$ represents the tumor response and $e_{i,j}(l)$ represents the residual term, which includes the noise and interference from breast skin, chest wall, and other responses. The structure of the data model is a data cube as shown in Fig. 5.

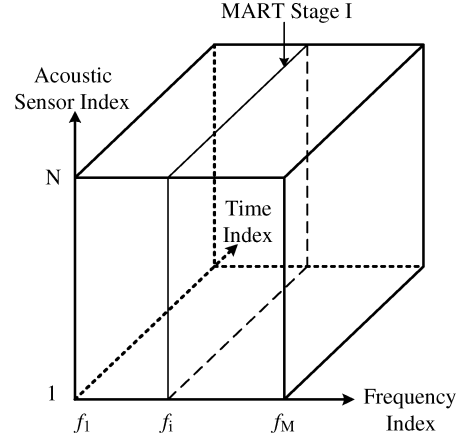


Fig. 5. Data cube model. In Stage I, MART slices the data cube for each frequency index. RCB is applied to each data slice to estimate the corresponding waveform.

MART is a three-stage time-domain signal processing algorithm. In Stage I, MART slices the data cube corresponding to each frequency index, and processes each data slice by the RCB to obtain the thermal acoustic waveform estimate for each stimulating frequency. Then, in Stage II, a scalar waveform is estimated from all frequencies based on the waveform estimates from Stage I. Finally, the positive peak and the negative peak of the estimated thermal acoustic waveform from Stage II are found in Stage III. The peak-to-peak difference is calculated as the image intensity at the focal point at $\mathbf{r}$. The details of all three stages are given below.

1) *Stage I:* In Stage I, MART approximates the data model as

$$\mathbf{y}_i(l) = \mathbf{a}_i s_i(l) + \mathbf{e}_i(l) \quad (16)$$

where $\mathbf{y}_i(l) = [y_{i,1}(l), \dots, y_{i,N}(l)]^T$ and $\mathbf{e}_i(l) = [e_{i,1}(l), \dots, e_{i,N}(l)]^T$. The scalar waveform $s_i(l)$ denotes the thermal acoustic signal generated at the focal location $\mathbf{r}$ corresponding to the i^{th} stimulating frequency. The vector $\mathbf{a}_i$ is referred to as the array steering vector, which is approximately equal to $\mathbf{1}_{N \times 1} = [1, \dots, 1]^T$ since all the signals have been aligned temporally and their attenuation compensated for in the preprocessing step. The residual $\mathbf{e}_i(l)$ is the noise and interference term, which is assumed uncorrelated with the signal.

There are two assumptions made to write the model given in (16). First, the steering vector is assumed to vary with the microwave frequency (i) but nearly constant with the time sample (l). Second, we assume that the thermal acoustic signal waveform depends only on the microwave frequency (i) but not on the acoustic sensor (j). The truth, however, is that the steering vector is not exactly known as it changes slightly with both the stimulating frequency and time due to array calibration errors and other factors. The signal waveform can also vary slightly with both the stimulating frequency and acoustic sensor, due to the inhomogeneous and frequency-dependent medium within the breast. The two aforementioned assumptions simplify the problem slightly. They cause little performance degradations when used with our adaptive and robust algorithm.

In practice, the true steering vector in (16) is not $\mathbf{1}_{N \times 1}$. We assume that the true steering vector $\mathbf{a}_i$ lies in the vicinity of the assumed steering vector $\bar{\mathbf{a}} = \mathbf{1}_{N \times 1}$, and the only knowledge we have about $\mathbf{a}_i$ is that

$$\|\mathbf{a}_i - \bar{\mathbf{a}}\|^2 \leq \epsilon_1 \quad (17)$$

where ϵ_1 is a user parameter, which may be determined depending on the various errors discussed previously.

The true steering vector $\mathbf{a}_i$ can be estimated via the following covariance fitting approach of RCB [26], [27]

$$\begin{aligned} \max_{\delta_i^2, \mathbf{a}_i} \delta_i^2 \quad \text{subject to} \quad & \hat{\mathbf{R}}_{\mathbf{Y}_i} - \delta_i^2 \mathbf{a}_i \mathbf{a}_i^T \geq 0 \\ & \|\mathbf{a}_i - \bar{\mathbf{a}}\|^2 \leq \epsilon_1 \end{aligned} \quad (18)$$

where δ_i^2 is the power of the signal $s_i(l)$ and

$$\hat{\mathbf{R}}_{\mathbf{Y}_i} \triangleq \frac{1}{L} \sum_{l=0}^{L-1} \mathbf{y}_i(l) \mathbf{y}_i^T(l) \quad (19)$$

is the sample covariance matrix. The above optimization problem can be solved as described in [26], and the estimated true steering vector is denoted here as $\hat{\mathbf{a}}_i$.

To obtain the signal waveform estimate, we pass the received signals through a Capon beamformer [27], [42]. The weight vector of the beamformer is determined by using the estimated steering vector $\hat{\mathbf{a}}_i$ in the following expression:

$$\mathbf{w}_i = \frac{\hat{\mathbf{R}}_{\mathbf{Y}_i}^{-1} \hat{\mathbf{a}}_i}{\hat{\mathbf{a}}_i^T \hat{\mathbf{R}}_{\mathbf{Y}_i}^{-1} \hat{\mathbf{a}}_i}. \quad (20)$$

Then the estimated signal waveform corresponding to the i^{th} stimulating frequency is

$$\hat{s}_i(l) = \mathbf{w}_i^T \mathbf{y}_i(l). \quad (21)$$

By repeating the aforementioned process for $i = 1$ through $i = M$, we obtain the complete set of M waveform estimates

$$\hat{\mathbf{s}}(l) = [\hat{s}_1(l), \dots, \hat{s}_M(l)]^T. \quad (22)$$

2) Stage II: Since the stimulating microwave sources with various frequencies are assumed to have the same power, we assume that the thermal acoustic responses from the tumor at different stimulating frequencies have nearly identical waveforms. Note that the thermal acoustic responses induced by the inhomogeneous microwave energy distribution (due to the inhomogeneous breast tissues) are different at different stimulating frequencies. This means that the elements of the vector $\hat{\mathbf{s}}(l)$ are all approximately equal to an unknown scalar signal $s(l)$, and the noise and interference term can be assumed uncorrelated with this signal. In Stage II of MART, we adopt the data model

$$\hat{\mathbf{s}}(l) = \mathbf{a}_s s(l) + \mathbf{e}_s(l) \quad (23)$$

where $\mathbf{a}_s$ is approximately equal to $\mathbf{1}_{M \times 1}$. However, the “steering vector” may again be imprecise, and, hence, RCB is needed again.

As we did in Stage I, we assume that the only knowledge about $\mathbf{a}_s$ is that

$$\|\mathbf{a}_s - \bar{\mathbf{a}}_s\|^2 \leq \epsilon_2, \quad (24)$$

where $\bar{\mathbf{a}}_s = \mathbf{1}_{M \times 1}$ is the assumed steering vector, and ϵ_2 is a user parameter. Again, the true steering vector $\mathbf{a}_s$ can be estimated via the covariance fitting approach

$$\begin{aligned} \max_{\delta^2, \mathbf{a}_s} \delta^2 \quad \text{subject to} \quad & \hat{\mathbf{R}}_s - \delta^2 \mathbf{a}_s \mathbf{a}_s^T \geq 0 \\ & \|\mathbf{a}_s - \bar{\mathbf{a}}_s\|^2 \leq \epsilon_2 \end{aligned} \quad (25)$$

where δ^2 is the power of the signal $s(l)$, and

$$\hat{\mathbf{R}}_s \triangleq \frac{1}{L} \sum_{l=0}^{L-1} \hat{\mathbf{s}}(l) \hat{\mathbf{s}}^T(l) \quad (26)$$

is the sample covariance matrix.

After obtaining the estimated steering vector $\hat{\mathbf{a}}_s$, we obtain the adaptive weight vector and the estimated signal waveform, respectively, as

$$\mathbf{w} = \frac{\hat{\mathbf{R}}_s^{-1} \hat{\mathbf{a}}_s}{\hat{\mathbf{a}}_s^T \hat{\mathbf{R}}_s^{-1} \hat{\mathbf{a}}_s} \quad (27)$$

and

$$\hat{s}(t) = \mathbf{w}^T \hat{\mathbf{s}}(t). \quad (28)$$

3) Stage III: Because the thermal acoustic pulse is usually bipolar: a positive peak, corresponding to the compression pulse, and a negative peak, corresponding to the rarefaction pulse [43], we use the peak-to-peak difference as the response intensity for the imaging location $\mathbf{r}$ in the third stage of MART. Compared with other energy or amplitude based response intensity estimation methods, peak-to-peak difference can be used to improve imaging quality with little additional computation costs.

The positive and negative peak values of the estimated waveform for the focal location $\mathbf{r}$ will be searched based on the estimated waveform (28) obtained in Stage II. Because of the nonuniform sound speed in biological tissues, the arrival time of the acoustic pulse generated at location $\mathbf{r}$ cannot be calculated accurately. However, it was reported in [18] that, when the heterogeneity is weak, such as in breast tissues, amplitude distortion caused by multipath is not severe. We assume that the original peak remains a peak in the estimated waveform, and the positive and negative peak values of the thermal acoustic pulse can be searched as

$$P^+ = \max \left\{ \max_{l \in [\Delta_1, \Delta_2]} \{\hat{s}(l)\}, 0 \right\} \quad (29)$$

and

$$P^- = \min \left\{ \min_{l \in [\Delta_1, \Delta_2]} \{\hat{s}(l)\}, 0 \right\} \quad (30)$$

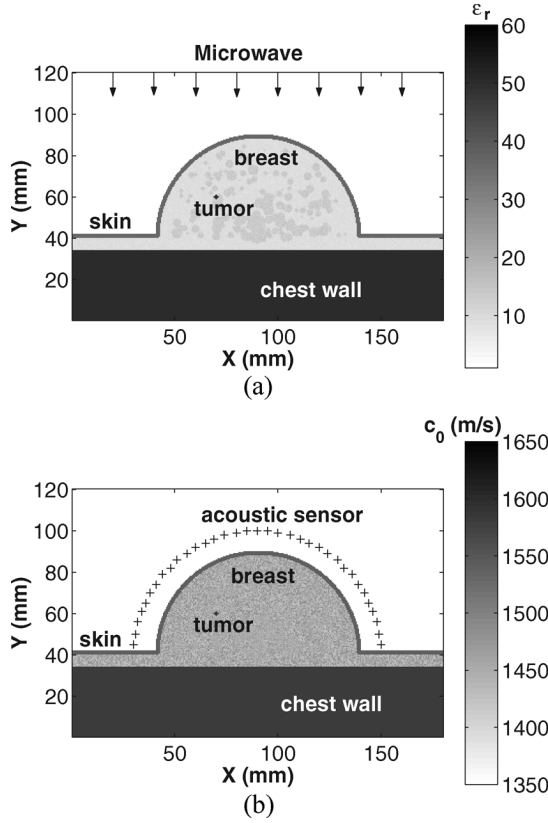


Fig. 6. Breast model for thermal acoustic simulation. (a) Model for electromagnetic simulation and (b) model for acoustic simulation.

where $[\Delta_1, \Delta_2] \in [0, L]$ is the searching range. Here Δ_1 and Δ_2 are user parameters, and the details on how to choose them can be found in [29].

After the positive and negative peak values are found, the response intensity for the focal point at location $\mathbf{r}$ is given as

$$I(\mathbf{r}) = P^+ - P^-. \quad (31)$$

IV. MODELING AND SIMULATIONS

We consider 2-D breast models simulated in two steps. In the first step, the electromagnetic field inside the breast model is simulated and the SAR distribution is calculated based on the simulated electromagnetic field. The second step is for the acoustic wave simulation, where the SAR distribution obtained in the first step is used as the acoustic pressure source through the thermal expansion coefficient. In both steps, the FDTD method is used for the simulation examples.

A. Electromagnetic Model and Simulation

For simulation purposes, the 2-D electromagnetic breast model used is as shown in Fig. 6(a). The breast model is a 10 cm in diameter semicircle, which includes the skin, breast fatty tissue, glandular tissues, and chest wall (muscle). A 1-mm-diameter tumor is embedded below the skin. The dielectric properties of the breast tissues as well as tumor at the

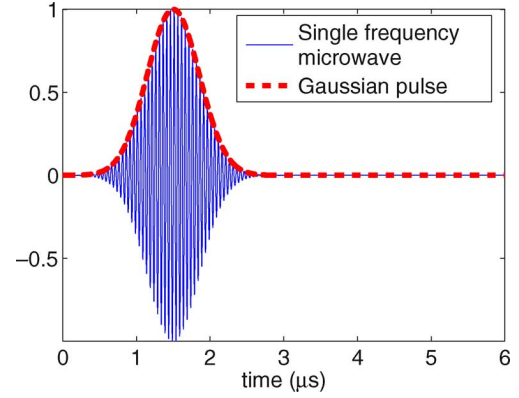


Fig. 7. Gaussian modulated microwave source.

microwave frequency f_i ($i = 1, \dots, M$) were calculated based on the Cole–Cole model in (7). The dielectric properties of the normal breast fatty tissue are assumed random with a variation of $\pm 10\%$ around the nominal values. The dielectric constants of glandular tissues are between $\epsilon_r = 11$ and $\epsilon_r = 15$.

Fig. 7 shows a Gaussian modulated electromagnetic wave used to irradiate the breast from the top of the model, as shown in Fig. 6(a). The time duration for the Gaussian pulse is $1 \mu\text{s}$. The electromagnetic field is simulated using the FDTD method [30], [31]. The grid cell size used by FDTD is $0.5 \times 0.5 \text{ mm}$ and the computational region is terminated by perfectly matched layer (PML) absorbing boundary conditions [44], [45].

The SAR distribution is given as [32], [33]

$$\text{SAR}(\mathbf{r}) = \frac{\sigma(\mathbf{r})E^2(\mathbf{r})}{2\rho(\mathbf{r})} \quad (32)$$

where $\sigma(\mathbf{r})$ is the conductivity of the biological tissues at location $\mathbf{r}$, $E(\mathbf{r})$ is the electric field at location $\mathbf{r}$, and $\rho(\mathbf{r})$ is the mass density of the biological tissues at location $\mathbf{r}$.

B. Acoustic Model and Simulation

In the microwave-induced TAI system, the microwave energy is small, and as a result, the acoustic pressure field induced by microwave is also small. So, the nonlinear acoustic effect does not need to be considered in the TAI system. For example, it is shown in [46] that the temperature rise is about 0.1 mK and the acoustic pressure change is only about 100 Pa in the microwave-induced TAI system. The shock distance in breast tissues is [47]

$$\chi_s = \frac{\frac{B}{2A}}{2\left(\frac{B}{2A} + 1\right)} \cdot \frac{\rho_0 c^2}{p} \cdot \lambda_{\min} = \frac{\frac{B}{2A}}{2\left(\frac{B}{2A} + 1\right)} \cdot \frac{\rho_0 c^2}{p} \cdot \frac{c}{f_{\max}} \quad (33)$$

where $(B/A) (\approx 10)$ is the nonlinear factor of the breast tissues, $\rho_0 (\approx 1000 \text{ kg/m}^3)$ is the mass density of the breast tissues, and $c (\approx 1500 \text{ m/s})$ is the sound speed inside the breast tissues [14]. p is the acoustic pressure rise, and $\lambda_{\min}$ and $f_{\max}$ are the minimal acoustic wavelength and the maximal acoustic frequency of the thermal acoustic signal, respectively. For our breast model, the acoustic pressure rise is $p = 100 \text{ Pa}$, and the

TABLE II
ACOUSTIC PARAMETERS FOR BIOLOGICAL TISSUES.
(* f IS THE ACOUSTIC FREQUENCY AND THE UNIT IS MEGAHERTZ)

Tissue	Breast	Skin	Muscle	Tumor
ρ (kg/m^3)	1020	1100	1041	1041
c (m/s)	1510	1537	1580	1580
α^* (dB/cm)	$0.75f^{1.5}$	3.5	$0.57f$	$0.57f$
β ($1/^\circ C$)	3E-4	3E-4	3E-4	3E-4
C_p ($J/(^\circ C \cdot kg)$)	3550	3500	3510	3510

maximal acoustic frequency is $f_{\max} = 500$ KHz. By substituting the parameters into (33), we obtain the shock distance in breast tissues to be

$$\begin{aligned} \chi_s &= \frac{\frac{B}{2A}}{2 \left(\frac{B}{2A} + 1 \right)} \cdot \frac{\rho_0 c^2}{p} \cdot \frac{c}{f_{\max}} \\ &= \frac{5}{2(5+1)} \cdot \frac{1000 \cdot 1500^2}{100} \cdot \frac{1500}{500 \times 10^3} \\ &= 2.8 \times 10^4 \text{ m.} \end{aligned} \quad (34)$$

Because the size of our breast model is only 10 cm, which is much smaller than the shock distance, it is reasonable to ignore the nonlinear acoustic effect in the microwave-induced TAI system.

The two basic linear acoustic wave generation equations are [13]

$$\rho \frac{\partial}{\partial t} \mathbf{u}(\mathbf{r}, t) = -\nabla p(\mathbf{r}, t) \quad (35)$$

and

$$\nabla \cdot \mathbf{u}(\mathbf{r}, t) = -\frac{1}{\rho c^2} \frac{\partial}{\partial t} p(\mathbf{r}, t) + \alpha p(\mathbf{r}, t) + \beta \frac{\partial}{\partial t} T(\mathbf{r}, t) \quad (36)$$

where $\mathbf{u}(\mathbf{r}, t)$ is the acoustic velocity vector, $p(\mathbf{r}, t)$ is the acoustic pressure field, ρ is the mass density, α is the attenuation coefficient, β is the thermal expansion coefficient, and $T(\mathbf{r}, t)$ is the temperature. The values for these acoustic properties for different breast tissues are listed in Table II [14], [46], [48]–[50]. The attenuation coefficient is calculated with $f = 0.15$ MHz. The values for the tumor are approximated using the values for muscle because we cannot find the values specific to the tumor.

Because the duration of the microwave pulse is much shorter than the thermal diffusion time, thermal diffusion can be neglected [13], and the thermal equation is

$$C_p \frac{\partial}{\partial t} T(\mathbf{r}, t) = \text{SAR}(\mathbf{r}, t) \quad (37)$$

where C_p is the specific heat. Substituting (37) into (36) gives

$$\nabla \cdot \mathbf{u}(\mathbf{r}, t) = -\frac{1}{\rho c^2} \frac{\partial}{\partial t} p(\mathbf{r}, t) + \alpha p(\mathbf{r}, t) + \frac{\beta}{C_p} \text{SAR}(\mathbf{r}, t). \quad (38)$$

FDTD is used again to compute the thermal acoustic wave based on (35) and (38). More details about FDTD for acoustic simulations can be found in [34], [35], and [51]–[57].

The breast model for the acoustic simulation is constructed similarly to the model for electromagnetic simulation. The

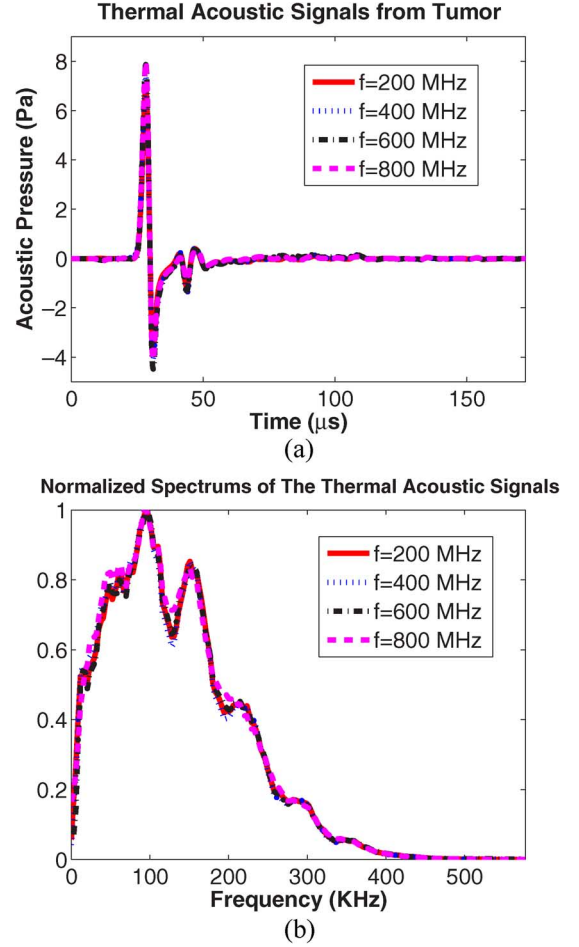


Fig. 8. Thermal acoustic signals at different stimulating frequencies $f = 200, 400, 600$, and 800 MHz. (a) Thermal acoustic responses from tumor only and (b) the normalized spectrums of the signals in (a).

velocities of the normal fatty breast tissue are also assumed random with a variation of $\pm 5\%$ around average values, as shown in Fig. 6(b). An acoustic sensor array with 35 elements deployed uniformly around the breast model is used to record the thermal acoustic signals. The distance between neighboring acoustic sensors is 4 mm. The grid cell size used by the acoustic FDTD is 0.1×0.1 mm and the computational region is terminated by PML absorbing boundary conditions [55]–[57]. Note that the size of the FDTD cell for acoustic simulation is much finer than that of the FDTD cell for electromagnetic simulation because the wavelength of an acoustic wave is much smaller than that of a microwave. The SAR distribution data is interpolated to achieve the designed grid resolution for the acoustic breast model.

V. NUMERICAL EXAMPLES

At the beginning of this section, the typical microwave-induced thermal acoustic responses from the tumor are plotted in Fig. 8(a) for stimulating frequencies of $f = 200, 400, 600$, and 800 MHz. The signals are simulated based on the aforementioned 2-D breast model. To obtain the signals, we perform the simulation twice at each stimulating frequency, with and

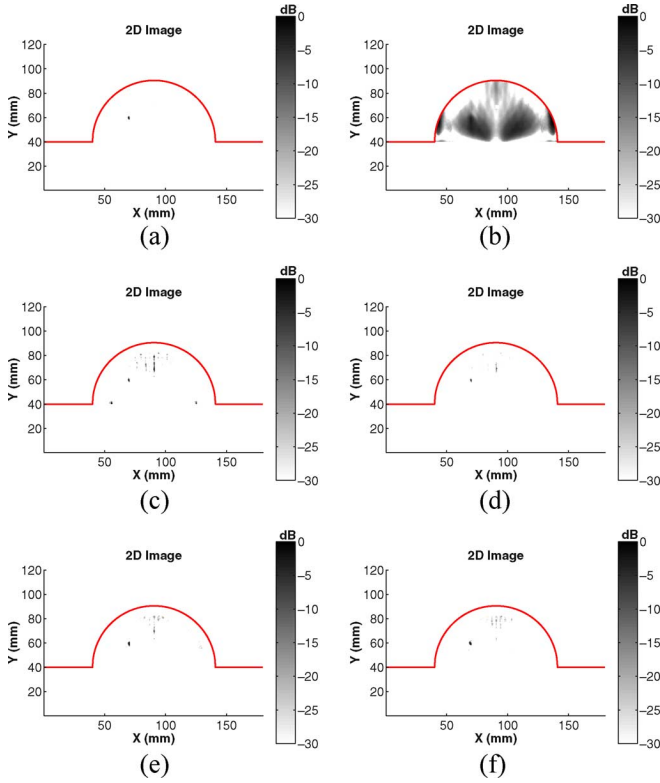


Fig. 9. Imaging results for the case of a single 1 mm-diameter tumor. (a) MART, (b) DAS, (c) SART at stimulating frequency $f = 200$ MHz, (d) SART at stimulating frequency $f = 400$ MHz, (e) SART at stimulating frequency $f = 600$ MHz, and (f) SART at stimulating frequency $f = 800$ MHz.

without the tumor, and record the acoustic signals in an acoustic sensor. The difference of the two received signals is referred to as the thermal acoustic response only from the tumor at the stimulating frequency. It can be seen that the thermal acoustic responses from the tumor at different stimulating frequencies are similar to one another. The figure also shows that the thermal acoustic signals are wideband bipolar pulses, with a large positive peak and a large negative peak. Fig. 8(b) shows the normalized spectra of the acoustic signals corresponding to the excitations in Fig. 8(a). It is seen that the frequency range of the acoustic signals is about from 1 to 400 KHz. The dominant band (3-dB band) of the signals ranges from 10 to 180 KHz, and the corresponding acoustic wavelength ranges from 150 to 8 mm in the breast tissues.

Several numerical examples are used in this section to demonstrate the effectiveness of MART. The thermal acoustic signals are simulated based on the 2-D breast model with tumor only. For comparison purposes, the DAS method is applied to the same data set also. We also present the imaging results for the single-frequency microwave-induced TAI at different stimulating frequencies. The corresponding image reconstruction method is referred to as the single-frequency adaptive and robust technique (SART), which is similar with MART but without Stage II of MART. In SART, the RCB is used to estimate the thermal acoustic waveform at a certain stimulating frequency just like in Stage I of MART. Then the peak search method used in MART Stage III is applied to the estimated waveform to determine the image intensity.

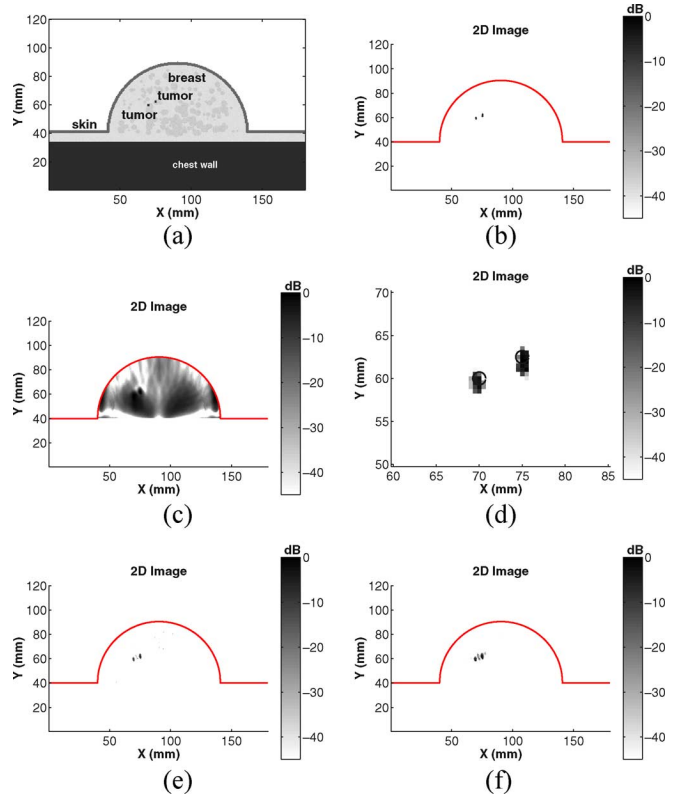


Fig. 10. Imaging results for the two 1.5-mm-diameter tumors case. (a) Breast model, (b) MART, (c) DAS, (d) zoom in of (b), (e) SART at stimulating frequency $f = 300$ MHz, and (f) SART at stimulating frequency $f = 700$ MHz.

In the first example, a 1-mm-diameter tumor is embedded in the breast model at the location $(x = 70 \text{ mm}, y = 60 \text{ mm})$. This is the challenging case of early breast cancer detection because of the small tumor size. Fig. 9(a) and (b) shows the imaging results for MART and DAS, respectively. The tumor is shown clearly in the MART image [Fig. 9(a)], and the size and location of the tumor is accurate. Because of the high side-lobe, poor resolution, and poor interference rejection capability of the DAS method, the tumor is essentially missed by DAS as shown in Fig. 9(b). Fig. 9(c)–(f) shows the imaging results for SART at the stimulating frequencies $f = 200, 400, 600$, and 800 MHz, respectively. The figures show that SART can determine the tumor correctly, but some clutter shows up in the SART images. Note that the clutter shows up at different locations with different stimulating frequencies. By comparing the images for MART and SART, it can be seen that the clutter are effectively suppressed by MART when multiple stimulating frequencies are used.

In the second numerical example, two small 1.5-mm-diameter tumors are set inside the breast model as shown in Fig. 10(a). Their locations are at $(x = 70 \text{ mm}, y = 60 \text{ mm})$ and $(x = 75 \text{ mm}, y = 62.5 \text{ mm})$. The distance between the two tumors is 4 mm. The imaging results using MART and DAS are shown in Fig. 10(b) and (c), respectively. The two tumors are seen clearly in the MART image. To show them clearly we zoom in onto the tumor locations in Fig. 10(d), where the two black circles mark the actual sizes and locations of the two tumors. It is shown that MART can be used to determine the locations and sizes of the two tumors accurately. The DAS

image contains much clutter. The two tumors cannot be separated clearly in the DAS image because of the poor resolution of DAS. Fig. 10(e) and (f) shows the imaging results of SART at stimulation frequencies $f = 300$ and 700 MHz, respectively. The tumors can be seen in both of the SART images, but clutter shows up between the two tumors in Fig. 10(e) and (f), and the sizes of the two tumors in Fig. 10(f) are larger than their true sizes.

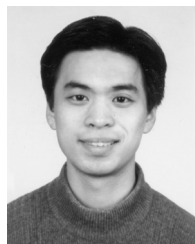
VI. CONCLUSION

An investigation of using a multifrequency microwave-induced TAI system for early breast cancer detection has been reported in this paper. The frequency band for this system has been given based on the cutoff frequency of the human breast. A simplified semicircular dielectric waveguide mode was used to calculate the cutoff frequency in this paper. By studying the microwave energy absorption properties of breast tissue and tumor, we have shown that the multifrequency microwave-induced TAI can offer higher SNR, higher imaging contrast, and more effective clutter suppression capability than the traditional single-frequency microwave-induced TAI. A MART has been presented for image formation. This data-adaptive algorithm can achieve better resolution and better interference rejection capability than its data-independent counterparts, such as DAS. The feasibility of this multifrequency microwave-induced TAI system as well as the performance of the proposed image reconstruction algorithm for early breast cancer detection have been demonstrated by using 2-D numerical electromagnetic and acoustic breast models. The absorbed microwave energy and the thermal acoustic field in the breast models have been simulated using the FDTD method. Numerical examples have been used to demonstrate the excellent performance of this multifrequency technique. More advanced models are being developed to further investigate and validate the preferential imaging capability of the technique for early breast cancer detection.

REFERENCES

- [1] E. C. Fear, P. M. Meaney, and M. A. Stuchly, "Microwaves for breast cancer detection," *IEEE Potentials*, vol. 22, no. 1, pp. 12–18, Feb./Mar. 2003.
- [2] P. M. Meaney, M. W. Fanning, D. Li, S. P. Poplack, and K. D. Paulsen, "A clinical prototype for active microwave imaging of the breast," *IEEE Trans. Microw. Theory Tech.*, vol. 48, no. 11, pp. 1841–1853, Nov. 2000.
- [3] E. J. Bond, X. Li, S. C. Hagness, and B. D. Van Veen, "Microwave imaging via space-time beamforming for early detection of breast cancer," *IEEE Trans. Antennas Propagat.*, vol. 51, no. 8, pp. 1690–1705, Aug. 2003.
- [4] B. Guo, Y. Wang, J. Li, P. Stoica, and R. Wu, "Microwave imaging via adaptive beamforming methods for breast cancer detection," *J. Electromagn. Waves Appl.*, vol. 20, no. 1, pp. 53–63, Jan. 2006.
- [5] Y. Xie, B. Guo, L. Xu, J. Li, and P. Stoica, "Multi-static adaptive microwave imaging for early breast cancer detection," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 8, pp. 1647–1657, Aug. 2006.
- [6] D. Li, P. M. Meaney, and K. D. Paulsen, "Conformal microwave imaging for breast cancer detection," *IEEE Trans. Microw. Theory Tech.*, vol. 51, no. 4, pp. 1179–1186, Apr. 2003.
- [7] Q. F. Fang, P. M. Meaney, S. D. Geimer, A. V. Streltsov, and K. D. Paulsen, "Microwave image reconstruction from 3-D fields coupled to 2-D parameter estimation," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 475–484, Apr. 2004.
- [8] Q. H. Liu, Z. Q. Zhang, T. T. Wang, J. A. Bryan, G. A. Ybarra, L. W. Nolte, and W. T. Joines, "Active microwave imaging I—2-D forward and inverse scattering methods," *IEEE Trans. Microw. Theory Tech.*, vol. 50, no. 1, pp. 123–133, Jan. 2002.
- [9] Z. Q. Zhang, Q. H. Liu, C. Xiao, E. Ward, G. Ybarra, and W. T. Joines, "Microwave breast imaging: 3-D forward scattering simulation," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 10, pp. 1180–1189, Oct. 2003.
- [10] R. K. Mishra, A. Swarup, and J. M. Thomas, "Dielectric properties of normal and malignant human breast tissues at radiowave and microwave frequencies," *Ind. J. Biochem. Biophys.*, vol. 21, pp. 76–79, Feb. 1984.
- [11] A. J. Surowiec, S. S. Stuchly, J. R. Barr, and A. Swarup, "Dielectric properties of breast carcinoma and the surrounding tissues," *IEEE Trans. Biomed. Eng.*, vol. 35, no. 4, pp. 257–263, Apr. 1988.
- [12] W. T. Joines, Y. Zhang, C. X. Li, and R. L. Jirtel, "The measured electrical properties of normal and malignant human tissues from 50 to 900 MHz," *Med. Phys.*, vol. 21, no. 4, pp. 547–550, Apr. 1994.
- [13] M. Xu and L. V. Wang, "Time-domain reconstruction for thermoacoustic tomography in a spherical geometry," *IEEE Trans. Med. Imag.*, vol. 21, no. 7, pp. 814–822, Jul. 2002.
- [14] T. Szabo, *Diagnostic Ultrasound Imaging: Inside Out*. Burlington, MA: Elsevier, 2004.
- [15] F. W. Kremkau, *Diagnostic Ultrasound: Principles and Instruments*. Philadelphia, PA: Saunders, 1993.
- [16] E. Steen and B. Olstad, "Volume rendering of 3-D medical ultrasound data using direct feature mapping," *IEEE Trans. Med. Imag.*, vol. 13, no. 6, pp. 517–525, Jun. 1994.
- [17] W. Joines, R. Jirtel, M. Rafal, and D. Schaeffer, "Microwave power absorption differences between normal and malignant tissue," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 6, no. 6, pp. 681–687, 1980.
- [18] Y. Xu and L. H. Wang, "Effects of acoustic heterogeneity in breast thermoacoustic tomography," *IEEE Trans. Ultrason. Ferroelect. Freq. Control*, vol. 50, no. 9, pp. 1134–1146, Sep. 2003.
- [19] M. Xu, Y. Xu, and L. V. Wang, "Time-domain reconstruction algorithms and numerical simulations for thermoacoustic tomography in various geometries," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 9, pp. 1086–1099, Sep. 2003.
- [20] Y. Xu, P. Kuchment, and L. H. Wang, "Limited-view thermoacoustic tomography and reconstruction by truncated conjugate gradient," *Med. Phys.*, vol. 31, pp. 724–733, Apr. 2004.
- [21] R. A. Kruger, K. K. Kopecky, A. M. Aisen, D. R. Reinecke, G. A. Kruger, and W. L. Kiser, "Thermoacoustic CT with radio waves: A medical imaging paradigm," *Radiology*, vol. 211, pp. 275–278, Apr. 1999.
- [22] R. A. Kruger, K. D. Miller, H. E. Reynolds, W. L. Kiser, D. R. Reinecke, and G. A. Kruger, "Breast cancer in vivo: Contrast enhancement with thermoacoustic CT at 434 MHz—Feasibility study," *Radiology*, vol. 216, pp. 279–283, Jul. 2000.
- [23] R. A. Kruger, W. L. Kiser, K. D. Miller, H. E. Reynolds, D. R. Reinecke, and G. A. Kruger, "Thermoacoustic CT," in *Proc. IEEE MTT-S Int. Microwave Symp. Dig.*, Jun. 2000, vol. 2, pp. 11–16.
- [24] Y. Xu, D. Feng, and L. H. Wang, "Exact frequency-domain reconstruction for thermoacoustic tomography I: Planar geometry," *IEEE Trans. Med. Imag.*, vol. 21, no. 7, pp. 823–828, Jul. 2002.
- [25] Y. Xu, M. Xu, and L. H. Wang, "Exact frequency-domain reconstruction for thermoacoustic tomography II: Cylindrical geometry," *IEEE Trans. Med. Imag.*, vol. 21, no. 7, pp. 829–833, Jul. 2002.
- [26] J. Li, P. Stoica, and Z. Wang, "On robust capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.
- [27] J. Li and P. Stoica, Eds., *Robust Adaptive Beamforming*. New York: Wiley, 2005.
- [28] Y. Xie, B. Guo, J. Li, and P. Stoica, "Novel multistatic adaptive microwave imaging methods for early breast cancer detection," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 1–12, 2006.
- [29] Y. Xie, B. Guo, J. Li, G. Ku, and L. V. Wang, "Adaptive and robust techniques (ART) for thermoacoustic tomography," presented at the Proc. 40th Asilomar Conf. Signals, Systems and Computers, Pacific Grove, CA, Nov. 2006.
- [30] A. Taflov and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 2nd ed. Boston, MA: Artech House, 2000.
- [31] D. M. Sullivan, *Electromagnetic Simulation Using the FDTD Method*. New York: IEEE, 2000.
- [32] P. Bernardi, M. Cavagnaro, S. Pisa, and E. Piuze, "Specific absorption rate and temperature elevation in a subject exposed in the far-field of radio-frequency sources operating in the 10–900-MHz range," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 295–304, Mar. 2003.
- [33] P. Bernardi, M. Cavagnaro, S. Pisa, and E. Piuze, "SAR distribution and temperature increase in an anatomical model of the human eye exposed to the field radiated by the user antenna in a wireless LAN," *IEEE Trans. Microw. Theory Tech.*, vol. 46, no. 12, pp. 2074–2082, Dec. 1998.

- [34] J. G. Maloney and K. E. Cummings, "Adaptation of FDTD techniques to acoustic modeling," in *Proc. 11th Annu. Rev. Progress in Applied Computational Electromagnetics*, Monterey, CA, 1995, vol. 2, pp. 724–731.
- [35] X. Yuan, D. Borup, J. Wiskin, M. Berggren, and S. A. Johnson, "Simulation of acoustic wave propagation in dispersive media with relaxation losses by using FDTD method with PML absorbing boundary condition," *IEEE Trans. Ultrason., Ferroelect., Freq. Control*, vol. 46, no. 1, pp. 14–23, Jan. 1999.
- [36] C. A. Balanis, *Advanced Engineering Electromagnetics*. New York: Wiley, 1989.
- [37] S. Yang, H. Kim, and H. J. Lee, "Circular-harmonic vector analysis of dielectric waveguide with a cross-cut-circle cross section," *Appl. Opt.*, vol. 34, pp. 7705–7713, Nov. 1995.
- [38] S. Gabriel, R. W. Lau, and C. Gabriel, "The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues," *Phys. Med. Biol.*, vol. 41, pp. 2271–2293, Nov. 1996.
- [39] C. Gabriel and S. Gabriel, "Compilation of the dielectric properties of body tissues at rf and microwave frequencies," [Online]. Available: <http://www.brooks.af.mil/AFRL/HED/hedr/reports/dielectric/home.html>
- [40] "Dielectric properties of body tissues," [Online]. Available: <http://nremf.ifac.cnr.it/tissprop/>
- [41] X. Li and S. C. Hagness, "A confocal microwave imaging algorithm for breast cancer detection," *IEEE Microw. Wireless Compon. Lett.*, vol. 11, no. 3, pp. 130–132, Mar. 2001.
- [42] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ: Prentice-Hall, 2005.
- [43] C. G. A. Hoelen and F. F. M. de Mul, "Image reconstruction for photoacoustic scanning of tissue structures," *Appl. Opt.*, vol. 39, pp. 5872–5883, Nov. 2000.
- [44] S. D. Gedney, "An anisotropic perfectly matched layer-absorbing medium for the truncation of FDTD lattices," *IEEE Trans. Antennas Propagat.*, vol. 44, no. 12, pp. 1630–1639, Dec. 1996.
- [45] Z. S. Sacks, D. M. Kingsland, R. Lee, and J. F. Lee, "A perfectly matched anisotropic absorber for use as an absorbing boundary condition," *IEEE Trans. Antennas Propagat.*, vol. 43, no. 12, pp. 1460–1463, Dec. 1995.
- [46] M. Xu and L. H. Wang, "Pulsed-microwave-induced thermoacoustic tomography: Filtered backprojection in a circular measurement configuration," *Med. Phys.*, vol. 29, no. 8, pp. 1661–1669, Aug. 2002.
- [47] M. F. Hamilton and D. T. Blackstock, *Nonlinear Acoustics*. San Diego, CA: Academic, 1997.
- [48] F. A. Duck, *Physical Properties of Tissue*. London, U.K.: Academic, 1990.
- [49] M. L. Palmeri and K. R. Nightingale, "On the thermal effects associated with radiation force imaging of soft tissue," *IEEE Trans. Ultrason., Ferroelect., Freq. Control*, vol. 51, no. 5, pp. 551–565, May 2004.
- [50] M. Converse, E. J. Bond, S. C. Hagness, and B. D. Van Veen, "Ultra-wide-band microwave space-time beamforming for hyperthermia treatment of breast cancer: A computational feasibility study," *IEEE Trans. Microw. Theory Tech.*, vol. 52, no. 8, pp. 1876–1889, Aug. 2004.
- [51] C. K. W. Tam and J. C. Webb, "Dispersion-relation-preserving finite-difference schemes for computational acoustics," *J. Comput. Phys.*, vol. 107, no. 8, pp. 262–281, Aug. 1993.
- [52] S. V. Tsynkov, "Numerical solution of problems on unbounded domains. A review," *Appl. Numer. Math.*, vol. 27, pp. 465–532, Aug. 1998.
- [53] T. Colonius, S. K. Lele, and P. Moin, "Boundary-conditions for direct computation of aerodynamic sound generation," *AIAA J.*, vol. 31, pp. 1574–1582, Sep. 1993.
- [54] C. W. Rowley and T. Colonius, "Discretely nonreflecting boundary conditions for linear hyperbolic systems," *J. Comput. Phys.*, vol. 157, pp. 500–538, Jan. 2000.
- [55] Q. H. Liu and J. Tao, "The perfectly matched layer for acoustic waves in absorptive media," *J. Acoust. Soc. Amer.*, vol. 102, pp. 2072–2082, October 1997.
- [56] X. Yuan, D. Borup, J. W. Wiskin, M. Berggren, R. Eidens, and S. A. Johnson, "Formulation and validation of Berenger's PML absorbing boundary for the FDTD simulation of acoustic scattering," *IEEE Trans. Ultrason., Ferroelect., Freq. Control*, vol. 44, no. 7, pp. 816–822, Jul. 1997.
- [57] T. K. Katsibas and C. S. Antonopoulos, "A general form of perfectly matched layers for three-dimensional problems of acoustic scattering in lossless and lossy fluid media," *IEEE Trans. Ultrason., Ferroelect., Freq. Control*, vol. 51, no. 8, pp. 964–972, Aug. 2004.



Bin Guo (S'06–M'07) received the B.E. and M.Sc. degrees in electrical engineering from Xian Jiaotong University, Xian, China, in 1997 and 2000, respectively. He is currently pursuing the Ph.D. degree in electrical engineering in the Department of Electrical and Computer Engineering, University of Florida, Gainesville.

From April 2002 to July 2003, he was an Associate Research Scientist with the Temasek Laboratories, National University of Singapore, Singapore. Since August 2003, he has been a Research Assistant with the Department of Electrical and Computer Engineering, University of Florida. His current research interests include biomedical applications of signal processing, microwave imaging, and computational electromagnetics.



Jian Li (S'88–M'90–SM'97–F'05) received the M.Sc. and Ph.D. degrees in electrical engineering from The Ohio State University, Columbus, in 1987 and 1991, respectively.

From July 1991 to June 1993, she was an Assistant Professor with the Department of Electrical Engineering, University of Kentucky, Lexington. Since August 1993, she has been with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, where she is currently a Professor. Her current research interests include spectral estimation, statistical and array signal processing, and their applications.

Dr. Li is a fellow of Institution of Electrical Engineers (IEE). She received the 1994 National Science Foundation Young Investigator Award and the 1996 Office of Naval Research Young Investigator Award. She has been a member of the Editorial Board of Signal Processing, a publication of the European Association for Signal Processing (EURASIP) since 2005. She is presently a member of two of the IEEE Signal Processing Society technical committees: the Signal Processing Theory and Methods (SPTM) Technical Committee and the Sensor Array and Multichannel (SAM) Technical Committee.



Henry Zmuda (M'75) received the B.E. degree from the Stevens Institute of Technology, Hoboken, NJ, in 1979, and the M.S. and Ph.D. degrees from Cornell University, Ithaca, NY, in 1982 and 1984, respectively, all in electrical engineering.

From 1984 to 1995, he was a member of the faculty at Stevens Institute of Technology. In 1995, he joined the Electrical Engineering Faculty, University of Florida. His research interests are primarily in the area of microwave photonics and applied electromagnetics.



Mark Sheplak received the B.S. degree in 1989, the M.S. degree in 1992, and the Ph.D. degree in 1995 in mechanical engineering from Syracuse University, Syracuse, NY.

He is an Associate Professor in the Department of Aerospace and Mechanical Engineering and an Affiliate Associate Professor of electrical and computer engineering at the University of Florida (UF). Prior to joining UF in 1998, he was a Postdoctoral Associate at the Massachusetts Institute of Technology's Microsystems Technology Laboratories, Cambridge, from 1995 to 1998. During his Ph.D. studies, he was a GSRP Fellow at the NASA-Langley Research Center, Hampton, VA, from 1992 to 1995. His current research focuses on the design, fabrication, and characterization of high-performance, instrumentation-grade, MEMS-based sensors and actuators that enable the measurement, modeling, and control of various physical properties.

USER PARAMETER FREE APPROACHES TO MULTISTATIC ADAPTIVE ULTRASOUND IMAGING

Lin Du[†] Jian Li[†] Petre Stoica[‡]

[†]Dept. of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6130, USA.

[‡]Dept. of Information Technology, Uppsala University, SE-75105 Uppsala, Sweden.

ABSTRACT

Delay-and-sum (DAS) beamforming is the standard technique for ultrasound imaging applications. Due to its data independent property, DAS may suffer from poorer resolution and worse interference suppression capability than the adaptive standard Capon beamformer (SCB). However, the performance of SCB is sensitive to the errors in the sample covariance matrix and the signal steering vector. Therefore, robust adaptive beamforming techniques are desirable. In this paper, we consider ultrasound imaging via applying a user parameter free robust adaptive beamformer, which uses a shrinkage-based general linear combination (GLC) algorithm to obtain an enhanced estimate of the array covariance matrix. We present several multistatic adaptive ultrasound imaging (MAUI) approaches based on GLC to achieve high resolution and good interference suppression capability. The performance of the proposed MAUI approaches is demonstrated via an experimental example.

Index Terms—Adaptive beamforming, Ultrasound imaging

1. INTRODUCTION

Delay-and-sum (DAS) beamforming is the standard technique for ultrasound imaging applications. Theoretically this data independent approach has lower resolution and worse interference suppression capability than an adaptive beamformer, e.g., the standard Capon beamformer (SCB) [1]. However, in practice, there is a clear performance degradation for SCB when the covariance matrix is inaccurately estimated due to limited data samples and when the knowledge of the steering vector is imprecise due to look direction errors or imperfect array calibration. Therefore, adaptive beamforming approaches that are robust to the aforementioned problems are desired.

Most of the early approaches to robust adaptive beamforming are ad-hoc techniques, e.g., the traditional diagonal loading algorithm [2], for which there is no clear way to choose the diagonal loading level. The diagonal loading algorithm has been previously applied to ultrasound imaging [3], where the diagonal loading level was set to be proportional to the received power. The robust Capon beamformer (RCB) presented in [4], on the other hand, can precisely calculate the diagonal loading level based on the uncertainty set of the steering vector. RCB was applied to ultrasound imaging in [5] and the results showed that RCB can provide much better imaging quality than DAS. However, we still need to specify the uncertainty set parameter for RCB, which may be hard to do in practice. To achieve user parameter free robust adaptive

beamforming, we have recently devised several beamformers in [6] based on the shrinkage method, which can compute the diagonal loading level automatically without specifying any user parameters. Among these beamformers, the general linear combination (GLC) algorithm performs well, especially when the number of snapshots is small.

In this paper, we present several user parameter free approaches based on GLC for multistatic adaptive ultrasound imaging (MAUI), which form images of the backscattered energy for each focal point within the region of interest. All the MAUI approaches are two-stage imaging algorithms and GLC is employed in each stage. A similar idea can be applied to microwave imaging to replace the user parameter dependent RCB in each stage [7]. The complete multistatic data set for a given focal point can be represented by the data cube shown in Fig. 1. In one of the MAUI methods, which we refer to as MAUI-1, GLC is used in Stage I to obtain a set of backscattered signal estimates at each time instant. Based on these estimates, a scalar waveform is recovered via GLC in Stage II, which is then used to compute the backscattered energy. An alternative way of signal processing in Stage I is to compute a set of backscattered waveforms for each transmitter, which is referred to as MAUI-2. In addition, we also consider a combined method MAUI-C, which uses the signal estimates from both MAUI-1 and MAUI-2 in Stage I for the computation of backscattered energy. An experimental example will be presented to illustrate the performance of the MAUI methods.

Notation: The superscript $(\cdot)^*$ denotes the conjugate transpose, $(\cdot)^T$ denotes the transpose, $\lfloor x \rfloor$ denotes rounding to the greatest integer less than x , $E(\cdot)$ is the expectation operator, $\text{tr}(\cdot)$ is the trace operator, and $\|\cdot\|$ denotes the Frobenius norm for a matrix or the Euclidean norm for a vector. Finally $\mathbf{R} \geq 0$ means that $\mathbf{R}$ is positive semi-definite.

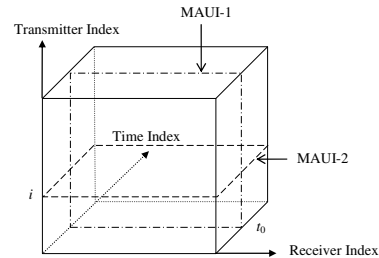


Fig. 1. The multistatic data cube. MAUI-1 processes the data set for a given time instant t_0 , while MAUI-2 processes the data set for a given transmitter index i .

II. PROBLEM FORMULATION

Consider an active array of M transducers using the multistatic (also called MIMO (multi-input multi-output) [8]) data

The work was supported in part by the U.S. Army Medical Command under Grant No. W81XWH-06-1-0389, by the National Science Foundation under Grant No. CCF-0634786 and ECCS-0729727 and by the Swedish Research Council (VR).

acquisition scheme. Each transducer takes turns to transmit the same pulse while all of the transducers record the backscattered signals. As a result, the received data set $\{\mathbf{P}_{i,j}(t), i, j = 1, \dots, M; t = 0, \dots, T-1\}$ comprises the A-scan data for all possible transmitter and receiver pairs of the array, where $\mathbf{P}_{i,j}(t)$ is the data sequence of the backscattered echo at the j^{th} transducer due to transmitting a pulse from the i^{th} transducer, and T is the number of data samples for the A-scan sequence.

To extend GLC to the wide-band ultrasound imaging application, we align the received signals from the data set $\{\mathbf{P}_{i,j}(t)\}$ to each focal point by inserting appropriate time delays. Let $\mathbf{r}_i$ and $\mathbf{r}_j$ denote the locations of the i^{th} transmitter and j^{th} receiver, respectively, and let $\mathbf{r}_f$ denote the location of a focal point in the imaging region of interest. The time delay due to the ultrasonic wave propagation from the i^{th} transmitter to the focal point $\mathbf{r}_f$ and then back to the j^{th} receiver is approximated as

$$\tau_{i,j}(\mathbf{r}_f) = \frac{1}{\Delta t} \left[\frac{\|\mathbf{r}_i - \mathbf{r}_f\|}{c} + \frac{\|\mathbf{r}_j - \mathbf{r}_f\|}{c} \right], \quad (1)$$

where c is the sound propagation speed in the medium under interrogation, and Δt denotes the sampling interval. Then, the time shifted signal for a given focal point of interest $\mathbf{r}_f$ can be represented as

$$y_{i,j}(\mathbf{r}_f, t) = \mathbf{P}_{i,j}(t + \tau_{i,j}(\mathbf{r}_f)), \quad i, j = 1, \dots, M; \quad t = 0, \dots, N-1, \quad (2)$$

where N is determined by the duration of the transmitted pulse and the sampling interval Δt .

The problem of interest here is to form an ultrasound image on a grid of points in the imaging region. The image is formed from the received data set $\{\mathbf{P}_{i,j}(t)\}$, or more precisely, $\{y_{i,j}(\mathbf{r}_f, t)\}$, for each focal point of interest.

III. MAUI

The two-stage MAUI algorithms use a GLC-based robust adaptive beamforming algorithm in each stage. We first review the GLC approach and then we show how to apply GLC to the data set $\{y_{i,j}(\mathbf{r}_f, t)\}$ in Stages I and II of the proposed MAUI approaches.

A. GLC

In the GLC approach, we replace the sample covariance matrix in SCB by an enhanced estimate obtained via a shrinkage-based method. The enhanced covariance matrix estimate $\tilde{\mathbf{R}}$ is obtained by linearly combining the sample covariance matrix $\hat{\mathbf{R}}$ and a shrinkage target (we use the identity matrix $\mathbf{I}$ here for lack of a better choice) in an optimal mean-squared error (MSE) sense:

$$\tilde{\mathbf{R}} = \alpha \mathbf{I} + \beta \hat{\mathbf{R}}, \quad (3)$$

where $\hat{\mathbf{R}} = \frac{1}{K} \sum_{k=1}^K \mathbf{y}(k) \mathbf{y}^*(k)$, with the $L \times 1$ vector $\mathbf{y}(k)$ denoting the k th snapshot and K representing the total number of snapshots. The shrinkage parameters α and β in (3) are estimated by minimizing the MSE of $\tilde{\mathbf{R}}$ with respect to α and

β , where

$$\begin{aligned} \text{MSE}(\tilde{\mathbf{R}}) &= E\{\|\tilde{\mathbf{R}} - \mathbf{R}\|^2\} \\ &= \|\alpha \mathbf{I} - (1 - \beta) \mathbf{R}\|^2 + \beta^2 E\{\|\hat{\mathbf{R}} - \mathbf{R}\|^2\} \\ &= \alpha^2 L - 2\alpha(1 - \beta) \text{tr}(\mathbf{R}) \\ &\quad + (1 - \beta)^2 \|\mathbf{R}\|^2 + \beta^2 E\{\|\hat{\mathbf{R}} - \mathbf{R}\|^2\}, \\ \mathbf{R} &= E[\mathbf{y}(k) \mathbf{y}^*(k)]. \end{aligned} \quad (4)$$

The optimal values for β and α can be readily obtained:

$$\beta_0 = \frac{\gamma}{\rho + \gamma}, \quad (5)$$

$$\alpha_0 = \nu(1 - \beta_0) = \nu \frac{\rho}{\gamma + \rho}, \quad (6)$$

where $\rho = E\{\|\hat{\mathbf{R}} - \mathbf{R}\|^2\}$, $\nu = \frac{\text{tr}(\mathbf{R})}{L}$, and $\gamma = \|\nu \mathbf{I} - \mathbf{R}\|^2$. Note that $\beta_0 \in [0, 1]$ and $\alpha_0 \geq 0$.

To estimate α_0 and β_0 from the given data, we need an estimate of ρ , which can be calculated as (see [9] for details):

$$\hat{\rho} = \frac{1}{K^2} \sum_{k=1}^K \|\mathbf{y}(k)\|^4 - \frac{1}{K} \|\hat{\mathbf{R}}\|^2. \quad (7)$$

Using (7) we can get estimates for β_0 and α_0 as

$$\hat{\beta}_0 = \frac{\hat{\gamma}}{\hat{\gamma} + \hat{\rho}}, \quad (8)$$

and

$$\hat{\alpha}_0 = \hat{\nu}(1 - \hat{\beta}_0), \quad (9)$$

where $\hat{\nu} = \frac{\text{tr}(\hat{\mathbf{R}})}{L}$, and $\hat{\gamma} = \|\hat{\nu} \mathbf{I} - \hat{\mathbf{R}}\|^2$. Note that $\hat{\alpha}_0 \geq 0$ and $\hat{\beta}_0 \geq 0$, which guarantees that the enhanced covariance matrix estimate $\tilde{\mathbf{R}} \geq 0$.

Substituting (8)-(9) in (3) yields the enhanced covariance matrix estimate $\tilde{\mathbf{R}}$. Using $\tilde{\mathbf{R}}$ instead of $\hat{\mathbf{R}}$ in the SCB formulation, we obtain the beamformer weight vector for GLC as follows:

$$\tilde{\mathbf{w}} = \frac{\tilde{\mathbf{R}}^{-1} \bar{\mathbf{a}}}{\bar{\mathbf{a}}^* \tilde{\mathbf{R}}^{-1} \bar{\mathbf{a}}}, \quad (10)$$

where $\bar{\mathbf{a}}$ denotes the assumed steering vector [10]. Note that GLC is a diagonal loading approach with the diagonal loading level $\hat{\alpha}_0/\hat{\beta}_0$ determined automatically from the observed data snapshots $\{\mathbf{y}(k)\}_{k=1}^K$.

B. Stage I

To apply the GLC-based robust adaptive beamformer to the data set $\{y_{i,j}(\mathbf{r}_f, t)\}$ in (2), we use two approximate signal models for $y_{i,j}(\mathbf{r}_f, t)$ by making different assumptions. Since we will concentrate on the focal point $\mathbf{r}_f$ in what follows, the dependence on $\mathbf{r}_f$ will be dropped for notational simplicity.

The MAUI-1 algorithm uses the following signal model:

$$\mathbf{y}_i(t) = \mathbf{a}(t) s_i(t) + \mathbf{e}_i(t), \quad (11)$$

where $\mathbf{y}_i(t) = [y_{i,1}(t), \dots, y_{i,M}(t)]^T$ represents the aligned array data vector of the i^{th} transmitter, $s_i(t)$ denotes the signal of interest (SOI) that is proportional to the ultrasound reflectivity or scattering strength, which is assumed to depend on the transmitter i but not on the receiver j , $\mathbf{e}_i(t)$ denotes the residual term due to noise and interferences, and $\mathbf{a}(t)$ denotes the array steering vector that is assumed to be approximately equal to $\mathbf{1}_{M \times 1}$. Here, we assume that $\mathbf{a}(t)$ may vary with t , but is constant with respect to the transmitter index i .

In Stage I, for a given time t_0 , we form a pseudo-covariance matrix by considering the number of transmitters as the number of snapshots:

$$\begin{aligned}\hat{\mathbf{R}}(t_0) &= \frac{1}{M} \mathbf{Y}(t_0) \mathbf{Y}^*(t_0), \\ \mathbf{Y}(t_0) &= [\mathbf{y}_1(t_0) \cdots \mathbf{y}_M(t_0)].\end{aligned}\quad (12)$$

By using $\hat{\mathbf{R}}(t_0)$ as the sample covariance matrix we obtain an enhanced covariance matrix estimate $\hat{\mathbf{R}}(t_0)$ as described in Section III.A, and then calculate the weight vector $\hat{\mathbf{w}}(t_0)$ for Stage I of MAUI-1 using (10) with $\bar{\mathbf{a}} = \mathbf{1}_{M \times 1}$ as:

$$\hat{\mathbf{w}}(t_0) = \frac{\hat{\mathbf{R}}(t_0)^{-1} \bar{\mathbf{a}}}{\bar{\mathbf{a}}^* \hat{\mathbf{R}}(t_0)^{-1} \bar{\mathbf{a}}}.\quad (13)$$

Once we got the weight vector, we can estimate $s_i(t_0)$ in (11) as:

$$\hat{s}_i(t_0) = \hat{\mathbf{w}}^*(t_0) \mathbf{y}_i(t_0).\quad (14)$$

Define a vector $\hat{\mathbf{s}}(t_0) = [\hat{s}_1(t_0), \dots, \hat{s}_M(t_0)]^T$ of the estimated signals for all transmitters. Repeating the above process from $t_0 = 0$ to $t_0 = N - 1$, we build the matrix $\hat{\mathbf{S}}_1 = [\hat{\mathbf{s}}(0), \dots, \hat{\mathbf{s}}(N - 1)]$.

The MAUI-2 algorithm considers another signal model:

$$\mathbf{y}_i(t) = \mathbf{a}_i s_i(t) + \mathbf{e}_i(t),\quad (15)$$

where $\mathbf{a}_i$ denotes the array steering vector, which is also assumed to be approximately equal to $\mathbf{1}_{M \times 1}$. However, different from MAUI-1, here $\mathbf{a}_i$ is assumed to change with i , but be constant with respect to t .

For a given transmitter i , the covariance matrix in Stage I of MAUI-2 is formulated as:

$$\begin{aligned}\hat{\mathbf{R}}_i &= \frac{1}{N} \mathbf{Y}_i \mathbf{Y}_i^*, \\ \mathbf{Y}_i &= [\mathbf{y}_i(0) \cdots \mathbf{y}_i(N - 1)].\end{aligned}\quad (16)$$

Using $\hat{\mathbf{R}}_i$ as the sample covariance matrix we get an enhanced estimate $\hat{\mathbf{R}}_i$, and then compute a weight vector $\hat{\mathbf{w}}_i$ using (10). The time sample vector of the corresponding beamformer output can be written as

$$\hat{\mathbf{s}}_i = [\hat{\mathbf{w}}_i^* \mathbf{Y}_i]^T.\quad (17)$$

Repeating the above process for $i = 1, \dots, M$ yields a set of waveforms $\hat{\mathbf{S}}_2 = [\hat{\mathbf{s}}_1, \dots, \hat{\mathbf{s}}_M]^T$.

As we mentioned before, the errors in the sample covariance matrix and the steering vector cause performance degradations for any adaptive beamforming algorithms. GLC is designed to improve the covariance matrix estimate. MAUI-1 and MAUI-2 use different sample covariance matrices. Hence the improvements obtained by using GLC may be different. This fact motivates us to combine MAUI-1 and MAUI-2 to achieve a better performance. We refer to this combined method, where $\hat{\mathbf{S}}_1$ of MAUI-1 and $\hat{\mathbf{S}}_2$ of MAUI-2 are used simultaneously, as MAUI-C. We denote the combined signal matrix as $\hat{\mathbf{S}}_C = [\hat{\mathbf{S}}_1^T \quad \hat{\mathbf{S}}_2^T]^T$.

Let the $M \times 1$ vectors $\{\hat{\mathbf{s}}_m(t)\}_{t=0, \dots, N-1}$ denote the columns of $\hat{\mathbf{S}}_m$ for $m = 1, 2$, and let the $2M \times 1$ vectors $\{\hat{\mathbf{s}}_C(t)\}_{t=0, \dots, N-1}$ denote the columns of $\hat{\mathbf{S}}_C$. Note that both MAUI-1 and MAUI-2 obtain M signal waveform estimates at the end of Stage I, while MAUI-C obtains $2M$ signal waveform estimates. We will apply GLC to these estimates in Stage II to recover a scalar waveform and compute the signal energy at the focal point.

C. Stage II

In Stage II, the signal model for both MAUI-1 and MAUI-2 can be represented as:

$$\hat{\mathbf{s}}_m(t) = \mathbf{a}_m s(t) + \mathbf{e}_m(t), \quad t = 0, \dots, N - 1, \quad m = 1, 2,\quad (18)$$

where the steering vector $\mathbf{a}_m$ is assumed to be $\mathbf{1}_{M \times 1}$, and $\mathbf{e}_m(t)$ represents the residual term. Similar to Stage I, the knowledge of $\mathbf{a}_m$ may be imprecise and the sample size N may be small. Hence the GLC-based robust adaptive beamformer is used again to estimate $s(t)$. Taking $\hat{\mathbf{R}}_m$ as the sample covariance matrix:

$$\hat{\mathbf{R}}_m = \frac{1}{N} \sum_{t=0}^{N-1} \hat{\mathbf{s}}_m(t) \hat{\mathbf{s}}_m^*(t), \quad m = 1, 2,\quad (19)$$

and paralleling the development in Stage I, we obtain the weight vector $\hat{\mathbf{w}}_m$ using (10). Then, the output signal estimate is computed as:

$$\hat{s}(t) = \hat{\mathbf{w}}_m^* \hat{\mathbf{s}}_m(t), \quad m = 1, 2.\quad (20)$$

Finally, the signal energy for a particular focal point $\mathbf{r}_f$ is computed as:

$$\mathcal{E}(\mathbf{r}_f) = \sum_{t=0}^{N-1} \hat{s}^2(t).\quad (21)$$

For Stage II of MAUI-C, the signal model can be written as:

$$\hat{\mathbf{s}}_C(t) = \mathbf{a}_C s(t) + \mathbf{e}_C(t), \quad t = 0, \dots, N - 1,\quad (22)$$

where the vector $\hat{\mathbf{s}}_C(t)$ is considered now to be a snapshot from a $2M$ -element ‘‘array’’, and the steering vector $\mathbf{a}_C$ is assumed to be $\mathbf{1}_{2M \times 1}$. $\mathbf{e}_C(t)$ denotes the residual term. We obtain the weight vector $\hat{\mathbf{w}}_C$ for MAUI-C via (10) by using the following sample covariance matrix:

$$\hat{\mathbf{R}}_C = \frac{1}{N} \sum_{t=0}^{N-1} \hat{\mathbf{s}}_C(t) \hat{\mathbf{s}}_C^*(t).\quad (23)$$

The beamformer $\hat{\mathbf{w}}_C$ yields an estimate of the signal:

$$\hat{s}(t) = \hat{\mathbf{w}}_C^* \hat{\mathbf{s}}_C(t).\quad (24)$$

Then, the backscattered energy at the focal point $\mathbf{r}_f$ is computed via (21).

IV. EXPERIMENTAL EXAMPLE

In this section, we present some experimental results to demonstrate the performance of the three MAUI algorithms. The complete multistatic data set was obtained by Bioacoustics Research Laboratory of the University of Illinois at Urbana-Champaign. The scene of interest contains several wire targets arranged in a complicated pattern. The data was collected using a 64-element linear array. The transducer center frequency was 2.6 MHz, the sampling rate was 25 MHz, and the sound velocity was assumed to be 1450 m/s. For comparison, the multistatic DAS scheme is also applied to the same data set. The DAS scheme estimates the signal waveform $s(t)$ as

$$\hat{s}(t) = \hat{\mathbf{w}}_{\text{DAS}}^* \mathbf{Y}(t) \hat{\mathbf{w}}_{\text{DAS}}, \quad t = 0, \dots, N - 1,\quad (25)$$

where $\hat{\mathbf{w}}_{\text{DAS}} = \bar{\mathbf{a}}/M$ is the weight vector for DAS. The backscattered energy at $\mathbf{r}_f$ is then estimated via (21).

Fig. 2 shows the ultrasound images for the wire data set under consideration. The images are displayed on a logarithmic scale with a 30 dB dynamic range. In Figs. 2 (a)-(d), we compare the images obtained via DAS and MAUI algorithms using only the central 32 elements of the array ($M = 32$). Since DAS simply sums all signals, the DAS image shown in Fig. 2 (a) has higher sidelobe level and poorer resolution than the MAUI images shown in Figs. 2 (b)-(d). Comparing Fig. 2 (b) and Fig. 2 (c), which correspond to MAUI-1 and MAUI-2 respectively, we note that MAUI-2 image has a lower background clutter level. However, MAUI-2 has poorer resolution: some wire targets are not discernable in the MAUI-2 image. On the other hand, the image obtained via MAUI-C has low sidelobe level similarly to MAUI-2 and high resolution similarly to MAUI-1. Moreover, all targets are clearly shown in the MAUI-C image. For comparison, we also include the DAS image obtained using the entire array ($M = 64$). Note that MAUI algorithms, especially MAUI-C, with 32 transducers can achieve similar imaging quality to DAS with a double sized array.

V. CONCLUSIONS

We have presented three user parameter free approaches to multistatic adaptive ultrasound imaging (MAUI). These two-stage MAUI approaches employ a GLC-based robust adaptive beamformer in each stage to achieve high resolution and good interference suppression capability, and also they are robust to small sample size problems and array steering vector errors. More importantly, GLC is a user parameter free approach as opposed to other existing robust adaptive beamforming algorithms, which makes it easy to use it in practice. The experimental results have demonstrated the effectiveness of the MAUI algorithms for ultrasound imaging. We have shown that the MAUI-C method, which combines MAUI-1 and MAUI-2, provides the best imaging quality.

REFERENCES

- [1] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, pp. 1408–1418, August 1969.
- [2] H. Cox, R. Zeskind, and M. Owen, "Robust adaptive beamforming," *Proceedings of IEEE*, vol. ASSP-35, pp. 1365–1375, October 1987.
- [3] J. Synnevåg, A. Austeng, and S. Holm, "Adaptive beamforming applied to medical ultrasound imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 54, no. 8, pp. 1606–1613, 2007.
- [4] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Transactions on Signal Processing*, vol. 51, pp. 1702–1715, July 2003.
- [5] Z. Wang, J. Li, and R. Wu, "Time-delay and time-reversal based robust Capon beamformers for ultrasound imaging," *IEEE Transactions on Medical Imaging*, vol. 24, no. 10, pp. 1308–1322, 2005.
- [6] J. Li, L. Du, and P. Stoica, "Fully automatic computation of diagonal loading levels for robust adaptive beamforming," submitted to *IEEE Transactions on Aerospace and Electronic Systems*, 2007.
- [7] Y. Xie, B. Guo, J. Li, and P. Stoica, "Novel multistatic adaptive microwave imaging methods for early breast cancer detection," *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp. 1–12, 2006.
- [8] J. Li and P. Stoica, "MIMO radar with colocated antennas: Review of some recent work," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 106–114, 2007.
- [9] P. Stoica, J. Li, X. Zhu, and J. R. Guerci, "On using a priori knowledge in space-time adaptive processing," to appear in *IEEE Transactions on Signal Processing*, 2007.
- [10] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ: Prentice-Hall, 2005.

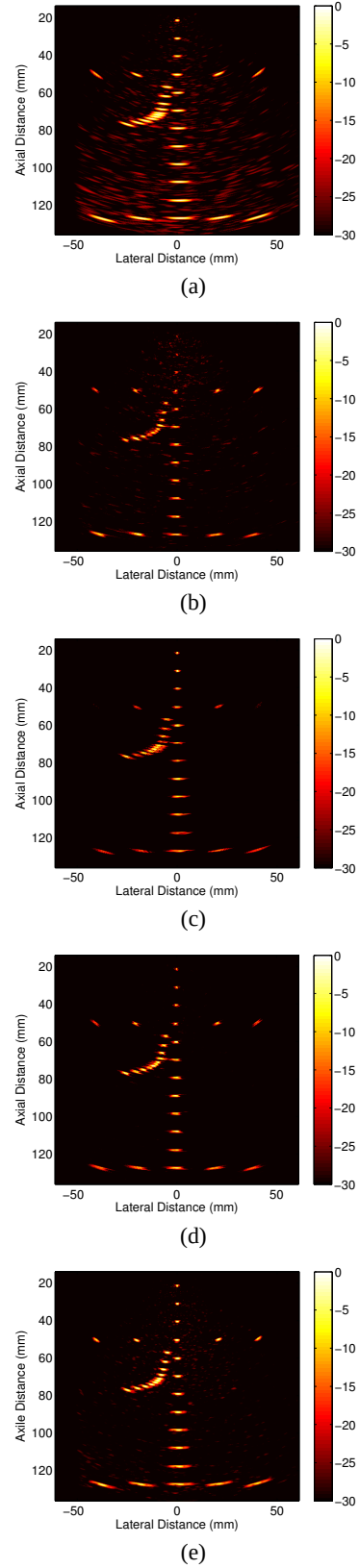


Fig. 2. Ultrasound images obtained via (a) DAS with $M = 32$, (b) MAUI-1 with $M = 32$, (c) MAUI-2 with $M = 32$, (d) MAUI-C with $M = 32$, and (e) DAS with $M = 64$.

Review of User Parameter-Free Robust Adaptive Beamforming Algorithms

Lin Du ^a, Tarik Yardibi ^a, Jian Li ^{a,*}, Petre Stoica ^b,

^a*Department of Electrical and Computer Engineering, University of Florida,
Gainesville, FL 32611-6130, USA*

^b*Department of Information Technology, Uppsala University, SE-751 05 Uppsala,
Sweden*

Abstract

This paper provides a comprehensive review of user parameter-free robust adaptive beamforming algorithms. We present ridge regression Capon beamformers (RRCB), the mid-way (MW) algorithm, and the convex combination (CC) as well as the generalized linear combination (GLC) approaches. The purpose of these methods is to mitigate the effect of small sample size and steering vector errors on the standard Capon beamformer (SCB). We also present sparsity based iterative beamforming algorithms, namely the iterative adaptive approach (IAA), maximum likelihood based IAA (referred to as IAA-ML) and M-SBL (multi-snapshot sparse Bayesian learning), which exploit sparsity to estimate the signal parameters. We provide a thorough evaluation of these beamforming methods in terms of power and spatial spectrum estimation accuracies, output signal-to-interference-plus-noise ratio (SINR) and resolution under various scenarios including coherent, non-coherent and distributed sources, steering vector mismatches, snapshot limitations and low signal-to-noise ratio (SNR) levels. Furthermore, we discuss the computational complexities of the algorithms and provide insights into which algorithm is the best choice under which circumstances.

Key words: Array processing; Beamforming; User parameter-free; Diagonal loading; Sparse signal representation

* Please address all correspondence to: Dr. Jian Li, Department of Electrical and Computer Engineering, P. O. Box 116130, University of Florida, Gainesville, FL 32611, USA. Phone: +1 352 392 2642. Fax: +1 352 392 0044. E-mail: li@dsp.ufl.edu.

1 Introduction

Beamforming refers to the process of combining the measurements from an array of sensors, e.g., antennas, microphones, with the goal of estimating the spatial and temporal information of the sources present in a certain environment. Many application areas including radar, sonar, acoustics, communications and medical imaging benefit from the superior performance of arrays over a single sensor [1][2]. Perhaps the simplest yet the most commonly used beamforming algorithm in practice is the data-independent delay-and-sum (DAS) beamformer, which weights and sums the measurements of each sensor so as to focus on different points in space. However, it is well-known that DAS, or more generally the data-independent beamformers, suffer from low resolution and high sidelobe problems [1][3]. Another widely employed beamforming algorithm is the standard Capon beamformer (SCB) [4] which is an optimal spatial filter that maximizes the array output signal-to-interference-plus-noise ratio (SINR) provided that the true covariance matrix and the signal steering vectors are known. In general however, the true covariance matrix is unknown and due to limited amount of data, its estimate may be quite poor and even ill-conditioned in which case the SCB cannot function at all. Moreover, array miscalibration and inaccurate data models may cause nonnegligible steering vector errors. When such problems occur, the performance of SCB degrades significantly whereas data-independent approaches show more steady performance [5]. The situation is similar with subspace based parametric beamforming methods such as multiple signal classification (MUSIC) [6][7], estimation of signal parameters via rotational invariance techniques (ESPRIT), see e.g., [3], method of direction estimation (MODE) [8][9] and weighted subspace fitting (WSF) [10], i.e., these methods are also quite sensitive to model errors. Therefore, adaptive beamforming approaches robust to the above-mentioned problems are of interest.

One of the most popular robust adaptive beamforming methods is the diagonal loading approach. Although there is quite some work in this area, see, e.g., [5], [11]-[16], most of these approaches determine the diagonal loading level either in an ad-hoc way or need user parameters that might not be available in practice. Therefore, user parameter-free robust adaptive approaches, such as the ridge regression Capon beamformers (RRCBs) [17], the mid-way (MW) method [18] and the shrinkage based beamformers [19][20] are desirable. RRCBs, which are based on the generalized sidelobe canceler (GSC) formulation of SCB, see, e.g., [1], use different ridge regression (RR) techniques such as HKB (proposed by Hoerl, Kennard and Baldwin in [21]) to improve the robustness of SCB. MW, on the other hand, makes use of the Pisarenko framework [22] for estimating the spatial power spectrum where the power estimates always lie between those obtained from SCB and DAS. MW is shown to be more robust than SCB and to have better resolution than

DAS [18]. The shrinkage based beamformers replace the sample covariance matrix used in SCB with the convex combination (CC), or more generally the linear combination (GLC), of the sample covariance matrix and a shrinkage target in an optimal mean-squared error (MSE) sense and the shrinkage parameters, which are related to the diagonal loading levels of the beamformers, can be calculated from the measurements automatically via both analytical and convex optimization techniques [19].

The rapidly growing area of sparse signal representation can be used for beamforming applications as well. The prominent methods in this area include convex optimization methods such as the least absolute shrinkage selection operator (LASSO) [23], basis pursuit (BP) [24], a weighted sparsity method [25], l_1 -SVD (singular value decomposition) [26][27] and iterative methods such as M-FOCUSS (focal underdetermined system solution) [28] and its variants [29]–[33], M-SBL (multi-snapshot sparse Bayesian learning) [34][35] and the most recent iterative adaptive approach (IAA) [36] and maximum likelihood based IAA (referred to as IAA-ML) [37]. The latter three of these algorithms are user parameter-free and hence the focus of our attention within this paper. M-SBL uses a Bayesian approach together with expectation maximization (EM) to eliminate the user parameters. IAA iteratively updates the spatial power estimates and weighting vectors based on a weighted least squares approach and it can work with few snapshots (even one) under considerable noise. IAA-ML, on the other hand, is based on likelihood maximization principles and can provide higher resolution than IAA when the snapshot number is comparable to the array size.

In this paper, we provide a thorough evaluation of the user parameter-free beamforming algorithms mentioned above in terms of array output SINR, signal-of-interest (SOI) power and spatial power spectrum estimation accuracies under various scenarios including uncorrelated, correlated and distributed sources, steering vector errors, snapshot limitations and low signal-to-noise ratio (SNR) levels.

The remainder of this paper is organized as follows. First, in Section 2, the data model used in array processing will be introduced and the basic beamforming algorithms, namely DAS and SCB, will be presented. Next in Section 3, RRCB, CC, GLC and MW, will be presented. IAA, IAA-ML and M-SBL will be presented in Section 4 and extensive numerical examples will be provided in Section 5. Finally, the paper will be concluded in Section 6.

Notation: Throughout the rest of this paper, we denote vectors and matrices by boldface lowercase and uppercase letters, respectively. The k th component of a vector $\mathbf{x}$ is written as x_k and the k th diagonal element of a diagonal matrix $\mathbf{P}$ is written as P_k . The superscript $(\cdot)^*$ denotes the conjugate transpose, $(\cdot)^T$ denotes the transpose, $\mathbf{R}^{1/2}$ denotes the Hermitian square root of the matrix

$\mathbf{R}$, $E(\cdot)$ denotes the expectation operator, $\text{tr}(\cdot)$ denotes the trace operator and $\|\cdot\|$ denotes the Frobenius norm for a matrix or the Euclidean norm for a vector. Finally, $\mathbf{I}$ denotes the identity matrix of appropriate dimension.

2 Data Model and Problem Formulation

Consider the wavefield generated by K far-field narrowband sources located at $\boldsymbol{\theta}$ where $\boldsymbol{\theta} \triangleq [\theta_1 \ \theta_2 \ \dots \ \theta_K]^T$ and θ_k is the impinging angle of the k th signal, $k = 1, \dots, K$. In the multi-snapshot case, the $M \times 1$ array output vector of an M element array in the presence of additive noise can be represented as [1][3]:

$$\mathbf{y}(n) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(n) + \mathbf{e}(n), \quad n = 1, \dots, N, \quad (1)$$

where N is the number of snapshots, $\mathbf{A}(\boldsymbol{\theta})$ is the $M \times K$ steering matrix defined as $\mathbf{A}(\boldsymbol{\theta}) \triangleq [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_K]$ and $\mathbf{s}(n) \triangleq [s_1(n) \ s_2(n) \ \dots \ s_K(n)]^T$, $n = 1, \dots, N$, is the source waveform vector. The array steering vector has different expressions depending on the array geometry and on whether the source is in the near-field or far-field of the array. For instance, in the far-field case the assumed steering vector corresponding to the k th source for a linear array is,

$$\mathbf{a}_k = \left[e^{-j\frac{2\pi f}{c_0}x_1 \sin(\theta_k)}, \dots, e^{-j\frac{2\pi f}{c_0}x_M \sin(\theta_k)} \right]^T, \quad (2)$$

where f is the center frequency, c_0 is the wave propagation velocity and x_m is the position of the m th sensor, $m = 1, \dots, M$ (see Fig. 1).

If one of the sources is considered as the SOI, the array output can be expressed as (see Eq. (1)):

$$\mathbf{y}(n) = \mathbf{a}_0 s_0(n) + \mathbf{i}(n) + \mathbf{e}(n), \quad (3)$$

where $\mathbf{a}_0$ is the array steering vector of the SOI with $\|\mathbf{a}_0\|^2 = M$, $s_0(n)$ is the signal waveform, and $\mathbf{i}(n)$ and $\mathbf{e}(n)$ are the interference and noise components, respectively (note that $\mathbf{a}_0$ can have a different form than the assumed one in Eq. (2) due to model errors). Therefore, under the assumption that the interference plus noise term and the SOI are uncorrelated, the covariance matrix of $\{\mathbf{y}(n)\}_{n=1}^N$ can be written as follows:

$$\mathbf{R} = P_0 \mathbf{a}_0 \mathbf{a}_0^* + \mathbf{Q}, \quad (4)$$

where P_0 denotes the power of the SOI, and $\mathbf{Q} = E\{(\mathbf{i}(n) + \mathbf{e}(n))(\mathbf{i}(n) + \mathbf{e}(n))^*\}$ is the interference-plus-noise covariance matrix. In practice, $\mathbf{R}$ and $\mathbf{a}_0$ are unavailable and hence are usually replaced by the sample covariance matrix $\hat{\mathbf{R}}$ and the assumed signal steering vector $\mathbf{a}$ (for instance, Eq. (2) for a far-field

linear array), respectively, where

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{n=1}^N \mathbf{y}(n) \mathbf{y}^*(n). \quad (5)$$

A beamforming algorithm aims to design a complex vector $\mathbf{w}$ in the best way to cancel the interferences and noise while keeping the SOI undistorted. Accordingly, the signal waveform and power are estimated by $\hat{s}_0(n) = \mathbf{w}^* \mathbf{y}(n)$ and $\hat{P}_0 = \mathbf{w}^* \mathbf{R} \mathbf{w}$, respectively. In addition, a spatial power spectrum of the sources present in a region can be obtained by estimating $\hat{P}_0$ over the set of all angles of interest. An important measure for the beamformer performance is the SINR which is defined as, see, e.g., [1][5]:

$$\text{SINR} = \frac{P_0 |\mathbf{w}^* \mathbf{a}_0|^2}{\mathbf{w}^* \mathbf{Q} \mathbf{w}}. \quad (6)$$

2.1 DAS

The classical DAS beamformer, which is nothing but a spatial matched filter, selects the weights as:

$$\mathbf{w}_{\text{DAS}} = \frac{\mathbf{a}}{M}. \quad (7)$$

Due to its data-independent property, DAS may suffer from lower resolution and worse interference suppression capability than the data-adaptive methods.

2.2 SCB

The most well-known adaptive beamforming technique SCB determines the array weight vector by minimizing the array output power subject to the constraint that the SOI is passed undistorted. Under ideal conditions, i.e., when the true $\mathbf{R}$ and $\mathbf{a}_0$ are known, SCB can be formulated as:

$$\underset{\mathbf{w}}{\text{minimize}} \quad \mathbf{w}^* \mathbf{R} \mathbf{w} \quad \text{subject to} \quad \mathbf{w}^* \mathbf{a}_0 = 1, \quad (8)$$

and the solution, which is given by

$$\mathbf{w}_{\text{opt}} = \frac{\mathbf{R}^{-1} \mathbf{a}_0}{\mathbf{a}_0^* \mathbf{R}^{-1} \mathbf{a}_0}, \quad (9)$$

maximizes the array output SINR defined in Eq. (6) and the corresponding optimal value is

$$\text{SINR}_{\text{opt}} = P_0 \mathbf{a}_0^* \mathbf{Q}^{-1} \mathbf{a}_0. \quad (10)$$

In practice, Eq. (9) is replaced by

$$\mathbf{w}_{\text{SCB}} = \frac{\hat{\mathbf{R}}^{-1} \mathbf{a}}{\mathbf{a}^* \hat{\mathbf{R}}^{-1} \mathbf{a}}. \quad (11)$$

In this case, SCB is no longer optimum and it is very sensitive to model errors including inaccurate covariance matrix estimates (especially when the sample size is small) and steering vector errors. The user parameter-free robust adaptive beamformers presented in the next two sections are designed to mitigate these problems.

3 Diagonal Loading Approaches

The weight vector used by the diagonal loading approach is given by:

$$\mathbf{w}_{\text{DL}} = \frac{(\hat{\mathbf{R}} + \lambda \mathbf{I})^{-1} \mathbf{a}}{\mathbf{a}^* (\hat{\mathbf{R}} + \lambda \mathbf{I})^{-1} \mathbf{a}}, \quad (12)$$

where the diagonal loading level λ is chosen manually, or sometimes based on the noise level or white noise gain constraint [11][12]. In some more advanced diagonal loading approaches, e.g., in [5], [12]–[15], the diagonal loading level is determined based on the uncertainty set of the signal steering vector. However, these approaches still require user parameters which may be hard to determine in practice. In the following, we present three user parameter-free beamforming algorithms which belong to the class of diagonal loading approaches.

3.1 RRCB

RRCB is based on the GSC reparameterization of SCB as follows:

$$\mathbf{w} = \frac{\mathbf{a}}{M} - \mathbf{B} \boldsymbol{\eta}, \quad (13)$$

where $\mathbf{B}$ is an $M \times (M - 1)$ semi-unitary matrix orthogonal to $\mathbf{a}$ (i.e., $\mathbf{B}^* \mathbf{a} = \mathbf{0}$ and $\mathbf{B}^* \mathbf{B} = \mathbf{I}$), the columns of which can be chosen as the last $M - 1$ columns of the unitary matrix in the QR-decomposition of $\mathbf{a}$. Note that $\mathbf{w}$ in Eq. (13) satisfies $\mathbf{w}^* \mathbf{a} = 1$. Then the minimization of $\mathbf{w}^* \hat{\mathbf{R}} \mathbf{w}$ with respect to $\boldsymbol{\eta}$ becomes (see Eq. (8)):

$$\underset{\boldsymbol{\eta}}{\text{minimize}} \left(\hat{\mathbf{R}}^{1/2} \frac{\mathbf{a}}{M} - \hat{\mathbf{R}}^{1/2} \mathbf{B} \boldsymbol{\eta} \right)^* \left(\hat{\mathbf{R}}^{1/2} \frac{\mathbf{a}}{M} - \hat{\mathbf{R}}^{1/2} \mathbf{B} \boldsymbol{\eta} \right). \quad (14)$$

The least-squares (LS) solution of Eq. (14) is

$$\boldsymbol{\eta}_{\text{LS}} = (\mathbf{B}^* \hat{\mathbf{R}} \mathbf{B})^{-1} \mathbf{B}^* \hat{\mathbf{R}} \mathbf{a} / M, \quad (15)$$

which when inserted into Eq. (13) yields the SCB weight vector in Eq. (11). On the other hand, RRCB applies ridge regression techniques for solving Eq. (14) to mitigate the ill-conditioning problems of SCB [17]. One of these techniques, the HKB regularization [21], provides a closed-form solution for $\boldsymbol{\eta}$ by using Tikhonov regularization [38]:

$$\boldsymbol{\eta} = (\mathbf{B}^* \hat{\mathbf{R}} \mathbf{B} + \rho_{\text{HKB}} \mathbf{I})^{-1} \mathbf{B}^* \hat{\mathbf{R}} \mathbf{a} / M, \quad (16)$$

where the loading level is specified automatically as

$$\rho_{\text{HKB}} = (M - 1) \hat{\sigma}_{\text{LS}}^2 / \|\boldsymbol{\eta}_{\text{LS}}\|^2, \quad (17)$$

and

$$\hat{\sigma}_{\text{LS}}^2 = \|\hat{\mathbf{R}}^{1/2} \mathbf{B} \boldsymbol{\eta}_{\text{LS}} - \hat{\mathbf{R}}^{1/2} \mathbf{a} / M\|^2. \quad (18)$$

It turns out that HKB is in fact a diagonal loading approach (see the derivations in [17]) and that the weight vector of HKB can be expressed as:

$$\mathbf{w}_{\text{HKB}} = \frac{(\hat{\mathbf{R}} + \rho_{\text{HKB}} \mathbf{I})^{-1} \mathbf{a}}{\mathbf{a}^* (\hat{\mathbf{R}} + \rho_{\text{HKB}} \mathbf{I})^{-1} \mathbf{a}}. \quad (19)$$

As shown in [19], in the absence of steering vector errors, an inherent problem of HKB is that ρ_{HKB} may become very large as N increases, which degrades its performance. In fact, when ρ_{HKB} is sufficiently large so that $\rho_{\text{HKB}} \mathbf{I}$ is the dominant term in Eq. (19), HKB behaves more like DAS¹, see Eq. (7).

3.2 Shrinkage Based Robust Capon Beamforming

To combat the small sample size problem introduced by the sample covariance matrix $\hat{\mathbf{R}}$, a class of shrinkage-based robust Capon beamformers have been presented in [19], where the covariance matrix estimate $\hat{\mathbf{R}}$ used in SCB is replaced by an enhanced estimate based on a shrinkage method; this enhanced estimate can be obtained by linearly combining $\hat{\mathbf{R}}$ and a shrinkage target (a given matrix with some structure) in an optimal MSE sense [39]. We consider two linear combinations here, i.e., convex combination (CC):

$$\tilde{\mathbf{R}} = \alpha \mathbf{I} + (1 - \alpha) \hat{\mathbf{R}}, \quad (20)$$

¹ Similar arguments cannot be made when there are steering vector errors and the performance of HKB may not degrade for large N as observed in our numerical examples later on.

and a more general linear combination (GLC):

$$\tilde{\mathbf{R}} = \alpha \mathbf{I} + \beta \hat{\mathbf{R}}, \quad (21)$$

where $\tilde{\mathbf{R}}$, which should satisfy $\tilde{\mathbf{R}} \geq 0$, is the enhanced estimate of $\mathbf{R}$ and we use the most commonly employed shrinkage target - the identity matrix $\mathbf{I}$ [40]. α and β in Eqs. (20) and (21) are the shrinkage parameters which are chosen by minimizing (an estimate of) the MSE of $\tilde{\mathbf{R}}$ [39], where $\text{MSE}(\tilde{\mathbf{R}}) = E\{\|\tilde{\mathbf{R}} - \mathbf{R}\|^2\}$.

The MSE minimization problem for GLC is considered first. To guarantee that $\tilde{\mathbf{R}} \geq 0$, we first get the unconstrained solutions for α and β , and then enforce them to be nonnegative². The MSE can be expressed as follows:

$$\begin{aligned} \text{MSE}(\tilde{\mathbf{R}}) &= \|\alpha \mathbf{I} - (1 - \beta)\mathbf{R}\|^2 + \beta^2 E\{\|\hat{\mathbf{R}} - \mathbf{R}\|^2\} \\ &= \alpha^2 M - 2\alpha(1 - \beta) \text{tr}(\mathbf{R}) \\ &\quad + (1 - \beta)^2 \|\mathbf{R}\|^2 + \beta^2 E\{\|\hat{\mathbf{R}} - \mathbf{R}\|^2\} \end{aligned} \quad (22)$$

and consequently the (unconstrained) optimal values for β and α can be obtained as [19]:

$$\beta_0 = \frac{\gamma}{\rho + \gamma}, \quad (23)$$

$$\alpha_0 = \nu(1 - \beta_0) = \nu \frac{\rho}{\gamma + \rho}, \quad (24)$$

where $\rho = E\{\|\hat{\mathbf{R}} - \mathbf{R}\|^2\}$, $\nu = \frac{\text{tr}(\mathbf{R})}{M}$, and $\gamma = \|\nu \mathbf{I} - \mathbf{R}\|^2$. We note that $\beta_0 \in [0, 1]$ and $\alpha_0 \geq 0$. To estimate α_0 and β_0 from the given data, we need an estimate of ρ which can be calculated as (see [41] for details):

$$\hat{\rho} = \frac{1}{N^2} \sum_{n=1}^N \|\mathbf{y}(n)\|^4 - \frac{1}{N} \|\hat{\mathbf{R}}\|^2. \quad (25)$$

In addition, we have $\gamma + \rho = E\{\|\hat{\mathbf{R}} - \nu \mathbf{I}\|^2\}$, an estimate of which is given by $\|\hat{\mathbf{R}} - \hat{\nu} \mathbf{I}\|^2$. Then we can get the nonnegativity enforced estimates of α_0 and β_0 :

$$\hat{\alpha}_0 = \min \left[\hat{\nu} \frac{\hat{\rho}}{\|\hat{\mathbf{R}} - \hat{\nu} \mathbf{I}\|^2}, \hat{\nu} \right], \quad (26)$$

² Note that the MSE minimization problems for GLC and CC can also be formulated as convex optimization problems (see [19] for details) where the constraint $\tilde{\mathbf{R}} \geq 0$ or $\alpha, \beta \geq 0$ ($\alpha \in [0, 1]$) for GLC (CC) can be readily incorporated into the convex formulation. In all of our numerical examples, the convex formulation and Eqs. (26)-(27) gave identical results as the constraints were found to be inactive (note however that they can be active depending on the application scenario, in which case the results could be different).

$$\hat{\beta}_0 = 1 - \frac{\hat{\alpha}_0}{\hat{\nu}}. \quad (27)$$

To get the shrinkage parameter estimate $\hat{\alpha}_0$ for CC, we note that α_0 and β_0 can be rewritten as $\alpha_0 = \nu\tau_0$ and $\beta_0 = 1 - \tau_0$ (see Eqs. (23) and (24), $\tau_0 = \frac{\rho}{\rho+\gamma}$), which implies that GLC reduces to CC when $\nu = 1$. Therefore, setting $\hat{\nu} = 1$, we can obtain $\hat{\alpha}_0$ from Eq. (26) for CC. Since $\nu = 1$ is generally not true in practice, GLC and CC are basically different.

Consequently, from Eqs. (20) and (21) together with Eqs. (26) and (27), we can obtain the enhanced estimates of the covariance matrix as follows:

$$\tilde{\mathbf{R}}_{\text{GLC}} = \hat{\alpha}_0 \mathbf{I} + \hat{\beta}_0 \hat{\mathbf{R}}, \quad (28)$$

and

$$\tilde{\mathbf{R}}_{\text{CC}} = \hat{\alpha}_0 \mathbf{I} + (1 - \hat{\alpha}_0) \hat{\mathbf{R}}. \quad (29)$$

Using $\tilde{\mathbf{R}}$ to replace $\hat{\mathbf{R}}$ in the SCB formulation (see Eq. (11)) yields the following shrinkage-based robust adaptive beamformers:

$$\mathbf{w}_{\text{GLC}} = \frac{\tilde{\mathbf{R}}_{\text{GLC}}^{-1} \mathbf{a}}{\mathbf{a}^* \tilde{\mathbf{R}}_{\text{GLC}}^{-1} \mathbf{a}} = \frac{\left(\frac{\hat{\alpha}_0}{\hat{\beta}_0} \mathbf{I} + \hat{\mathbf{R}} \right)^{-1} \mathbf{a}}{\mathbf{a}^* \left(\frac{\hat{\alpha}_0}{\hat{\beta}_0} \mathbf{I} + \hat{\mathbf{R}} \right)^{-1} \mathbf{a}}, \quad (30)$$

and

$$\mathbf{w}_{\text{CC}} = \frac{\tilde{\mathbf{R}}_{\text{CC}}^{-1} \mathbf{a}}{\mathbf{a}^* \tilde{\mathbf{R}}_{\text{CC}}^{-1} \mathbf{a}} = \frac{\left(\frac{\hat{\alpha}_0}{1 - \hat{\alpha}_0} \mathbf{I} + \hat{\mathbf{R}} \right)^{-1} \mathbf{a}}{\mathbf{a}^* \left(\frac{\hat{\alpha}_0}{1 - \hat{\alpha}_0} \mathbf{I} + \hat{\mathbf{R}} \right)^{-1} \mathbf{a}}. \quad (31)$$

We observe that the shrinkage-based robust adaptive beamformers are in fact diagonal loading approaches with the loading levels determined from the measurements fully automatically. Finally, the SOI power estimate for GLC (CC) is given by $\mathbf{w}_{\text{GLC}}^* \tilde{\mathbf{R}}_{\text{GLC}} \mathbf{w}_{\text{GLC}}$ ($\mathbf{w}_{\text{CC}}^* \tilde{\mathbf{R}}_{\text{CC}} \mathbf{w}_{\text{CC}}$).

3.3 MW

MW power estimation method is developed based on the Pisarenko framework [22], originally devised for temporal power spectrum estimation, which yields the following class of power estimates [18]:

$$\hat{P}_0(r) = \begin{cases} \frac{1}{M} \left(\frac{\mathbf{a}^* \hat{\mathbf{R}}^r \mathbf{a}}{M} \right)^{1/r}, & \text{for } r \neq 0, \\ \frac{1}{M} \exp \left(\frac{\mathbf{a}^* \ln(\hat{\mathbf{R}}) \mathbf{a}}{M} \right), & \text{for } r = 0. \end{cases} \quad (32)$$

To deal with the problems of DAS and SCB, MW uses $r = 0$ in Eq. (32), which takes a mid-way position between $r = -1$ (corresponding to the power

estimate of SCB) and $r = 1$ (corresponding to the power estimate of DAS). The power estimate of MW can be expressed as:

$$\hat{P}_{0,\text{MW}} = \frac{1}{M} \exp \left(\frac{\mathbf{a}^* \ln(\hat{\mathbf{R}}) \mathbf{a}}{M} \right). \quad (33)$$

Note that taking the logarithm of $\hat{\mathbf{R}}$ reduces the dynamic range of the eigenvalues of $\hat{\mathbf{R}}$, which is also one of the main objectives of the diagonal loading approaches. However, MW compensates this dynamic range reduction by using the inverse logarithmic function, i.e., the exponential, when estimating the powers unlike the shrinkage approaches [18].

MW can obtain the power estimates without obtaining an explicit expression for the weighting vector. For SOI waveform estimation, however, a beamformer has to be designed which can be obtained by minimizing the white-noise gain under the constraints that the output power is equal to the MW power estimate given by Eq. (33) and that the SOI is passed undistorted:

$$\begin{aligned} & \underset{\mathbf{w}}{\text{minimize}} \quad \|\mathbf{w}\|^2 \\ & \text{subject to} \quad \mathbf{w}^* \hat{\mathbf{R}} \mathbf{w} = \hat{P}_{0,\text{MW}}, \quad \mathbf{w}^* \mathbf{a} = 1. \end{aligned} \quad (34)$$

The above problem can be solved by using the Lagrange method and the solution is [18]:

$$\mathbf{w}_{\text{MW}} = \frac{(\mathbf{I} + \lambda \hat{\mathbf{R}})^{-1} \mathbf{a}}{\mathbf{a}^* (\mathbf{I} + \lambda \hat{\mathbf{R}})^{-1} \mathbf{a}}, \quad (35)$$

where the Lagrange parameter λ can be calculated using Newton's method based on the following equation:

$$\frac{\mathbf{a}^* (\mathbf{I} + \lambda \hat{\mathbf{R}})^{-1} \hat{\mathbf{R}} (\mathbf{I} + \lambda \hat{\mathbf{R}})^{-1} \mathbf{a}}{(\mathbf{a}^* (\mathbf{I} + \lambda \hat{\mathbf{R}})^{-1} \mathbf{a})^2} = \hat{P}_{0,\text{MW}}. \quad (36)$$

Note from Eq. (35) that the MW beamformer can also be considered as a user parameter-free diagonal loading approach where the loading level is calculated by solving Eq. (36).

4 Sparse approaches

In this section we consider sparse approaches to beamforming, namely IAA, IAA-ML and M-SBL. For this purpose, consider the data model in Eq. (1) where K is replaced by the number of potential source locations in the field (or the number of scanning points) to avoid the need for estimating the true number of sources which is usually unknown in practice. As a result, K will

be much larger than the true number of sources and $\mathbf{s}(n)$, $n = 1, \dots, N$, in Eq. (1) will contain only a few non-zero elements, i.e., it will be sparse. Consequently, sparse approaches could be used for signal waveform (and hence power) estimation.

4.1 IAA

Let $\mathbf{P}$ be a $K \times K$ diagonal matrix, whose diagonal contains the signal power at each angle on the scanning grid and define the covariance matrix of the interference and noise as (see Eq. (4))

$$\bar{\mathbf{Q}}_k = \bar{\mathbf{R}} - P_k \mathbf{a}_k \mathbf{a}_k^*, \quad (37)$$

where $\bar{\mathbf{R}} = \mathbf{A}(\boldsymbol{\theta}) \mathbf{P} \mathbf{A}^*(\boldsymbol{\theta})$ and k is the grid index of the current SOI. Then, the weighted least squares cost function is given by, see, e.g., [3][42][43],

$$\sum_{n=1}^N \|\mathbf{y}(n) - s_k(n) \mathbf{a}_k\|_{\bar{\mathbf{Q}}_k}^2, \quad (38)$$

where $\|\mathbf{x}\|_{\bar{\mathbf{Q}}_k}^2 \triangleq \mathbf{x}^* \bar{\mathbf{Q}}_k^{-1} \mathbf{x}$. Minimizing Eq. (38) with respect to $s_k(n)$, $n = 1, \dots, N$, yields,

$$\hat{s}_k(n) = \frac{\mathbf{a}_k^* \bar{\mathbf{Q}}_k^{-1} \mathbf{y}(n)}{\mathbf{a}_k^* \bar{\mathbf{Q}}_k^{-1} \mathbf{a}_k} = \frac{\mathbf{a}_k^* \bar{\mathbf{R}}^{-1} \mathbf{y}(n)}{\mathbf{a}_k^* \bar{\mathbf{R}}^{-1} \mathbf{a}_k}, \quad (39)$$

where the second equality follows from Eq. (37) and the matrix inversion lemma, see, e.g., [3]. The IAA power estimates are then given by $\hat{P}_k = \frac{1}{N} \sum_{n=1}^N |\hat{s}_k(n)|^2$, $k = 1, \dots, K$. Since IAA requires $\bar{\mathbf{R}}$, which depends on the unknown signal powers, it must be implemented as an iterative approach; the initialization is done by DAS. The IAA algorithm is summarized in Table 1. Note that when $\bar{\mathbf{R}} = \mathbf{I}$, the power estimate of IAA reduces to the DAS power estimate (see Eqs. (7) and (39)). Also, IAA can be thought in the form of Eq. (9) where $\mathbf{R}$ is estimated first and then the filter weights are calculated using this estimate. In IAA, $\hat{\mathbf{P}}$ and hence $\bar{\mathbf{R}}$ are obtained from the signal estimates of the previous iteration and not from the data snapshots.

Note that for IAA (as well as IAA-ML and M-SBL, see below) a scanning grid has to be set and the signal parameters corresponding to these grids are estimated jointly, i.e., the signal parameter estimate for every grid point is readily available once the algorithm is run. This is unlike in the previously mentioned methods (DAS, SCB, HKB, CC, GLC and MW) in which only the SOI power is estimated. To get the signal parameter estimates at the SOI, we can define the SOI weight vector of IAA as $\mathbf{w}_{\text{IAA}} = \mathbf{w}_i$, where i is the index of the scanning grid point corresponding to the assumed direction-of-arrival (DOA) of the SOI.

Table 1
The IAA algorithm

$$\begin{aligned} \hat{s}_k(n) &= \mathbf{a}_k^* \mathbf{y}(n) / M, \quad n = 1, \dots, N, k = 1 \dots, K \\ \hat{P}_k &= \frac{1}{N} \sum_{n=1}^N |\hat{s}_k(n)|^2, \quad k = 1, \dots, K \\ \mathbf{repeat} \\ &\quad \bar{\mathbf{R}} = \mathbf{A}(\boldsymbol{\theta}) \hat{\mathbf{P}} \mathbf{A}^*(\boldsymbol{\theta}) \\ &\quad \mathbf{for} \ k = 1, \dots, K \\ &\quad \quad \mathbf{w}_k = \frac{\bar{\mathbf{R}}^{-1} \mathbf{a}_k}{\mathbf{a}_k^* \bar{\mathbf{R}}^{-1} \mathbf{a}_k} \\ &\quad \quad \hat{P}_k = \mathbf{w}_k^* \hat{\mathbf{R}} \mathbf{w}_k \\ &\quad \mathbf{end} \ \mathbf{for} \\ &\quad \mathbf{until} \ (\text{convergence}) \end{aligned}$$

4.2 IAA-ML

IAA-ML [37] minimizes the negative log-likelihood function of $\{\mathbf{y}(n)\}_{n=1}^N$, i.e.,

$$\ln |\bar{\mathbf{R}}| + \frac{1}{N} \sum_{n=1}^N \mathbf{y}^*(n) \bar{\mathbf{R}}^{-1} \mathbf{y}(n), \quad (40)$$

with respect to the unknowns in $\bar{\mathbf{R}}$, where it was assumed that the received signal is a complex multivariate zero-mean Gaussian random vector with covariance matrix $\bar{\mathbf{R}}$ and that the snapshots are independent and identically distributed (i.i.d.) [1]. Assume that $\mathbf{Q}_k$ is known and that the signal power at θ_k is to be estimated. Using the fact that $|\mathbf{I} + \mathbf{A}\mathbf{B}| = |\mathbf{I} + \mathbf{B}\mathbf{A}|$ and the matrix inversion lemma together with Eq. (37), it can be shown that minimizing Eq. (40) with respect to P_k is equivalent to minimizing,

$$f(P_k) \triangleq \ln(1 + P_k \mathbf{a}_k^* \bar{\mathbf{Q}}_k^{-1} \mathbf{a}_k) - \frac{P_k \mathbf{a}_k^* \bar{\mathbf{Q}}_k^{-1} \hat{\mathbf{R}} \bar{\mathbf{Q}}_k^{-1} \mathbf{a}_k}{1 + P_k \mathbf{a}_k^* \bar{\mathbf{Q}}_k^{-1} \mathbf{a}_k}. \quad (41)$$

Setting the first derivative of Eq. (41) with respect to P_k to zero, i.e., $f'(\tilde{P}_k) = 0$, gives:

$$\tilde{P}_k = \frac{\mathbf{a}_k^* \bar{\mathbf{Q}}_k^{-1} (\hat{\mathbf{R}} - \bar{\mathbf{Q}}_k) \bar{\mathbf{Q}}_k^{-1} \mathbf{a}_k}{(\mathbf{a}_k^* \bar{\mathbf{Q}}_k^{-1} \mathbf{a}_k)^2}. \quad (42)$$

Moreover, it can be shown that $f''(\tilde{P}_k) > 0$, which means that $\tilde{P}_k$ is the unique minimizer of $f(P_k)$. In principle $\tilde{P}_k$ may be negative; therefore, the nonnegativity of the power estimates is enforced at each iteration by setting the negative estimates to zero. Accordingly, the IAA-ML power estimate is

obtained as

$$\hat{P}_k = \max \left(0, P_k + \frac{\mathbf{a}_k^* \bar{\mathbf{R}}^{-1} (\hat{\mathbf{R}} - \bar{\mathbf{R}}) \bar{\mathbf{R}}^{-1} \mathbf{a}_k}{(\mathbf{a}_k^* \bar{\mathbf{R}}^{-1} \mathbf{a}_k)^2} \right), \quad (43)$$

where we have used the matrix inversion lemma in Eq. (42) to replace $\bar{\mathbf{Q}}_k$ by $\bar{\mathbf{R}}$. As $P_k = \tilde{P}_k$ is the unique minimizer of $f(P_k)$, $\hat{P}_k$ minimizes $f(P_k)$ subject to $P_k \geq 0$. Since computing $\hat{P}_k$ requires knowledge of P_k and $\bar{\mathbf{R}}$, the algorithm must be implemented iteratively; the initialization of $\hat{P}_k$ is done with DAS. IAA-ML is outlined in Table 2, where the inverse covariance matrix is calculated efficiently using the matrix inversion lemma. The sorting procedure helps driving the estimates for the potentially source free locations to zero so that the other estimations can be done more accurately. Finally, IAA-ML is locally convergent due to cyclically maximizing the likelihood function [37].

IAA-ML estimates the signal powers, but if desired the waveforms can be obtained as well by a minimum mean-squared error (MMSE) estimator. Since $\{\mathbf{y}(n)\}_{n=1}^N$ and $\{\mathbf{s}(n)\}_{n=1}^N$ are jointly Gaussian distributed with means zero, the MMSE estimate of $\{\mathbf{s}(n)\}_{n=1}^N$ given the observations $\{\mathbf{y}(n)\}_{n=1}^N$ is [44][45]

$$\hat{\mathbf{s}}(n) = \mathbf{W}^* \mathbf{y}(n) = \mathbf{P} \mathbf{A}^* (\boldsymbol{\theta}) (\mathbf{A}(\boldsymbol{\theta}) \mathbf{P} \mathbf{A}^* (\boldsymbol{\theta}))^{-1} \mathbf{y}(n) \quad (44)$$

for $n = 1, \dots, N$. The IAA-ML signal waveform estimates are obtained by using $\hat{\mathbf{P}}$ obtained with IAA-ML in lieu of $\mathbf{P}$ in Eq. (44). Accordingly, the SOI weight vector of IAA-ML, namely $\mathbf{w}_{\text{IAA-ML}}$, is given by the i th column of the matrix $\mathbf{W}$ in Eq. (44), where i is the grid index corresponding to the assumed DOA of the SOI.

4.3 M-SBL

A Bayesian approach can also be used to estimate the signal waveforms using various priors to enforce sparsity. An important algorithm in this context is the sparse Bayesian learning (SBL) approach [46][47], and M-SBL, the multi-snapshot extension of it [34][35], which uses a zero-mean Gaussian prior with a distinct variance for each $\tilde{\mathbf{s}}_k = [s_k(1), \dots, s_k(N)]^T$, i.e.,

$$p(\tilde{\mathbf{s}}_k; \gamma_k) = \mathcal{CN}(\mathbf{0}, \gamma_k \mathbf{I}), \quad k = 1, \dots, K \quad (45)$$

where $\boldsymbol{\gamma} = [\gamma_1, \dots, \gamma_K]^T$ is a vector of K hyperparameters controlling the prior variance of the elements of $\tilde{\mathbf{s}}_k$. The Bayesian estimate of the signal waveforms is obtained by using a type-II likelihood maximization and an EM algorithm resulting in the algorithm summarized in Table 3. From the table, we can define the SOI weight vector of M-SBL, namely $\mathbf{w}_{\text{M-SBL}}$, as the i th column of the matrix $\mathbf{W}$ with i denoting the grid index corresponding to the assumed DOA of the SOI.

Table 2
The IAA-ML algorithm

$$\hat{s}_k(n) = \mathbf{a}_k^* \mathbf{y}(n) / M, \quad n = 1, \dots, N, k = 1, \dots, K$$

$$\hat{P}_k = \frac{1}{N} \sum_{n=1}^N |\hat{s}_k(n)|^2, \quad k = 1, \dots, K$$

$$\bar{\mathbf{R}}^{-1} = (\mathbf{A}(\boldsymbol{\theta}) \hat{\mathbf{P}} \mathbf{A}^*(\boldsymbol{\theta}))^{-1}$$

repeat

Adjust $[i_1, \dots, i_K]$ such that $\hat{P}_{i_1} \leq \hat{P}_{i_2} \leq \dots \leq \hat{P}_{i_K}$

for $k = 1, \dots, K$

$$\hat{P}_{i_k}^{\text{previous}} = \hat{P}_{i_k}$$

$$\hat{P}_{i_k} = \max \left(0, \hat{P}_{i_k} + \frac{\mathbf{a}_{i_k}^* \bar{\mathbf{R}}^{-1} (\hat{\mathbf{R}} - \bar{\mathbf{R}}) \bar{\mathbf{R}}^{-1} \mathbf{a}_{i_k}}{(\mathbf{a}_{i_k}^* \bar{\mathbf{R}}^{-1} \mathbf{a}_{i_k})^2} \right)$$

$$\bar{\mathbf{R}}^{-1} = \bar{\mathbf{R}}^{-1} - \frac{(\hat{P}_{i_k} - \hat{P}_{i_k}^{\text{previous}}) \bar{\mathbf{R}}^{-1} \mathbf{a}_{i_k} \mathbf{a}_{i_k}^* \bar{\mathbf{R}}^{-1}}{1 + (\hat{P}_{i_k} - \hat{P}_{i_k}^{\text{previous}}) \mathbf{a}_{i_k}^* \bar{\mathbf{R}}^{-1} \mathbf{a}_{i_k}}$$

end for

until (convergence)

5 Numerical Examples

In this section we present a number of experiments demonstrating the benefits and limitations of SCB, HKB, MW, GLC, CC, M-SBL, IAA and IAA-ML. We consider the beamformer output SINR, SOI power estimate and spatial power spectrum estimate as our performance metrics. The examples will contain: *i*) uncorrelated and correlated point sources and distributed sources, *ii*) different snapshot numbers (N can range from one, for instance, in underwater acoustics measurements [48][49], to hundreds, for instance, in aeroacoustic measurements [50]), *iii*) different SNR values, and *iv*) steering vector errors. Finally, we discuss the computational complexities of each algorithm and provide some insights into which algorithm is more suitable under what circumstances.

5.1 Simulation Details

We consider a uniform linear array with $M = 10$ sensors and half-wavelength inter-element spacing. The far-field narrowband signal waveforms and the additive noise signals are assumed to be temporally white circularly symmetric complex Gaussian random processes with zero mean and a certain variance.

Table 3
The M-SBL algorithm

$\gamma = \mathbf{1}, \lambda_s > 0$

repeat

 $\mathbf{\Gamma} = \mathbf{diag}(\gamma_k)$

 $\mathbf{\Sigma}_t = \mathbf{A}(\boldsymbol{\theta})\mathbf{\Gamma}\mathbf{A}^*(\boldsymbol{\theta}) + \lambda_s\mathbf{I}$

 $\mathbf{\Sigma} = \mathbf{\Gamma}(\mathbf{I} - \mathbf{A}^*(\boldsymbol{\theta})\mathbf{\Sigma}_t^{-1}\mathbf{A}(\boldsymbol{\theta})\mathbf{\Gamma})$

 $\hat{\mathbf{s}}(n) = \mathbf{W}^*\mathbf{y}(n) = \mathbf{\Gamma}\mathbf{A}^*(\boldsymbol{\theta})\mathbf{\Sigma}_t^{-1}\mathbf{y}(n), \quad n = 1, \dots, N$

 $\lambda_s = \frac{\frac{1}{N} \sum_{n=1}^N \|\mathbf{y}(n) - \mathbf{A}(\boldsymbol{\theta})\hat{\mathbf{s}}(n)\|^2}{M - K + \sum_{k=1}^K (\Sigma_k / \gamma_k)}$

 $\gamma_k = \frac{1}{N} \sum_{n=1}^N |\hat{s}_k(n)|^2 / (1 - \gamma_k^{-1} \Sigma_k), \quad k = 1, \dots, K$

until (convergence)

The noise is further assumed to be spatially white and independent of the sources. The SNR for each source is defined as:

$$\text{SNR}_k = 10 \log_{10} \left(\frac{P_k}{\sigma^2} \right) \text{ dB}, \quad k = 1, \dots, K, \quad (46)$$

where P_k and σ^2 denote the variances of the k th source and noise, respectively. The scanning grid for IAA, IAA-ML and M-SBL is uniform in the range from -90° to 90° with 1° increment between adjacent grid points. For each example, 100 Monte-Carlo trials are performed and average results are presented. The steering vector errors are simulated by perturbing each element of the assumed steering vector (corresponding to both the SOI and interferences) using independent identically distributed (i.i.d.) zero-mean circularly symmetric complex Gaussian random variables with variance γ^2 . The perturbed steering vectors are then normalized so that their norm square equals M (see Eq. (3)).

5.2 Point Sources

First, we consider the scenario of uncorrelated sources. Three sources with powers 10, 20 and 20 dB are assumed to be present at 0° , 20° and 60° , re-

spectively. We consider the first signal as the SOI and the other two signals as interferences.

Fig. 2 shows the beamformer output SINR and SOI power estimates versus the snapshot number N in the absence of steering vector errors for SNR = 10 dB (here SNR means the power ratio between the SOI and noise, and the noise power is 0 dB). Among the tested methods, SCB, HKB and MW cannot work properly when the sample covariance matrix is rank deficient and hence the results of these algorithms are not shown when $N < M$. In addition, GLC and CC cannot function properly when $N = 1$ since $\hat{\rho} = 0$ (see Eqs. (25)-(27)) in that case and as a result GLC and CC reduce to SCB. From Fig. 2(a), we observe that the performance of HKB degrades when N exceeds 100 which is due to the problem in choosing the diagonal loading level as mentioned earlier. All the other robust adaptive beamformers converge to SINR_{opt} faster than SCB as N increases with IAA having the best performance. Fig. 2(b) shows that IAA and GLC provide more accurate SOI power estimates than the other algorithms for all sample sizes considered.

Fig. 3 shows the output SINR and SOI power estimates for different SNR values obtained by varying the noise power for $N = 20$. We observe that IAA shows the best SINR performance within the SNR range considered and its performance is quite close to the optimum value. Moreover, IAA together with GLC and MW provide good SOI power estimates for all SNR values considered. IAA-ML and M-SBL also provide good SINR performance. However, they underestimate the SOI power for relatively low SNR.

Next, we examine the robustness of the beamformers to array steering vector errors. Fig. 4 shows the output SINR and SOI power estimates versus the perturbation variance γ^2 for SNR = 10 dB and $N = 20$. In this scenario, GLC and MW show the best SINR performance. GLC and IAA followed by MW provide more accurate SOI power estimates than the other methods. However, when N is relatively large, GLC tends to choose a small diagonal loading level and thus is less effective in combatting the steering vector errors. As shown in Fig. 5, which considers the same scenario as Fig. 4 except that N is increased to 100, the performance of GLC degrades compared to the case when $N = 20$. With more snapshots, MW shows the best SINR performance while IAA still provides the most accurate SOI power estimates among all the methods.

We now consider an example consisting of two correlated sources. The SOI with 10 dB power is at 0° and an interference with 20 dB power is at 20° . Fig. 6 shows the beamformer output SINR and SOI power estimates versus the correlation coefficient between the SOI and the interference for $N = 20$. Interestingly, we observe that the performances (both SINR and SOI power estimates) of IAA, IAA-ML and M-SBL remain almost unchanged when the correlation coefficient increases. On the other hand, the performances of the

other methods degrade as the interference becomes more correlated with the SOI and these algorithms fail almost completely when the two sources become coherent. IAA shows the best performance in this case as well.

Finally, we present the spatial power spectrum estimates of the beamformers for various scenarios when $N = 20$. DAS is also included in this comparison as a reference. First, we consider five uncorrelated sources with powers 10, 15, 30, 25 and 20 dB at -45° , -35° , 0° , 5° and 60° , respectively. The noise power is assumed to be 0 dB. From Fig. 7, we note that IAA-ML and M-SBL give the best resolution followed by SCB, CC and HKB. However, these algorithms underestimate the powers of some sources, especially those with relatively low SNR. On the other hand, IAA always provides accurate power estimates. The second example is the same as the previous one, except that the angular locations of the sources are -45.6° , -35.1° , 0° , 5.2° and 60.5° , i.e., the sources are no longer on the grid points of the sparse algorithms. The results are shown in Fig. 8. Comparing Fig. 8 with Fig. 7, we observe that the performances of the algorithms with relatively low resolution, i.e., IAA, GLC, MW and DAS are virtually unchanged and the performances of the other methods degrade slightly. Thirdly, we compare the spatial power spectrum estimates in the presence of steering vector errors for a case where three uncorrelated sources with powers 15, 30 and 20 dB are located at -15° , 0° and 60° and the perturbation variance is -10 dB. The performances of the beamformers are shown in Fig. 9, from which we can observe that IAA gives the best power estimates. Finally, Fig. 10 shows the power and location estimates for two coherent sources at -5° and 5° with powers 10 and 20 dB, respectively. Note that only the sparse algorithms can function well in this case, which is consistent with the results shown in Fig. 6. Again, IAA-ML and M-SBL have better resolution than IAA, while IAA provides more accurate power estimates, especially for relatively low SNR.

5.3 Distributed Sources

So far we have studied point sources, i.e., sources impinging from a single location in space. In the following examples, we examine the performance of the beamformers for distributed sources as well [51]-[55]. The array output vector for K spatially distributed sources can be represented as [52][53] (see also Eq. (1)):

$$\begin{aligned} \mathbf{y}(n) &= \sum_{k=1}^K \int_{-\pi/2}^{\pi/2} \mathbf{a}(\theta) s_k(n) g_k(\theta, \boldsymbol{\psi}_k) d\theta + \mathbf{e}(n), \\ &= \sum_{k=1}^K \mathbf{c}(\boldsymbol{\psi}_k) s_k(n) + \mathbf{e}(n) \end{aligned} \quad (47)$$

where $\mathbf{c}(\boldsymbol{\psi}_k) = \int_{-\pi/2}^{\pi/2} \mathbf{a}(\theta) g_k(\theta, \boldsymbol{\psi}_k) d\theta$ is the generalized steering vector, $g(\theta, \boldsymbol{\psi}_k)$ is the angular signal density and $\boldsymbol{\psi}_k$ is the location parameter vector of the k th source [53]. For a uniformly distributed source, $\boldsymbol{\psi}_k = [\theta_k, \delta\theta_k]^T$, where θ_k denotes the central angle and $\delta\theta_k$ denotes the angular spread of the k th source. Accordingly, the m th component of the steering vector $\mathbf{c}(\boldsymbol{\psi}_k)$ due to the k th source can be approximated as (for small values of $\delta\theta_k$):

$$\mathbf{c}(\boldsymbol{\psi}_k)_m \approx e^{-j\frac{2\pi f}{c_0} x_m \sin(\theta_k)} \text{sinc}\left(\frac{2f}{c_0} x_m \delta\theta_k \cos(\theta_k)\right) \quad (48)$$

where $\text{sinc}(x) = \sin(\pi x)/(\pi x)$ [52].

Fig. 11 shows the SINR and SOI power estimates of the algorithms versus N for two uncorrelated spatially distributed sources, simulated using Eqs. (47) and (48), that are located at $\theta_1 = 0^\circ$ and $\theta_2 = 20^\circ$ with $\delta\theta_1 = \delta\theta_2 = 2^\circ$. The steering vectors are normalized such that $\|\mathbf{c}(\boldsymbol{\psi}_k)\|^2 = M$, $k = 1, 2$. We assume that the first source is the SOI with 10 dB power and the other one is the interference with 20 dB power. Fig. 12 shows the corresponding results obtained by varying SNR. We note that IAA shows the best performance. MW and GLC also provide good SOI power estimates. In addition, the SINR performances of IAA-ML and M-SBL are good while their SOI power estimates are not so accurate especially for relatively low SNR. Finally, the spatial power estimates are shown in Fig. 13 for two uncorrelated spatially distributed sources, where one of the sources is located at $\theta_1 = -5^\circ$ with 10 dB power and the other at $\theta_2 = 5^\circ$ with 20 dB power and $\delta\theta_1 = \delta\theta_2 = 3^\circ$. As can be observed from the plots, IAA provides the more accurate power and location estimates in this case.

5.4 Complexity Analysis

The computational complexity of DAS is $\mathcal{O}(M^2)$ whereas the complexities of SCB, HKB, CC, GLC and MW are $\mathcal{O}(M^3)$, mainly because of the matrix inversion operation. Assuming $K \gg M$, the complexity of each IAA, IAA-ML and M-SBL iteration is $\mathcal{O}(M^2 K)$ per iteration³. Note however that if the spatial estimate of the sources in the whole region is desired, DAS has complexity $\mathcal{O}(M^2 K)$ and all the other methods have complexity $\mathcal{O}(M^2 K)$ as well, assuming $K \gg M$. In all our simulations, IAA and IAA-ML were run for a maximum of 15 iterations and, in general, IAA and IAA-ML converged in at most half the number of iterations necessary for M-SBL to converge.

³ The complexity of M-SBL does not include N per the discussion in [34]. This stems from the fact that the M-SBL iterations can be implemented by using $\hat{\mathbf{R}}$ instead of $\{\mathbf{y}(n)\}_{n=1}^N$.

Table 4
Summary of results

	Correct steering vector	Steering vector errors	Correlated Sources	Distributed Sources
SINR	IAA	MW ^a	IAA	IAA
	M-SBL	GLC ^b	M-SBL	M-SBL
	IAA-ML	IAA	IAA-ML	IAA-ML
SOI Power Estimate	IAA	IAA	IAA	IAA
	GLC ^b	GLC ^b	IAA-ML	GLC ^b
	MW ^a	MW ^a	M-SBL	MW ^a
Angle estimation accuracy	IAA-ML	IAA-ML	IAA-ML	IAA-ML
	M-SBL	M-SBL	M-SBL	M-SBL
	SCB ^a	MW ^a	IAA	IAA

^a $N \approx M$ or larger, ^b N not very large

5.5 Overall Assessments

Based on all the numerical examples above, IAA shows the best beamformer output SINR except in the presence of array calibration errors. Also, IAA provides the most accurate power estimates in all the cases. IAA-ML and M-SBL provide the highest resolution, which is useful for DOA estimation. However, the SOI power estimation performances of these two methods are more sensitive to low SNR than for the other methods. One desirable property of the sparse algorithms is that their performances are not affected much by the presence of correlated (even coherent) sources, while all the other algorithms fail in this case. Except for the correlated source case, GLC and MW provide good overall performance, especially in the presence of steering vector errors. However, GLC does not work well for large snapshot numbers when there are steering vector errors and in the spatially distributed sources case, while MW cannot be used for low snapshot number cases. CC provides similar performance as SCB except for small sample sizes in which case it outperforms SCB. We summarize our empirical observations in Table 4 which lists the algorithms that show good performance for a given scenario.

6 Conclusions

In this paper we have reviewed and compared the following beamforming methods: the delay-and-sum (DAS) method, the standard Capon beamformer

(SCB), the diagonal loading approaches of ridge regression Capon beamformers (RRCB) and of mid-way (MW), the shrinkage based diagonal loading approaches of convex combination (CC) and of generalized linear combination (GLC), and the sparsity based beamformers of the iterative adaptive approach (IAA), of the maximum likelihood based IAA (referred to as IAA-ML) and of M-SBL (multi-snapshot sparse Bayesian learning). These algorithms have been evaluated under various scenarios according to their SINR as well as to their SOI power estimation and spatial power estimation accuracies. General guidelines have been offered to assist selecting the most suitable algorithm in a given application scenario.

Acknowledgments

This work was supported in part by the National Science Foundation (NSF) under Grant No. ECCS-0729727, by the U.S. Army Medical Command under Grant No. DAAB15-00-C-1024, by the National Aeronautics and Space Administration (NASA) under Grant No. NNX07AO15A, and by the Swedish Research Council (VR). Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

References

- [1] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. New York, NY: John Wiley & Sons, 2002.
- [2] S. Haykin, J. Reilly, V. Kezys, and E. Vertatschitsch, "Some aspects of array signal processing," *IEEE Proceedings of Radar and Signal Processing*, vol. 139, pp. 1–26, February 1992.
- [3] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ: Prentice-Hall, 2005.
- [4] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, pp. 1408–1418, August 1969.
- [5] J. Li and P. Stoica, eds., *Robust Adaptive Beamforming*. New York, NY: John Wiley & Sons, 2005.
- [6] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. AP-34, pp. 276–280, March 1986.

- [7] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-37, pp. 720–741, May 1989.
- [8] P. Stoica and K. C. Sharman, "Novel eigenanalysis method for direction estimation," *IEE Proceedings, Pt. F*, vol. 137, pp. 19–26, February 1990.
- [9] P. Stoica and K. C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-38, pp. 1132–1143, July 1990.
- [10] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Transactions on Signal Processing*, vol. 39, pp. 2436–2449, November 1991.
- [11] H. Cox, R. Zeskind, and M. Owen, "Robust adaptive beamforming," *Proceedings of IEEE*, vol. ASSP-35, pp. 1365–1375, October 1987.
- [12] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization," *IEEE Transactions on Signal Processing*, vol. 51, pp. 313–324, February 2003.
- [13] P. Stoica, Z. Wang, and J. Li, "Robust Capon beamforming," *IEEE Signal Processing Letters*, vol. 10, pp. 172–175, June 2003.
- [14] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Transactions on Signal Processing*, vol. 51, pp. 1702–1715, July 2003.
- [15] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Transactions on Signal Processing*, vol. 53, pp. 1684–1696, May 2005.
- [16] Y. C. Eldar, A. Nehorai, and P. S. LaRosa, "A competitive mean-squared error approach to beamforming," *IEEE Transactions on Signal Processing*, vol. 55, pp. 5143–5154, November 2007.
- [17] Y. Selén, R. Abrahamsson, and P. Stoica, "Automatic robust adaptive beamforming via ridge regression," *Signal Processing*, vol. 88, pp. 33–49, 2008.
- [18] P. Stoica, J. Li, and X. Tan, "On spatial power spectrum and signal estimation using the Pisarenko framework," *IEEE Transactions on Signal Processing*, vol. 56, pp. 5109–5119, October 2008.
- [19] L. Du, J. Li, and P. Stoica, "Fully automatic computation of diagonal loading levels for robust adaptive beamforming," submitted to *IEEE Transactions on Aerospace and Electronic Systems*, 2007.
- [20] J. Li, L. Du, and P. Stoica, "Fully automatic computation of diagonal loading levels for robust adaptive beamforming," *The 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Las Vegas, Nevada, USA, March 2008.

- [21] A. E. Hoerl, R. W. Kennard, and K. F. Baldwin, "Ridge regression: some simulations," *Communication in Statistics: Theory and Methods*, vol. 4, pp. 105–123, 1975.
- [22] V. F. Pisarenko, "On the estimation of spectra by means of non-linear functions of the covariance matrix," *Geophysical Journal of the Royal Astronomical Society*, vol. 28, pp. 511–531, June 1972.
- [23] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society*, vol. 58, no. 1, pp. 267–288, 1996.
- [24] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.
- [25] E. J. Candes, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *ArXiv e-prints*, vol. 711, Nov. 2007.
- [26] D. M. Malioutov, "A sparse signal reconstruction perspective for source localization with sensor arrays," Master's thesis, MIT, July 2003.
- [27] D. M. Malioutov, M. Çetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3010–3022, 2005.
- [28] S. F. Cotter, B. D. Rao, E. Kjersti, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Transactions on Signal Processing*, vol. 53, no. 7, pp. 2477–2488, 2005.
- [29] B. D. Rao and K. Kreutz-Delgado, "An affine scaling methodology for best basis selection," *IEEE Transactions on Signal Processing*, vol. 47, no. 1, pp. 187–200, 1999.
- [30] K. Kreutz-Delgado, J. F. Murray, B. D. Rao, K. Engan, T. Lee, and T. J. Sejnowski, "Dictionary learning algorithms for sparse representation," *Neural Computation*, vol. 15, no. 2, pp. 349–396, 2003.
- [31] M. A. T. Figueiredo, "Adaptive sparseness for supervised learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1150–1159, 2003.
- [32] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," *The 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Las Vegas, Nevada, USA, April 2008.
- [33] J. J. Fuchs, "Convergence of a sparse representations algorithm applicable to real or complex data," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 598–605, 2007.
- [34] D. P. Wipf and B. D. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3704–3716, 2007.

- [35] D. P. Wipf and S. Nagarajan, "Beamforming using the relevance vector machine," *ICML '07: Proceedings of the 24th international conference on Machine learning*, pp. 1023–1030, 2007.
- [36] T. Yardibi, J. Li, P. Stoica, M. Xue, and A. B. Baggeroer, "Source localization and sensing: A nonparametric iterative adaptive approach based on weighted least squares," submitted to *IEEE Transactions on Aerospace and Electronic Systems*, 2008.
- [37] T. Yardibi, J. Li, and P. Stoica, "Nonparametric and sparse signal representations in array processing via iterative adaptive approaches," *42nd IEEE Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, October 2008.
- [38] A. Tikhonov and V. Arsenin, *Solution of Ill-posed Problems*. Winston, Washington, DC: Winston & Sons, 1977.
- [39] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, vol. 88, pp. 365–411, 2004.
- [40] J. Schafer and K. Strimmer, "A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics," *Statistical Applications in Genetics and Molecular Biology*, vol. 4, Art. No. 32, 2005.
- [41] P. Stoica, J. Li, X. Zhu, and J. R. Guerci, "On using a priori knowledge in space-time adaptive processing," *IEEE Transactions on Signal Processing*, vol. 56, pp. 2598–2602, June 2008.
- [42] J. Li and P. Stoica, "An adaptive filtering approach to spectral estimation and SAR imaging," *IEEE Transactions on Signal Processing*, vol. 44, pp. 1469–1484, June 1996.
- [43] P. Stoica, H. Li, and J. Li, "A new derivation of the APES filter," *IEEE Signal Processing Letter*, vol. 6, pp. 205–206, August 1999.
- [44] H. Stark and J. W. Woods, *Probability and Random Processes with Applications to Signal Processing*. Upper Saddle River, NJ 07458: Prentice-Hall, 2002.
- [45] T. Söderström and P. Stoica, *System Identification*. London, U.K.: Prentice-Hall International, 1989.
- [46] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, 2001.
- [47] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *SIAM Journal on Scientific Computing*, vol. 52, no. 8, pp. 2153–2164, 2004.
- [48] A. B. Baggeroer and H. Cox, "Passive sonar limits upon nulling multiple moving ships with large aperture arrays," *33th Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 103–108, 1999.

- [49] A. L. Kraay and A. B. Baggeroer, “A physically constrained maximum-likelihood method for snapshot-deficient adaptive array processing,” *IEEE Transactions on Signal Processing*, vol. 55, pp. 4048–4063, 2007.
- [50] W. M. Humphreys, Jr., T. F. Brooks, W. W. Hunter, Jr., and K. R. Meadows, “Design and use of microphone directional arrays for aeroacoustic measurements,” *AIAA Paper 98-0471*, AIAA, 36th Aerospace Sciences Meeting and Exhibit, Reno, NV, January 1998.
- [51] M. Bengtsson, *Antenna Array Signal Processing for High Rank Models*. Ph.D. dissertation, Royal Institute of Technology, Sweden, 1999.
- [52] S. Valaee, B. Champagne, and P. Kabal, “Parametric localization of distributed sources,” *IEEE Transactions on Signal Processing*, vol. 43, pp. 2144–2153, September 1995.
- [53] A. Zoubir, Y. Wang, and P. Charge, “The generalized beamforming techniques for estimating the coherently distributed sources,” *The European Conference on Wireless Technology*, pp. 157–160, November 2005.
- [54] O. Besson and P. Stoica, “Decoupled estimation of DOA and angular spread for a spatially distributed source,” *IEEE Transactions on Signal Processing*, vol. 48, no. 7, pp. 1872–1882, 2000.
- [55] O. Besson and P. Stoica, “Decoupled estimation of DOA and angular spread for spatially distributed sources,” *33th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, vol. 1, pp. 253–257, Oct 24-27 1999.

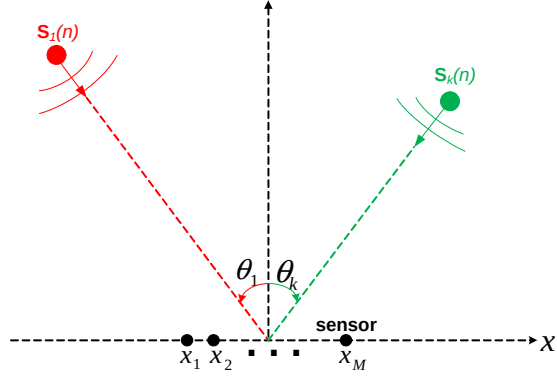


Fig. 1. A linear array.

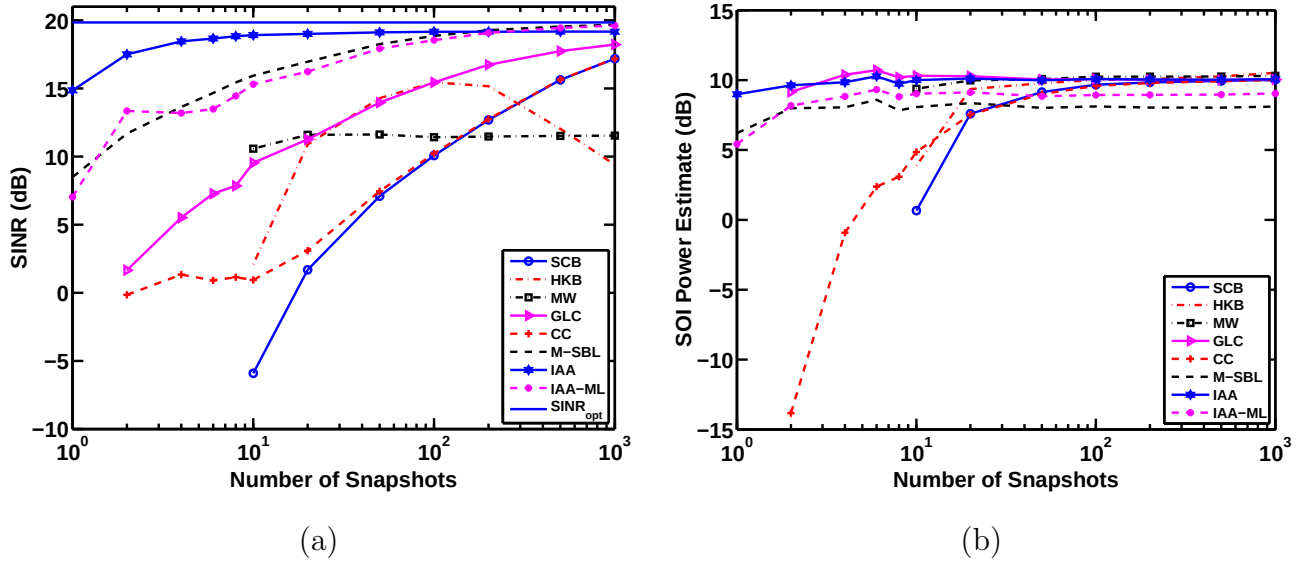
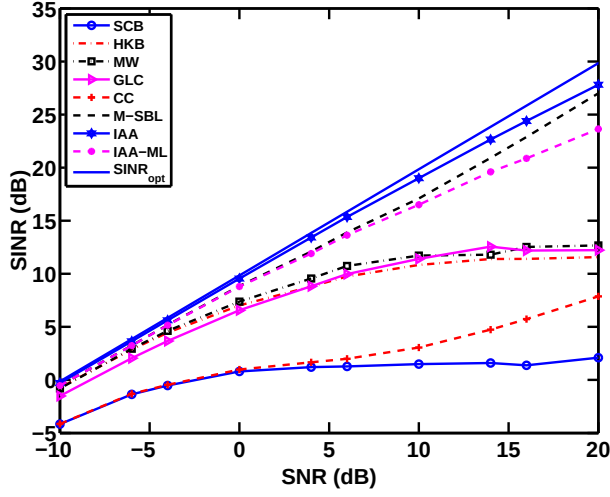
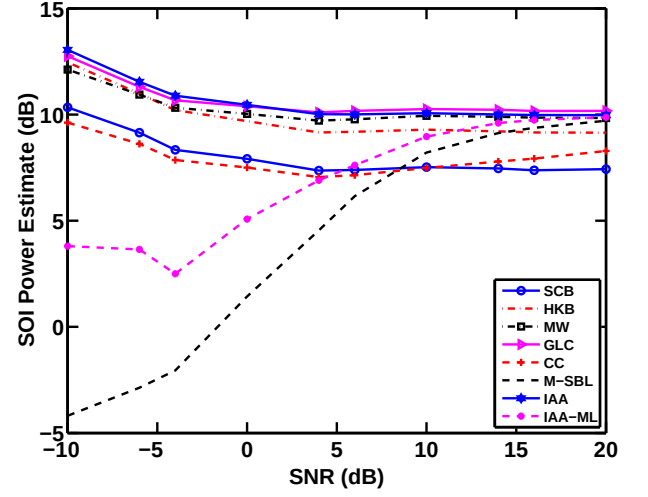


Fig. 2. (a) Beamformer output SINR and (b) SOI power estimates versus the number of snapshots N for $\text{SNR} = 10$ dB. The SOI is at 0° with 10 dB power and the two interferences each with 20 dB power are located at 20° and 60° .

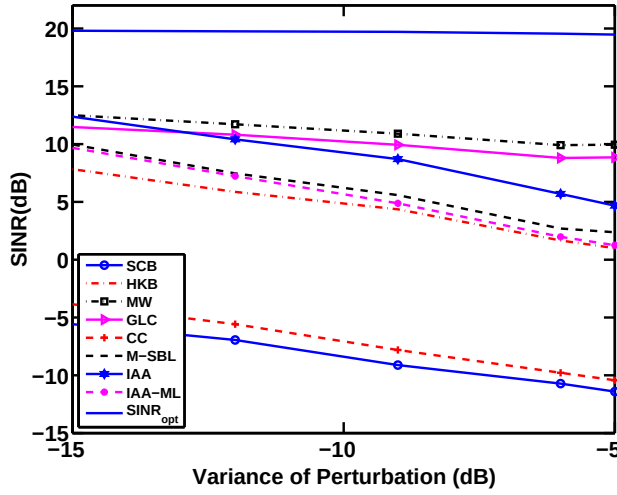


(a)

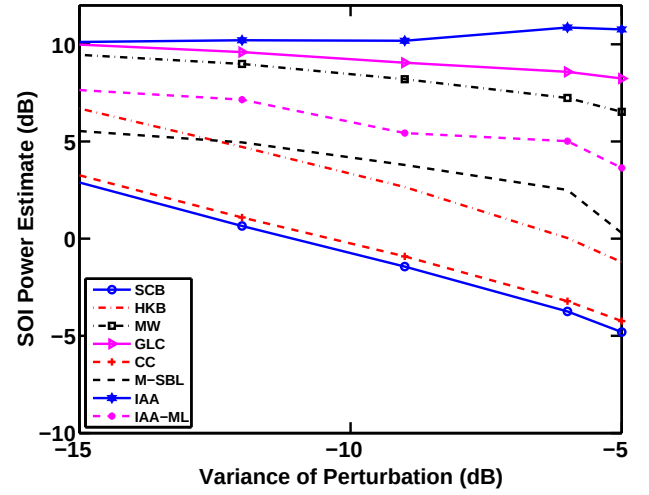


(b)

Fig. 3. (a) Beamformer output SINR and (b) SOI power estimates versus SNR for $N = 20$. The SOI is at 0° with 10 dB power and the two interferences each with 20 dB power are located at 20° and 60° .

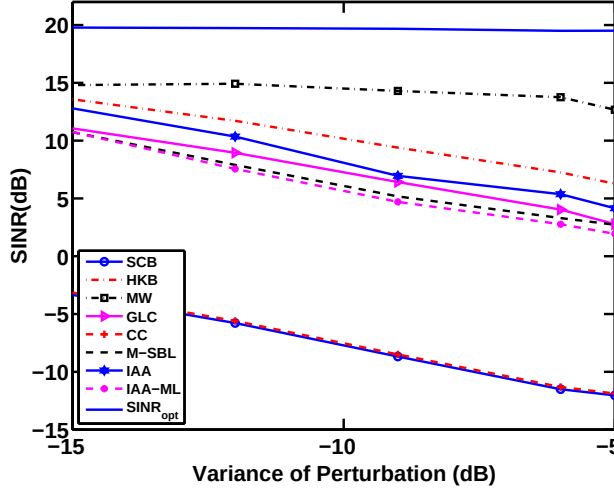


(a)

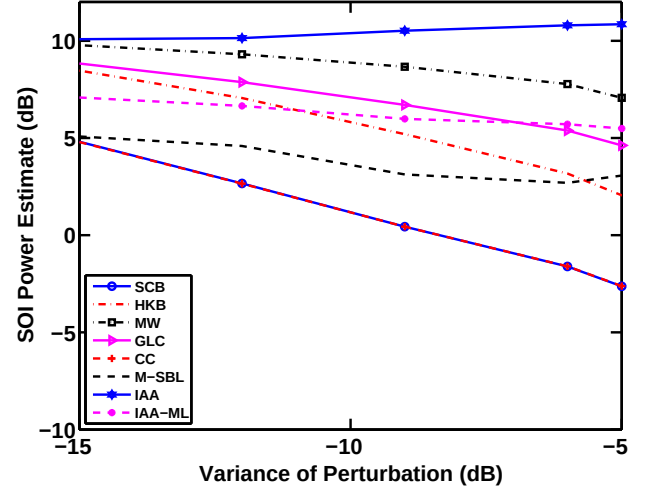


(b)

Fig. 4. (a) Beamformer output SINR and (b) SOI power estimates versus the perturbation variance for $\text{SNR} = 10$ dB and $N = 20$. The SOI is at 0° with 10 dB power and the two interferences each with 20 dB power are located at 20° and 60° .

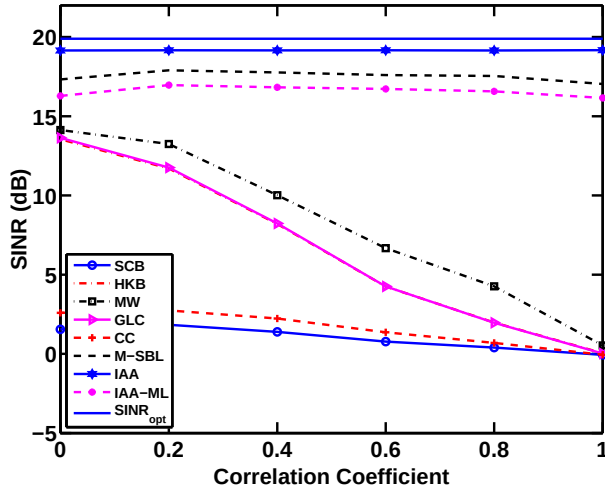


(a)

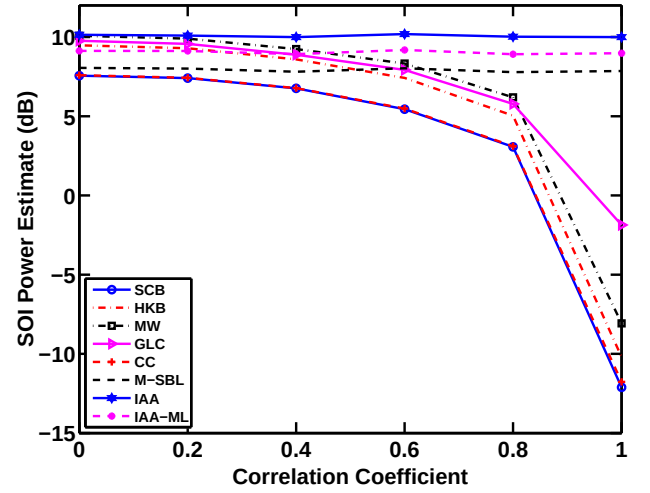


(b)

Fig. 5. (a) Beamformer output SINR and (b) SOI power estimates versus the perturbation variance for SNR = 10 dB and $N = 100$. The SOI is at 0° with 10 dB power and the two interferences each with 20 dB power are located at 20° and 60° .



(a)



(b)

Fig. 6. (a) Beamformer output SINR and (b) SOI power estimates versus the correlation coefficient between the SOI at 0° with 10 dB power and an interference at 20° with 20 dB power. SNR = 10 dB and $N = 20$.

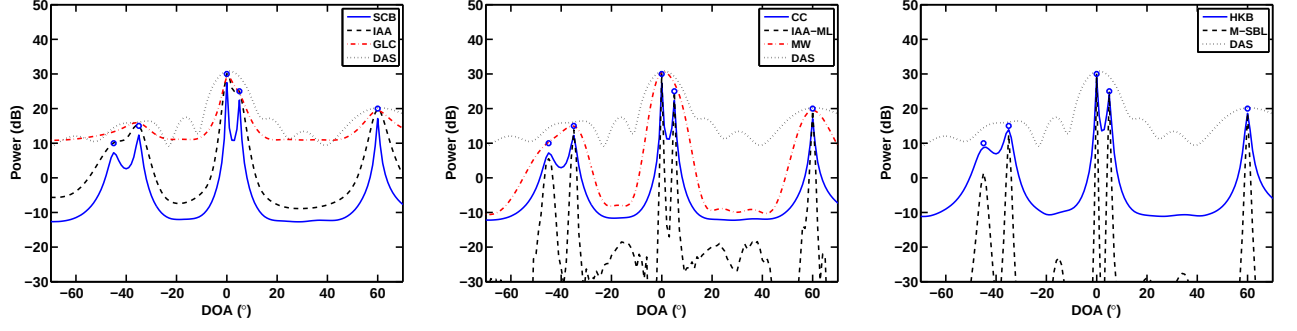


Fig. 7. Spatial power estimates for uncorrelated sources. Circles represents the true source locations and powers. The sources are located at $\theta_1 = -45^\circ$, $\theta_2 = -35^\circ$, $\theta_3 = 0^\circ$, $\theta_4 = 5^\circ$ and $\theta_5 = 60^\circ$ with $\text{SNR}_1=10$ dB, $\text{SNR}_2=15$ dB, $\text{SNR}_3=30$ dB, $\text{SNR}_4=25$ dB and $\text{SNR}_5=20$ dB.

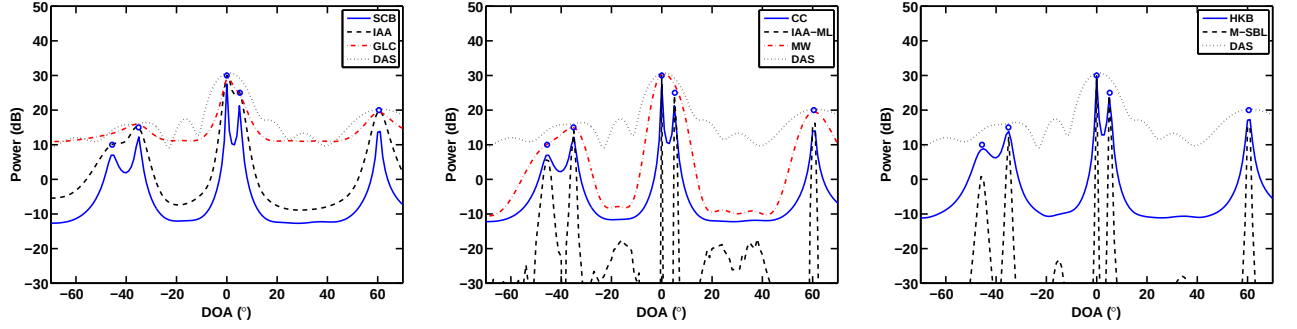


Fig. 8. Spatial power estimates for uncorrelated sources when the sources are not on the scanning grid. Circles represents the true source locations and powers. The sources are located at $\theta_1 = -45.6^\circ$, $\theta_2 = -35.1^\circ$, $\theta_3 = 0^\circ$, $\theta_4 = 5.2^\circ$ and $\theta_5 = 60.5^\circ$ with $\text{SNR}_1=10$ dB, $\text{SNR}_2=15$ dB, $\text{SNR}_3=30$ dB, $\text{SNR}_4=25$ dB and $\text{SNR}_5=20$ dB.

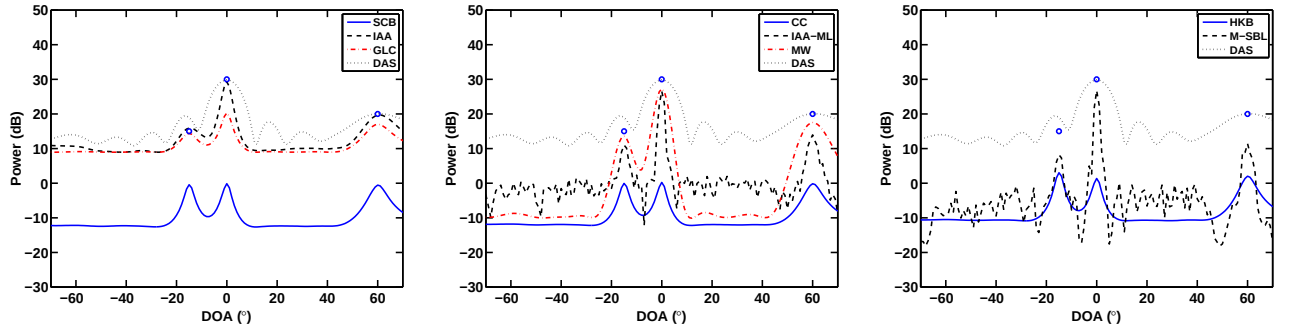


Fig. 9. Spatial power estimates for uncorrelated sources in the presence of array calibration errors. Circles represents the true source locations and powers. The sources are located at $\theta_1 = -15^\circ$, $\theta_2 = 0^\circ$ and $\theta_3 = 60^\circ$ with $\text{SNR}_1=15$ dB, $\text{SNR}_2=30$ dB and $\text{SNR}_3=20$ dB.

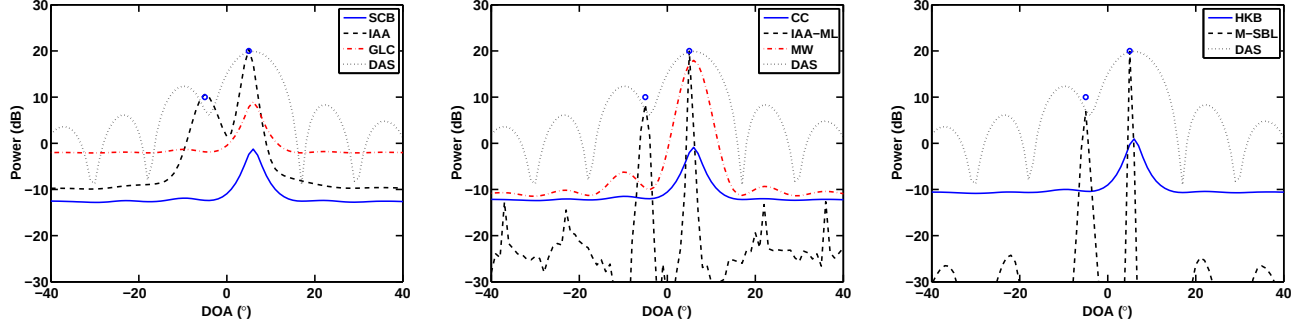


Fig. 10. Spatial power estimates for two coherent sources. Circles represents the true source locations and powers. The sources are located at $\theta_1 = -5^\circ$ and $\theta_2 = 5^\circ$ with $\text{SNR}_1=10$ dB and $\text{SNR}_2=20$ dB.

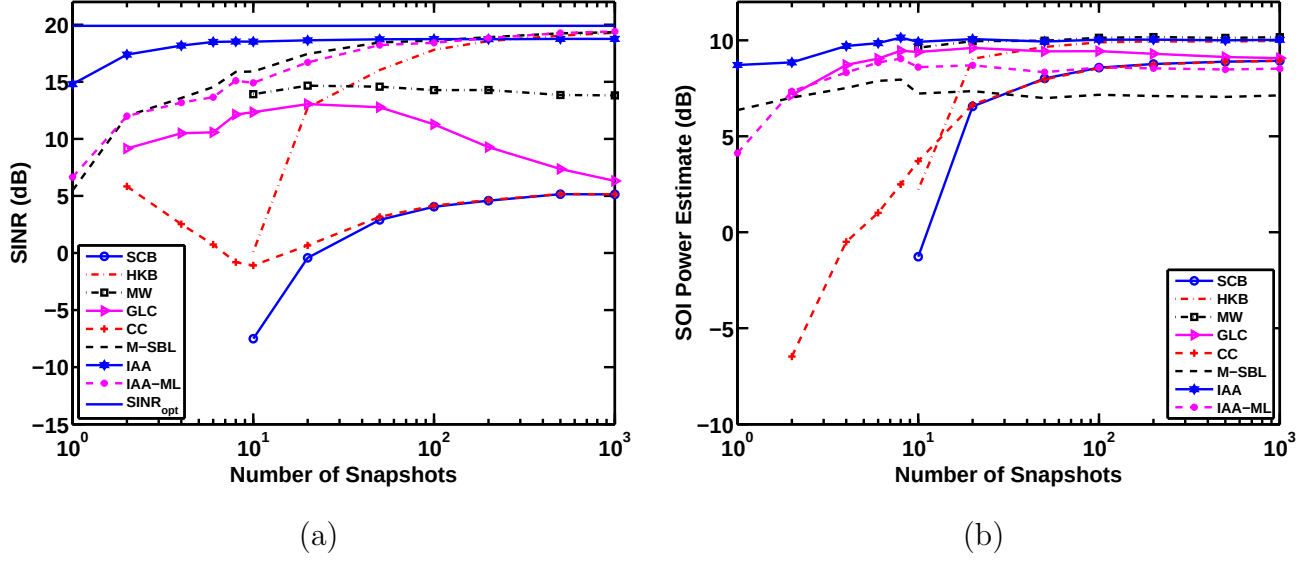


Fig. 11. (a) Beamformer output SINR and (b) SOI power estimates versus the number of snapshots N for uncorrelated spatially distributed sources with $\text{SNR} = 10$ dB. The SOI is at 0° with angular spread 2° and 10 dB power, and the interference is at 20° with angular spread 2° and 20 dB power.

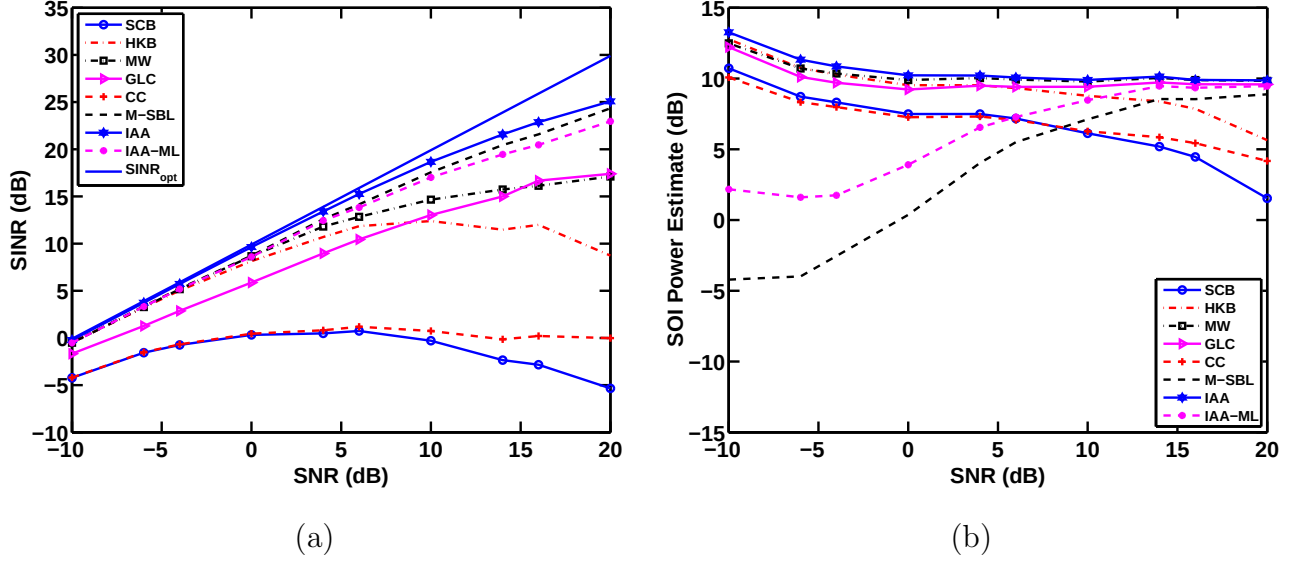


Fig. 12. (a) Beamformer output SINR and (b) SOI power estimates versus SNR for uncorrelated spatially distributed sources with $N = 20$. The SOI is at 0° with angular spread 2° and 10 dB power, and the interference is at 20° with angular spread 2° and 20 dB power.

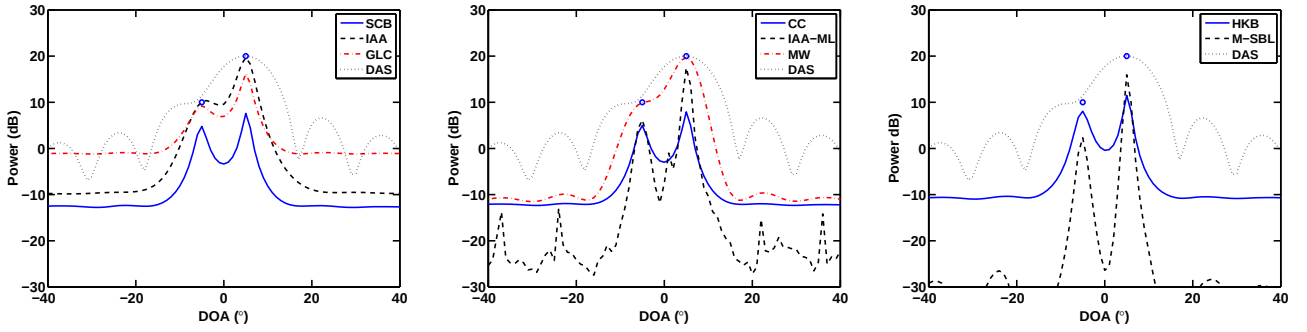


Fig. 13. Spatial power estimates for two uncorrelated spatially distributed sources. Circles represent the true central directions and powers of the sources. The central angles of the sources are $\theta_1 = -5^\circ$ and $\theta_2 = 5^\circ$ with angular spreads $\delta\theta_1 = \delta\theta_2 = 3^\circ$, $\text{SNR}_1 = 10$ dB and $\text{SNR}_2 = 20$ dB.

FULLY AUTOMATIC COMPUTATION OF DIAGONAL LOADING LEVELS FOR ROBUST ADAPTIVE BEAMFORMING

Jian Li, Lin Du*

University of Florida
Dept. of Electrical and Computer Engineering
Gainesville, FL 32611-6130, USA.

Petre Stoica

Uppsala University
Dept. of Information Technology
SE-75105, Uppsala, Sweden.

ABSTRACT

One of the most well-known robust adaptive beamforming approaches is diagonal loading. However, there are usually no clear guidelines on how to choose the diagonal loading level reliably. In this paper, we present algorithms that can compute the diagonal loading level fully automatically from the given data without the need of specifying any user parameters. The proposed diagonal loading algorithms use shrinkage-based covariance matrix estimates, instead of the conventional sample covariance matrix, in the standard Capon beamforming formulation. The performance of the resulting beamformers is illustrated via numerical examples and compared with other adaptive beamforming techniques.

Index Terms— Diagonal loading, Adaptive beamforming

1. INTRODUCTION

The Standard Capon Beamformer (SCB) is an optimal spatial filter that maximizes the array output signal to interference plus noise ratio (SINR), provided that the true covariance matrix and the signal steering vector are accurately known. However, the covariance matrix can be inaccurately estimated due to limited data samples and the knowledge of the steering vector can be imprecise due to look direction errors or imperfect array calibration. Whenever these factors exist, there is a clear performance degradation for SCB. Therefore, adaptive beamforming approaches robust to small sample size problems and steering vector errors are needed.

One of the most well-known robust adaptive beamforming approaches is diagonal loading [1]. The main drawback of this method is that there is no clear way to choose the diagonal loading level reliably. Several recent robust adaptive beamformers have been proposed [2], which can be regarded as diagonal loading approaches, with the diagonal loading level calculated based on the uncertainty set of the array steering vector. However, we still need to specify the param-

eter related to the size of the uncertainty set. Indeed, fully parameter-free robust adaptive beamformers are scarce. One example is the HKB-based SCB [3], which is also a diagonal loading algorithm. However, it may have an inherent problem in choosing an appropriate diagonal loading level.

We provide alternative approaches for the fully automatic computation of the diagonal loading level. We replace the conventional sample covariance matrix used in SCB by an enhanced estimate based on a shrinkage method [4]. Numerical examples are presented to compare the performance of the proposed beamformers with that of HKB and SCB in terms of output SINR and signal-of-interest (SOI) power estimation.

Notation: The superscript $(\cdot)^*$ denotes the conjugate transpose, $(\cdot)^T$ denotes the transpose, $E(\cdot)$ is the expectation operator, $\text{tr}(\cdot)$ is the trace operator, and $\|\cdot\|$ denotes the Frobenius norm for a matrix or the Euclidean norm for a vector.

2. PROBLEM FORMULATION

Consider an array comprising M sensors and let $\mathbf{R}$ denote the theoretical covariance matrix of the array output vector. We assume that $\mathbf{R} > 0$ (positive definite) has the following form:

$$\mathbf{R} = \sigma_0^2 \mathbf{a}_0 \mathbf{a}_0^* + \mathbf{Q}, \quad (1)$$

where σ_0^2 denotes the power of the SOI, $\mathbf{a}_0$ is the array steering vector of the SOI with $\|\mathbf{a}_0\|^2 = M$, and $\mathbf{Q}$ is the interference-plus-noise covariance matrix.

Under ideal conditions, i.e., $\mathbf{a}_0$ and $\mathbf{R}$ are accurately known, the SCB maximizes the output SINR and the optimal value is $\text{SINR}_{\text{opt}} = \sigma_0^2 \mathbf{a}_0^* \mathbf{Q}^{-1} \mathbf{a}_0$. In practice, the exact covariance matrix $\mathbf{R}$ is unavailable. Therefore, $\mathbf{R}$ is replaced by the sample covariance matrix $\hat{\mathbf{R}} = \frac{1}{N} \sum_{n=1}^N \mathbf{y}(n) \mathbf{y}^*(n)$, with N denoting the number of snapshots and $\mathbf{y}(n)$ representing the n th snapshot. As N increases, $\hat{\mathbf{R}}$ converges to $\mathbf{R}$, and the value of the corresponding SINR will approach SINR_{opt} eventually. However, when $\hat{\mathbf{R}}$ contains samples from SOI (e.g., in mobile communications applications), the convergence rate of SCB can be very slow ($N \gg M$ is required). Consequently, the performance of SCB degrades substantially in the presence of small sample size problems, even when $\mathbf{a}_0$ is exactly known.

*The work was supported in part by the Office of Naval Research under Grant No. N00014-07-1-0193, by the National Science Foundation under Grant No. CCF-0634786, ECCS-0729727 and by the Swedish Research Council (VR).

Moreover, the mismatch between the true and assumed steering vectors ($\mathbf{a}_0$ and $\mathbf{a}$) can also significantly deteriorate the performance of SCB.

To improve the performance of SCB, we replace $\hat{\mathbf{R}}$ by an enhanced covariance matrix estimate based on a shrinkage-based method [4]. The enhanced estimate is obtained by linearly combining $\hat{\mathbf{R}}$ and a shrinkage target (a given matrix with some structure) in an optimal mean-squared error (MSE) sense, which can be done via both analytical and convex optimization approaches as shown in the next section.

3. SHRINKAGE-BASED COVARIANCE MATRIX ESTIMATION

A linear shrinkage estimate, which we refer to as the Convex Combination (CC), has the form:

$$\tilde{\mathbf{R}} = \alpha \mathbf{I} + (1 - \alpha) \hat{\mathbf{R}}, \quad (2)$$

where α is the shrinkage intensity, $\tilde{\mathbf{R}}$ is an enhanced estimate of $\mathbf{R}$ and we use the most commonly employed shrinkage target - the identity matrix $\mathbf{I}$. We also consider a more general linear combination (GLC):

$$\tilde{\mathbf{R}} = \alpha \mathbf{I} + \beta \hat{\mathbf{R}}. \quad (3)$$

The shrinkage parameters for both CC and GLC can be chosen by minimizing (an estimate of) the MSE of the estimator $\tilde{\mathbf{R}}$ [4], where $\text{MSE}(\tilde{\mathbf{R}}) = E\{\|\tilde{\mathbf{R}} - \mathbf{R}\|^2\}$.

Note that the constraints $\alpha \in [0, 1]$ for CC and $\alpha \geq 0$, $\beta \geq 0$ for GLC can be imposed to guarantee that $\tilde{\mathbf{R}} \geq 0$. Alternatively, we can impose $\tilde{\mathbf{R}} \geq 0$ directly for both CC and GLC. In the rest of the section, we first review the approaches in [5], where the constraints in the former case are used, and then we extend the approaches further by formulating the MSE minimization problems as convex optimization problems, where all the aforementioned constraints can be imposed.

3.1. Review of the Approaches in [5]

We consider the MSE minimization problem for GLC first.

$$\begin{aligned} \text{MSE}(\tilde{\mathbf{R}}) &= \|\alpha \mathbf{I} - (1 - \beta) \mathbf{R}\|^2 + \beta^2 E\{\|\hat{\mathbf{R}} - \mathbf{R}\|^2\} \\ &= \alpha^2 M - 2\alpha(1 - \beta) \text{tr}(\mathbf{R}) \\ &\quad + (1 - \beta)^2 \|\mathbf{R}\|^2 + \beta^2 E\{\|\hat{\mathbf{R}} - \mathbf{R}\|^2\}. \end{aligned} \quad (4)$$

The optimal values for β and α can be readily obtained:

$$\beta_0 = \frac{\gamma}{\rho + \gamma}, \quad (5)$$

$$\alpha_0 = \nu(1 - \beta_0) = \nu \frac{\rho}{\gamma + \rho}, \quad (6)$$

where $\rho = E\{\|\hat{\mathbf{R}} - \mathbf{R}\|^2\}$, $\nu = \frac{\text{tr}(\mathbf{R})}{M}$, and $\gamma = \|\nu \mathbf{I} - \mathbf{R}\|^2$. We note that $\beta_0 \in [0, 1]$ and $\alpha_0 \geq 0$.

To estimate α_0 and β_0 from the given data, we need an estimate of ρ , which can be calculated as (see [5] for details):

$$\hat{\rho} = \frac{1}{N^2} \sum_{n=1}^N \|\mathbf{y}(n)\|^4 - \frac{1}{N} \|\hat{\mathbf{R}}\|^2. \quad (7)$$

Consequently, we can get estimates for α_0 and β_0 :

$$\hat{\beta}_0^{(1)} = \frac{\hat{\gamma}}{\hat{\gamma} + \hat{\rho}}, \quad (8)$$

and

$$\hat{\alpha}_0^{(1)} = \hat{\nu}(1 - \hat{\beta}_0^{(1)}), \quad (9)$$

where $\hat{\nu} = \frac{\text{tr}(\hat{\mathbf{R}})}{M}$, and $\hat{\gamma} = \|\hat{\nu} \mathbf{I} - \hat{\mathbf{R}}\|^2$. Note that $\hat{\alpha}_0^{(1)}$ and $\hat{\beta}_0^{(1)}$ satisfy the constraints $\alpha \geq 0$ and $\beta \geq 0$. In addition, note that $\gamma + \rho = E\{\|\hat{\mathbf{R}} - \nu \mathbf{I}\|^2\}$, an estimate of which is given by $\|\hat{\mathbf{R}} - \hat{\nu} \mathbf{I}\|^2$. Then we can get alternative estimates of α_0 and β_0 (we need to guarantee that they are nonnegative):

$$\hat{\alpha}_0^{(2)} = \min \left[\hat{\nu} \frac{\hat{\rho}}{\|\hat{\mathbf{R}} - \hat{\nu} \mathbf{I}\|^2}, \hat{\nu} \right], \quad (10)$$

$$\hat{\beta}_0^{(2)} = 1 - \frac{\hat{\alpha}_0^{(2)}}{\hat{\nu}}. \quad (11)$$

We will refer to the GLC using (8)-(9) as GLC₁, and to the GLC using (10)-(11) as GLC₂.

Note that α_0 and β_0 can be rewritten as $\alpha_0 = \nu \tau_0$ and $\beta_0 = 1 - \tau_0$ ($\tau_0 = \frac{\rho}{\rho + \gamma}$), which implies that GLC reduces to CC when $\nu = 1$. Therefore, setting $\hat{\nu} = 1$, we can obtain $\hat{\alpha}_0^{(1)}$ from (9) and $\hat{\alpha}_0^{(2)}$ from (10) for CC, which we refer to as CC₁ and CC₂, respectively.

3.2. Extensions

Below the MSE minimization problems for GLC and CC are formulated as convex optimization problems.

Consider first the convex formulation of GLC. From (4),

$$\begin{aligned} \text{MSE}(\tilde{\mathbf{R}}) &= \alpha^2 M + 2\alpha\beta \text{tr}(\mathbf{R}) + \beta^2 \|\mathbf{R}\|^2 + \beta^2 \rho \\ &\quad - 2\alpha \text{tr}(\mathbf{R}) - 2\beta \|\mathbf{R}\|^2 + \text{const} \\ &= \boldsymbol{\theta}^T \mathbf{A} \boldsymbol{\theta} - 2\mathbf{b}^T \boldsymbol{\theta} + \text{const}. \end{aligned} \quad (12)$$

where $\boldsymbol{\theta} = [\alpha \quad \beta]^T$, $\mathbf{b} = [\text{tr}(\mathbf{R}) \quad \|\mathbf{R}\|^2]^T$, and

$$\mathbf{A} = \begin{bmatrix} M & \text{tr}(\mathbf{R}) \\ \text{tr}(\mathbf{R}) & \|\mathbf{R}\|^2 + \rho \end{bmatrix}. \quad (13)$$

Note that $\mathbf{A} > 0$ and hence, (12) has a unique (unconstrained) minimum solution given by:

$$\boldsymbol{\theta}_0 = [\alpha_0 \quad \beta_0]^T = \mathbf{A}^{-1} \mathbf{b}, \quad (14)$$

which is equivalent to the optimal solution in (5) and (6). Next, we rewrite (12) as:

$$[\boldsymbol{\theta} - \mathbf{A}^{-1} \mathbf{b}]^T \mathbf{A} [\boldsymbol{\theta} - \mathbf{A}^{-1} \mathbf{b}] + \text{const}. \quad (15)$$

Then, the MSE minimization problem for GLC under the constraint $\tilde{\mathbf{R}} \geq 0$ can be formulated as the following Semidefinite Program (SDP):

$$\begin{aligned} & \min_{\delta, \boldsymbol{\theta}} \quad \delta \\ & \text{subject to} \quad \begin{bmatrix} \delta & [\boldsymbol{\theta} - \hat{\mathbf{A}}^{-1}\hat{\mathbf{b}}]^T \\ [\boldsymbol{\theta} - \hat{\mathbf{A}}^{-1}\hat{\mathbf{b}}] & \hat{\mathbf{A}}^{-1} \end{bmatrix} \geq 0 \\ & \quad \tilde{\mathbf{R}}(\boldsymbol{\theta}) \geq 0. \end{aligned} \quad (16)$$

which can be solved in polynomial time using public domain software [6]. In addition, it is easy to obtain a convex optimization formulation for CC by adding the constraint:

$$\mathbf{u}^T \boldsymbol{\theta} = 1, \quad \mathbf{u} = [1 \quad 1]^T \quad (17)$$

to (16). The so-obtained problem is still a SDP.

Note that $\mathbf{A}$ and $\mathbf{b}$ are replaced by their estimates $\hat{\mathbf{A}}$ and $\hat{\mathbf{b}}$ in (16). One way to obtain $\hat{\mathbf{A}}$ and $\hat{\mathbf{b}}$ is to use $\hat{\rho}$ (7) and $\hat{\mathbf{R}}$, respectively, in lieu of ρ and $\mathbf{R}$ in $\mathbf{A}$ and $\mathbf{b}$. Then, we can obtain estimates $\hat{\alpha}_0^{(1')}$ and $\hat{\beta}_0^{(1')}$ of α_0 and β_0 by solving (16). We refer to this method as GLC_{1'}. Similarly, we can also obtain $\hat{\alpha}_0^{(1')}$ by adding (17) to (16) for CC, which we refer to as CC_{1'}. Note that GLC_{1'} and CC_{1'} can be readily shown to be equivalent to GLC₁ and CC₁, respectively. Indeed, the constraint $\tilde{\mathbf{R}} \geq 0$ is inactive due to the GLC_{1'} solution satisfying $\hat{\alpha}_0^{(1')} \geq 0$ and $\hat{\beta}_0^{(1')} \geq 0$, and to the CC_{1'} solution satisfying $\hat{\alpha}_0^{(1')} \in [0, 1]$, which guarantees that $\tilde{\mathbf{R}} \geq 0$.

Exactly as in CC₂ and GLC₂, we can also use alternative estimates of the unknown quantities in $\mathbf{A}$ and $\mathbf{b}$. Noting that $\rho + \|\mathbf{R}\|^2 = E\{\|\hat{\mathbf{R}}\|^2\}$, so we can estimate $\rho + \|\mathbf{R}\|^2$ in $\mathbf{A}$ by $\|\hat{\mathbf{R}}\|^2$, and estimate $\|\mathbf{R}\|^2$ in $\mathbf{b}$ by $\|\hat{\mathbf{R}}\|^2 - \hat{\rho}$. We also replace $\mathbf{R}$ by $\hat{\mathbf{R}}$ in $\text{tr}(\mathbf{R})$. Consequently, we can obtain estimates $\hat{\alpha}_0^{(3)}$ and $\hat{\beta}_0^{(3)}$ from (16) for GLC, which we refer to as GLC₃, and an estimate $\hat{\alpha}_0^{(3)}$ from (17) and (16) for CC, which we refer to as CC₃. GLC₃ and CC₃ are in general different from GLC₂ and CC₂, respectively, due to GLC₃ and CC₃ enforcing $\tilde{\mathbf{R}} \geq 0$ directly while minimizing (15) (with $\mathbf{A}$ and $\mathbf{b}$ replaced by $\hat{\mathbf{A}}$ and $\hat{\mathbf{b}}$). GLC₂ and CC₂, on the other hand, minimize (15) (with $\mathbf{A}$ and $\mathbf{b}$ replaced by the same $\hat{\mathbf{A}}$ and $\hat{\mathbf{b}}$) without imposing any constraints, and then clip the solutions to satisfy $\hat{\alpha}_0^{(2)} \geq 0$ and $\hat{\beta}_0^{(2)} \geq 0$ for GLC and $\hat{\alpha}_0^{(2)} \in [0, 1]$ for CC. Therefore, GLC₂ and CC₂ are suboptimal. The optimal version of GLC₂, which we refer to as GLC₄, can be obtained by using the constraints $\alpha \geq 0$ and $\beta \geq 0$ instead of $\tilde{\mathbf{R}}(\boldsymbol{\theta}) \geq 0$ in (16) and calculating $\hat{\mathbf{A}}$ and $\hat{\mathbf{b}}$ in the same way as in GLC₃. We can similarly get CC₄, which is the optimal version of CC₂.

4. SHRINKAGE-BASED ROBUST CAPON BEAMFORMERS

We have 8 methods to obtain the enhanced estimates of the covariance matrix, i.e.,

$$\tilde{\mathbf{R}}_{\text{GLC}_i} = \hat{\alpha}_0^{(i)} \mathbf{I} + \hat{\beta}_0^{(i)} \hat{\mathbf{R}}, \quad i = 1, \dots, 4, \quad (18)$$

and

$$\tilde{\mathbf{R}}_{\text{CC}_i} = \hat{\alpha}_0^{(i)} \mathbf{I} + (1 - \hat{\alpha}_0^{(i)}) \hat{\mathbf{R}}, \quad i = 1, \dots, 4. \quad (19)$$

Using one of the above enhanced estimates $\tilde{\mathbf{R}}$ in lieu of $\hat{\mathbf{R}}$ in the SCB formulation yields the shrinkage-based robust adaptive beamformer: $\tilde{\mathbf{w}} = \frac{\tilde{\mathbf{R}}^{-1} \mathbf{a}}{\mathbf{a}^* \tilde{\mathbf{R}}^{-1} \mathbf{a}}$. The resulting beamformer output SINR is given by $\text{SINR} = \frac{\sigma_0^2 |\tilde{\mathbf{w}}^* \mathbf{a}_0|^2}{\tilde{\mathbf{w}}^* \mathbf{Q} \tilde{\mathbf{w}}}$, and the SOI power estimate is $\hat{\sigma}_0^2 = \tilde{\mathbf{w}}^* \tilde{\mathbf{R}} \tilde{\mathbf{w}}$.

From (18)-(19), we note that the shrinkage-based robust adaptive beamformers are diagonal loading approaches with the diagonal loading levels ($\hat{\alpha}_0/\hat{\beta}_0$ for GLC and $\hat{\alpha}_0/(1 - \hat{\alpha}_0)$ for CC) determined automatically from the observed data snapshots $\{\mathbf{y}(n)\}_{n=1}^N$.

5. NUMERICAL EXAMPLES

We present below several numerical examples comparing the performance of the shrinkage-based robust adaptive beamformers with that of HKB and SCB. Interestingly, in all of these examples, the GLC₂ solutions did not need clipping, and hence GLC₂, GLC₃, and GLC₄ were equivalent, in addition, GLC₁ performed similarly to GLC₂. The same was true for CC. Therefore, only the results obtained by GLC₂ and CC₂ will be presented, and we will refer to GLC₂ as GLC and CC₂ as CC for short. In all examples, we assume a uniform linear array with $M = 10$ sensors and half-wavelength inter-element spacing. The noise is assumed to be white complex Gaussian random process with zero-mean and covariance matrix $\mathbf{I}$. A SOI with a 10 dB power is assumed to impinge on the array from 0° , and two interferences, each with a 20 dB power, are assumed to be present at 10° and 60° . For each scenario, 1000 Monte-Carlo trials are performed.

First, we examine the output SINR convergence performance of the beamformers. Fig. 1 shows the mean of the output SINRs versus the number of snapshots N when $\mathbf{a}_0$ is known. As shown in the figure, SCB converges to SINR_{opt} very slowly. Both GLC and CC outperform SCB. Whereas, unlike CC, GLC provides a significant improvement over SCB for all values of N considered. Figs. 2(a) and (b) show the mean values of the diagonal loading levels of GLC and CC, respectively, as a function of N . We observe that the diagonal loading level of CC is much lower than that of GLC. Another observation from Fig. 1 is that the output SINR of HKB decreases when N is beyond a certain number. Unlike GLC and CC, as shown in Fig. 2(c), the mean of the diagonal loading level for HKB starts from a very small value and

monotonically and quickly increases with N . This behavior limits HKB's performance improvement over SCB when N is small and deteriorates its performance when N is large.

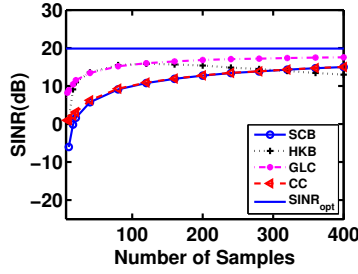


Fig. 1. Beamformer output SINR versus N .

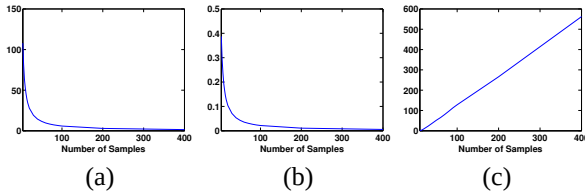


Fig. 2. Average diagonal loading levels versus N : (a) GLC (b) CC (c) HKB.

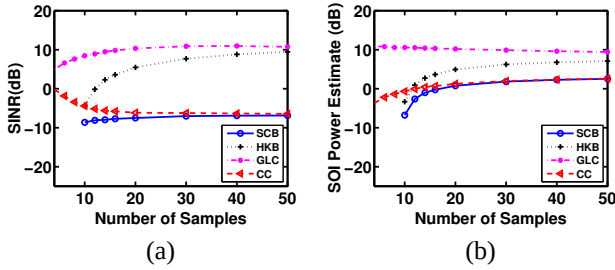


Fig. 3. Performance comparison in the presence of a 2° steering angle error: (a) SINR, (b) SOI power estimates versus N .

Next, we examine the robustness of the beamformers to small sample size problems and to steering vector errors. Figs. 3 and 4, respectively, show the performance in the presence of look direction errors (2° SOI steering angle mismatch) and array calibration errors. The array calibration errors are simulated by perturbing each element of the array steering vector using independent zero-mean complex Gaussian random variables with variance 0.01. GLC shows the best performance, especially when N is small. Moreover, HKB and SCB are not applicable when $\hat{\mathbf{R}}$ is rank deficient ($N < M$), whereas, GLC and CC can be used. Note that when $N \gg M$, $\hat{\mathbf{R}}$ becomes very close to $\mathbf{R}$, and hence the shrinkage-based approaches will choose small diagonal loading levels. Then their ability to combat steering vector errors will diminish. Yet the shrinkage based methods are very useful for the case of small sample sizes. This case is often encountered in practice and is most critically in need of performance improvement.

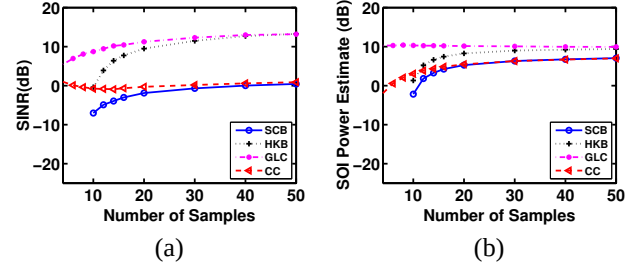


Fig. 4. Performance comparison in the presence of array calibration errors: (a) SINR, (b) SOI power estimates versus N .

6. CONCLUSIONS

We have presented several approaches to the fully automatic computation of diagonal loading levels. In our diagonal loading algorithms, the conventional sample covariance matrix used in the SCB formulation is replaced by an enhanced covariance matrix estimate based on shrinkage. We have shown how to efficiently obtain the shrinkage covariance matrix estimates from the available data. Several numerical examples have been used to compare the performance of the proposed beamformers with that of SCB and HKB. The shrinkage-based approaches improve the robustness of SCB against small sample size problems and steering vector errors, with GLC having the best performance among the methods tested. More importantly, we have demonstrated that GLC is very useful in the case of small sample sizes - the case in which the users of adaptive arrays are most interested.

7. REFERENCES

- [1] H. Cox, R. Zeskind, and M. Owen, "Robust adaptive beamforming," *Proceedings of IEEE*, vol. ASSP-35, pp. 1365–1375, October 1987.
- [2] J. Li and P. Stoica, eds., *Robust Adaptive Beamforming*. New York, NY: John Wiley & Sons, 2005.
- [3] Y. Selen, R. Abrahamsson, and P. Stoica, "Automatic robust adaptive beamforming via ridge regression," to appear in *Signal Processing*, 2008.
- [4] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of Multivariate Analysis*, vol. 88, pp. 365–411, 2004.
- [5] P. Stoica, J. Li, X. Zhu, and J. R. Guerci, "On using a priori knowledge in space-time adaptive processing," to appear in *IEEE Transactions on Signal Processing*.
- [6] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," Available: <http://www2.unimaas.nl/~sturm/software/sedumi.html>.