MAXIMUM LIKELIHOOD ESTIMATION OF POINT SCATTERERS FOR COMPUTATIONAL TIME-REVERSAL IMAGING*

GANG SHI[†] AND ARYE NEHORAI[†]

Abstract. We present a statistical framework for the fixed-frequency computational time-reversal imaging problem assuming point scatterers in a known background medium. Our statistical measurement models are based on the physical models of the multistatic response matrix, the distorted wave Born approximation and Foldy-Lax multiple scattering models. We develop maximum likelihood (ML) estimators of the locations and reflection parameters of the scatterers. Using a simplified single-scatterer model, we also propose a likelihood time-reversal imaging technique which is suboptimal but computationally efficient and can be used to initialize the ML estimation. We generalize the fixed-frequency likelihood imaging to multiple frequencies, and demonstrate its effectiveness in resolving the grating lobes of a sparse array. This enables to achieve high resolution by deploying a large-aperture array consisting of a small number of antennas while avoiding spatial ambiguity. Numerical and experimental examples are used to illustrate the applicability of our results.

1. Introduction. The time reversal approach, which traces its origin to "phase conjugation" in nonlinear optics [1], was later developed using acoustic experiments [2] and now attracts increasing interests with broad applications. The key idea behind the so-called *physical time-reversal* methods is to record a signal emitted by sources or reflected by targets using an array of transducers; then transmit the time-reversed and complex conjugated version of the measurements back into the medium. In a reciprocal medium, the back-propagated wave will then retrace the original trajectory and focus around the original source locations without the need to solve the inverse of the channel. In a homogeneous medium, the diffraction resolution of the refocused field in the direction parallel to the array (or cross-range resolution) is $\lambda R/a$ [3], [4], where λ is the carrier wavelength, R is the range between the array and the source, and a is the array aperture. The resolution in the perpendicular direction (or range resolution) is $\lambda(R/a)^2$ [3], [4]. For physical time-reversal methods in a non-homogeneous medium, experimental and theoretical evidence show that the refocusing is much tighter and the Fresnel zones are reduced [5]-[19]. This superior focusing resolution quality is interpreted intuitively due to taking advantage of the inhomogeneity which distributes the wave over a larger part of the medium and therefore carries more information about the source location [6]. The enhanced resolution is called *super-resolution*, making the time-reversal methods attractive.

Field measurements conducted by a group of the University of California at San Diego demonstrated that a time-reversed acoustic signal can be focused up to 30 kilometers away in water about 120 meters deep [2], [5]. Recent experiments [20]-[22]

†ECE Department, The University of Illinois at Chicago. E-mail: gshi1@uic.edu

confirm that the spatial and temporal focusing property holds also for electromagnetic waves. This refocusing property is currently being explored and has a great potential for numerous applications, including radar, underwater acoustics, biomedicine, detection of defects in metals, communications, etc. In [23], [24] Cramér-Rao bounds are computed to evaluate the performance of physical time reversal for estimating the reflectivity of an unknown environment.

One application of this time-reversal refocusing property is to detect and locate a target by *computational* or *virtual time reversal* through imaging [3], [6]. In this case, after receiving the signal reflected from the target, a back-propagated process is *computed* rather than implemented in the real medium. The reversed signal then refocus in the computed image around the target location, and the peaks indicate existence of possible targets. Various computational imaging strategies have been proposed, which in general fall under three categories: (a) time domain methods that use mostly times of arrival and amplitude information recorded by the array, (b) fixed-frequency methods that use mostly differential phase information on the array, and (c) intensity measurement based methods (see [3] for more details).

In Section 2, we employ two physical models used in computational time-reversal imaging methods assuming point scatterers: one ignores the multiple scattering among the scatterers using the distorted wave Born approximation [25] and the other incorporates them using the Foldy-Lax models [26], [27]. We then formulate the computational time-reversal imaging in a statistical framework by establishing a measurement model. In Section 3 we develop maximum likelihood estimation (MLE) and suboptimal methods to estimate the parameters of interest, including the locations and scattering coefficients of the scatterers. In Section 4 we propose a likelihood computational time-reversal imaging method using a simplified single-scatterer model, and generalize it into multiple-frequency version in Section 5, which is demonstrated to be useful in resolving the spatial ambiguity of a sparse array. Numerical and experimental examples are presented in Section 6, and conclusions are given in Section 7.

2. Physical and Statistical Measurement Models. We present physical and statistical measurement models of the multistatic response matrix [28] made by an array of antennas, which we will employ in later sections for solving the computational time-reversal imaging in a statistical signal processing framework.

2.1. Multistatic Response Matrix. We introduce the so-called multistatic response matrix [28] by considering transmit and receive antenna arrays of N_t and N_r isotropic point antennas, centered at known positions $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_{N_t}$ and $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots,$ $\boldsymbol{\beta}_{N_r}$, respectively. Suppose a known signal $\boldsymbol{s}(t) = [s_1(t), s_2(t), \ldots, s_{N_t}(t)]^T$ is transmitted to illuminate the scenario of interest, i.e., the j-th antenna transmits $s_j(t)$, $j = 1, \ldots, N_t$, and the resulting backscattered returns are measured by all the receive

antennas. Consider this system in the frequency domain and ignore noise, the received signal is represented as $\boldsymbol{y}(\omega) = K(\omega)\boldsymbol{s}(\omega)$, where $\boldsymbol{y}(\omega) = [y_1(\omega), y_2(\omega), \ldots, y_{N_r}(\omega)]^T$, $\boldsymbol{s}(\omega) = [s_1(\omega), s_2(\omega), \ldots, s_{N_t}(\omega)]^T$ are the Fourier transforms of the received and transmitted signals, respectively. The matrix $K = [K_{j,k}(\omega)]$ is the *multistatic response matrix* [28] of dimension $N_r \times N_t$. The element $K_{j,k}(\omega)$ can be interpreted as the received signal at a frequency ω at the j-th receive antenna due to an impulse excitation (in the frequency domain) applied by the k-th transmit antenna [28].

Let $\boldsymbol{y}(t)$ be the received signal in the time domain, i.e., the inverse Fourier transform of $\boldsymbol{y}(\omega)$, then its time reversed complex conjugated version $\boldsymbol{y}^*(-t)$ in the time domain corresponds to phase conjugation in frequency domain $\boldsymbol{y}^*(\omega)$, where "*" denotes the complex conjugate. Consider a special case with co-located transmit and receive arrays, i.e., $N_t = N_r = N$ and $\boldsymbol{\alpha}_1 = \boldsymbol{\beta}_1, \boldsymbol{\alpha}_2 = \boldsymbol{\beta}_2 \cdots \boldsymbol{\alpha}_N = \boldsymbol{\beta}_N$, after the time reversed complex conjugated $\boldsymbol{y}^*(-t)$ is transmitted back into the medium, the measured backscattered signal on the receive array could be modeled as

$$(1) \qquad \boldsymbol{y}^{(b)}(\omega) = K(\omega)\boldsymbol{y}^*(\omega) = K(\omega)K^*(\omega)s^*(\omega),$$

where the superscript "(b)" denotes the measurement of the back-propagated time-reversed and complex conjugated signal, or equivalently,

$$(2) \qquad \boldsymbol{y}^{(b)*}(\omega) = K^*(\omega)K(\omega)s(\omega).$$

We can see that $K^*(\omega)K(\omega)$ works as a time-reversal operator, defined as *time-reversal matrix* in [29]-[31], and the information about the probed scenario is encoded in the multistatic matrix $K(\omega)$. It is shown in [31] that for well-resolved scatterers, the eigenvectors of the time-reversal matrix that correspond to the nonzero eigenvalues are associated in a one-to-one manner with the scatterers. In the computational time-reversal imaging, the back-propagated field or "pseudo-spectrum" [31] is computed from $K(\omega)$ and the inference on the scenario is made through the created image. It is worth noting that the time-reversal methods are different from inverse filtering where the former exploits the reciprocity of the channel whereas the latter tries to find the inverse of a channel.

In the rest of this paper, the numbers of transmit and receive antennas N_t and N_r are not necessarily equal, we consider a scenario where M discrete point scatterers are embedded in a known background medium (see Figure 1 for an illustration). The locations and (frequency-dependent) scattering coefficients of the scatterers are assumed unknown, which are denoted by $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_M$, and $\tau_1(\omega), \tau_2(\omega), \ldots, \tau_M(\omega)$, respectively. Alternatively to the conventional computational time-reversal imaging [3], we will estimate the locations and scattering coefficients using statistical maximum-likelihood (ML) method, which is known asymptotically unbiased and efficient [32] under mild conditions.
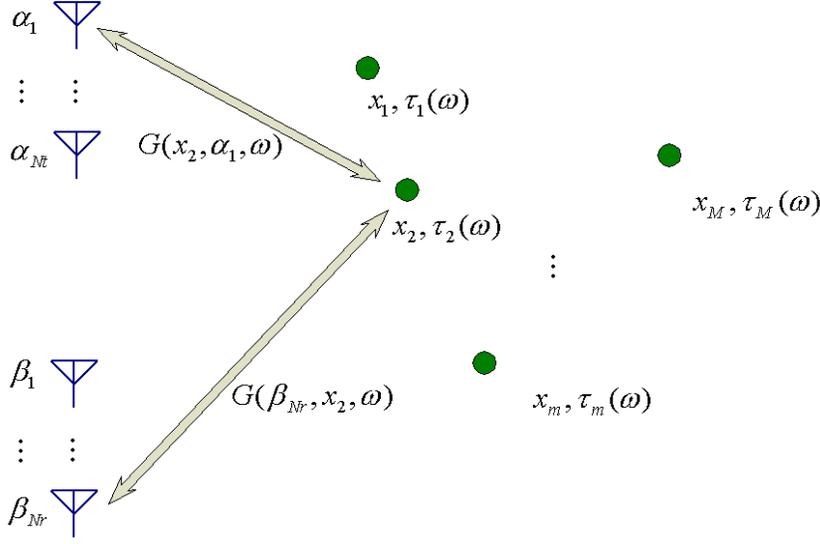
Fig. 1. *scenario illustration.*

2.2. Multistatic Matrix Using the Distorted Wave Born Approximation. In this section we use the distorted wave Born approximation (DWBA) [25], meaning that we neglect the multiple scattering among the scatterers. Then, the scattered field at a position r induced by the k-th transmit antenna is given by [31]

$$(3) \quad \psi_k^{(s)}(\boldsymbol{r}, \omega) = \sum_{m=1}^{M} G(\boldsymbol{r}, \boldsymbol{x}_m, \omega) \tau_m(\omega) G(\boldsymbol{x}_m, \boldsymbol{\alpha}_k, \omega) s_k(\omega), \qquad k = 1, 2, \ldots, N_{\mathrm{t}},$$

where $G(\cdot)$ is the scalar time harmonic background Green function satisfying the reduced wave equation [3]

$$(4) \qquad \nabla^2 G(\boldsymbol{r}, \boldsymbol{x}', \omega) + (\omega/c_0)^2 n^2(\boldsymbol{x}) G(\boldsymbol{r}, \boldsymbol{x}'\omega) = -\delta(\boldsymbol{r} - \boldsymbol{x}'),$$

here c_0 is a reference propagation speed and $n(\boldsymbol{x})$ is the refraction index of the medium at a position $\boldsymbol{x}$. Hence, the knowledge of the background medium is fully incorporated into the Green function, and in this paper we assume that the background Green function is known. For example, in a homogeneous medium the three-dimensional Green function that satisfies the radiation condition is given by [33]

$$(5) \qquad G(\boldsymbol{r}, \boldsymbol{x}', \omega) = \frac{\exp(i2\pi|\boldsymbol{r} - \boldsymbol{x}'|/\lambda)}{4\pi|\boldsymbol{r} - \boldsymbol{x}'|} = \frac{\exp(i\omega|\boldsymbol{r} - \boldsymbol{x}'|/c)}{4\pi|\boldsymbol{r} - \boldsymbol{x}'|},$$

where λ is the wavelength and c is the propagation speed; note that we use $|\cdot|$ to represent both the Euclidean norm of a vector and modulus of a complex number. The Green function (5) represents a spherical wave case, and its variations have been used for general channel modeling in [23], [24], [34].

According to (3), the multistatic matrix $K(\omega)$ is found as

(6)
$$K_{j,k}(\omega) = \sum_{m=1}^{M} G(\boldsymbol{\beta}_j, \boldsymbol{x}_m, \omega)\tau_m(\omega)G(\boldsymbol{x}_m, \boldsymbol{\alpha}_k, \omega), j = 1, 2, \ldots, N_{\mathrm{r}}, k = 1, 2, \ldots, N_{\mathrm{t}}.$$

Define $\boldsymbol{x} = [\boldsymbol{x}_1^T, \boldsymbol{x}_2^T, \ldots, \boldsymbol{x}_M^T]^T$ as the unknown scatterer location parameter vector of dimension $3M$, and $\boldsymbol{\tau}(\omega) = [\tau_1(\omega), \tau_2(\omega), \ldots, \tau_M(\omega)]^T$ unknown scattering coefficients. The above model is formulated in a matrix form as

(7)
$$K(\boldsymbol{x}, \boldsymbol{\tau}(\omega), \omega) = \sum_{m=1}^{M} \tau_m(\omega)\boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}_m, \omega)\boldsymbol{g}_{\mathrm{t}}^T(\boldsymbol{x}_m, \omega),$$

(8)
$$= A_{\mathrm{r}}(\boldsymbol{x}, \omega)T(\boldsymbol{\tau}(\omega))A_{\mathrm{t}}^T(\boldsymbol{x}, \omega),$$

where "^{T}" stands for a matrix transpose, $T(\boldsymbol{\tau}(\omega)) = \mathrm{diag}\{\boldsymbol{\tau}(\omega)\}$,

(9)
$$A_{\mathrm{r}}(\boldsymbol{x}, \omega) = [\boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}_1, \omega), \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}_2, \omega) \cdots \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}_M, \omega)],$$

(10)
$$A_{\mathrm{t}}(\boldsymbol{x}, \omega) = [\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}_1, \omega), \boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}_2, \omega) \cdots \boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}_M, \omega)],$$

and the N_{r}-dimensional receive Green function vector $\boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}', \omega)$ and N_{t}-dimensional transmit Green function vector $\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}', \omega)$ are defined as

(11)
$$\boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}', \omega) = [G(\boldsymbol{\beta}_1, \boldsymbol{x}', \omega), G(\boldsymbol{\beta}_2, \boldsymbol{x}', \omega), \ldots, G(\boldsymbol{\beta}_{N_{\mathrm{r}}}, \boldsymbol{x}', \omega)]^T,$$

(12)
$$\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}', \omega) = [G(\boldsymbol{x}', \boldsymbol{\alpha}_1, \omega), G(\boldsymbol{x}', \boldsymbol{\alpha}_2, \omega), \ldots, G(\boldsymbol{x}', \boldsymbol{\alpha}_{N_{\mathrm{t}}}, \omega)]^T.$$

The receive and transmit Green function vectors incorporate all of the spatial characteristics of the response of the receive and transmit arrays to a point scatterer located at $\boldsymbol{x}'$, hence can be interpreted as the generalization of the conventional *array response vector* or *steering vector* [35]. In the following, we will drop the dependence on frequency ω in all notations for notational simplicity except for Section 5 where we consider multiple frequencies.

2.3. Multistatic Matrix Using the Foldy-Lax Multiple Scattering Model. To incorporate the multiple scattering among the scatterers, the physical model of the multistatic matrix (6) could be generalized using the so-called Foldy-Lax multiple scattering model [26], [27]

(13)
$$K_{j,k} = \sum_{m=1}^{M} G(\boldsymbol{\beta}_j, \boldsymbol{x}_m)\tau_m \mathcal{G}(\boldsymbol{x}_m, \boldsymbol{\alpha}_k), \quad j = 1, 2, \ldots, N_{\mathrm{r}}, \quad k = 1, 2, \ldots, N_{\mathrm{t}},$$

where $G(\cdot)$ is the background Green function and $\mathcal{G}(\cdot)$ represents the full Green function of the background plus scatterers system. In a matrix form, the multistatic matrix

(14)
$$K(\boldsymbol{x}, \boldsymbol{\tau}) = \sum_{m=1}^{M} \tau_m \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}_m)\mathbf{g}_{\mathrm{t}}^T(\boldsymbol{x}_m) = A_{\mathrm{r}}(\boldsymbol{x})T(\boldsymbol{\tau})\mathcal{A}_{\mathrm{t}}^T(\boldsymbol{x}),$$

where $\mathcal{A}_t(\boldsymbol{x}) = [\mathbf{g}_t(\boldsymbol{x}_1), \mathbf{g}_t(\boldsymbol{x}_2) \cdots \mathbf{g}_t(\boldsymbol{x}_M)]$, and $\mathbf{g}_t(\boldsymbol{x}_m) = [\mathcal{G}(\boldsymbol{x}_m, \boldsymbol{\alpha}_1), \mathcal{G}(\boldsymbol{x}_m, \boldsymbol{\alpha}_2), \ldots,$
$\mathcal{G}(\boldsymbol{x}_m, \boldsymbol{\alpha}_{N_t})]^T$, for $m = 1, 2, \ldots, M$.

The Green functions G and $\mathcal{G}$ satisfy the Foldy-Lax equations [26], [27]:

$$(15) \qquad \mathcal{G}(\boldsymbol{x}_m, \boldsymbol{\alpha}_k) = G(\boldsymbol{x}_m, \boldsymbol{\alpha}_k) + \sum_{m' \neq m} G(\boldsymbol{x}_m, \boldsymbol{x}_{m'}) \tau_{m'} \mathcal{G}(\boldsymbol{x}_{m'}, \boldsymbol{\alpha}_k),$$

$$m = 1, 2, \ldots, M, k = 1, 2, \ldots, N_t.$$

We reformulate the above MN_t Foldy-Lax equations into an $M \times N_t$ matrix form as

$$(16) \qquad \mathcal{A}_t^T(\boldsymbol{x}) = A_t^T(\boldsymbol{x}) + S(\boldsymbol{x})T(\boldsymbol{\tau})\mathcal{A}_t^T(\boldsymbol{x}),$$

where

$$S(\boldsymbol{x}) = \begin{pmatrix} 0 & G(\boldsymbol{x}_1, \boldsymbol{x}_2) & \cdots & G(\boldsymbol{x}_1, \boldsymbol{x}_M) \\ G(\boldsymbol{x}_2, \boldsymbol{x}_1) & 0 & \cdots & G(\boldsymbol{x}_2, \boldsymbol{x}_M) \\ \vdots & & \ddots & \vdots \\ G(\boldsymbol{x}_{M-1}, \boldsymbol{x}_1) & \cdots & 0 & G(\boldsymbol{x}_{M-1}, \boldsymbol{x}_M) \\ G(\boldsymbol{x}_M, \boldsymbol{x}_1) & \cdots & G(\boldsymbol{x}_M, \boldsymbol{x}_{M-1}) & 0 \end{pmatrix}.$$

Solving $\mathcal{A}_t(\boldsymbol{x}) = A_t(\boldsymbol{x})[I - T(\boldsymbol{\tau})S(\boldsymbol{x})]^{-1}$ from (16) and substituting it into (14), we derive the closed-form physical model of the multistatic matrix K that includes the multiple scatterings as

$$(17) \qquad K(\boldsymbol{x}, \boldsymbol{\tau}) = A_r(\boldsymbol{x})[T^{-1}(\boldsymbol{\tau}) - S(\boldsymbol{x})]^{-1} A_t^T(\boldsymbol{x}).$$

Observing (17), we could see that when $|\tau_m|$, $m = 1, 2, \ldots, M$ are small or the distances between the scatterers are large enough, $[T^{-1}(\boldsymbol{\tau}) - S(\boldsymbol{x})]^{-1} \approx T(\boldsymbol{\tau})$, then the physical model $K(\boldsymbol{x}, \boldsymbol{\tau}) = A_r(\boldsymbol{x})[T^{-1}(\boldsymbol{\tau}) - S(\boldsymbol{x})]^{-1} A_t^T(\boldsymbol{x}) \approx A_r(\boldsymbol{x})T(\boldsymbol{\tau})A_t^T(\boldsymbol{x})$, namely it becomes the model (8) that uses the Born approximation.

2.4. Statistical Measurement Model. In the fixed-frequency computational time-reversal imaging, the multistatic matrix is evaluated at only one specific frequency. The back-propagated field or "pseudo-spectrum" [31] is computed from K directly assuming it is obtained in advance. Obviously the measurement and modeling of the multistatic matrix will have inaccuracies, hence we assume that it is perturbed by an additive noise. The statistical measurement model of the multistatic matrix is

$$(18) \qquad Y = K(\boldsymbol{x}, \boldsymbol{\tau}) + W,$$

where Y is the $N_r \times N_t$ measurement matrix, W is a noise matrix whose entries $w_{j,k}$ are assumed zero-mean jointly circularly symmetric complex Gaussian distributed, i.e.,

$$\mathrm{E}\left[\mathrm{Re}\{w_{j,k}\}\right] = \mathrm{E}\left[\mathrm{Im}\{w_{j,k}\}\right] = 0,$$

$$\mathrm{E}\left[\mathrm{Re}\{w_{j,k}\}\,\mathrm{Im}\{w_{j,k}\}\right] = 0,$$

$$(19) \qquad \mathrm{E}\left[(\mathrm{Re}\{w_{j,k}\})^2\right] = \mathrm{E}\left[(\mathrm{Im}\{w_{j,k}\})^2\right] = \sigma_{j,k}^2/2,$$

where $j = 1, 2, \ldots, N_r$, $k = 1, 2, \ldots, N_t$, and $\text{Re}\{\cdot\}$, $\text{Im}\{\cdot\}$ take the real and imaginary parts, respectively. We further assume that the elements of W are independent identically distributed, i.e.,

$$(20) \qquad \text{E}[w_{j,k}w_{j',k'}^*] = \sigma^2 \delta(j - j')\delta(k - k'),$$

where $j, j' = 1, 2, \ldots, N_r$, $k, k' = 1, 2, \ldots, N_t$, and $\delta(\cdot)$ stands for Kronecker delta function. Here the assumptions of whiteness and homogeneity are made for simplicity, we plan to extend the covariance structure of the noise matrix to a more general case in future work.

Note that the Green function representations in the multistatic matrix $K(\boldsymbol{x}, \boldsymbol{\tau})$ are completely general in the sense that the formulation could easily be adapted to modeling different operating scenarios by applying appropriate Green functions. Furthermore, since we make no assumptions on the antenna locations, our model can include various calibrated array configurations for instance, linear, planar, three-dimensional, etc., as long as the coherence among the antennas are preserved. Throughout this paper, we assume the number of scatterers M is *a priori* known, otherwise it could be estimated by examining the profile of the singular values of the multistatic matrix [36], or determined according to the information theoretic criteria [37], [38].

3. Scattering Parameter Estimation. Using the above statistical measurement models for the fixed-frequency computational time-reversal imaging methods, we develop maximum likelihood (ML) and suboptimal methods for estimating the location $\boldsymbol{x}$ and scattering coefficients $\boldsymbol{\tau}$ vectors.

Based on the same measurements of the scenario as in the computational time-reversal imaging, we derive ML estimators, which is known asymptotically unbiased and efficient [32]. Given the measurement Y, the likelihood function of unknown parameters $\boldsymbol{x}$, $\boldsymbol{\tau}$, and σ^2 is given in the form of the probability density function of the measurement as

$$(21) \qquad l_0(\boldsymbol{x}, \boldsymbol{\tau}, \sigma^2; Y) = p(Y; \boldsymbol{x}, \boldsymbol{\tau}, \sigma^2) = \frac{1}{(\pi\sigma^2)^{N_r N_t}} \exp(-\|Y - K(\boldsymbol{x}, \boldsymbol{\tau})\|_F^2/\sigma^2),$$

where $\|\cdot\|_F$ represents the Frobenius norm of a matrix and $\exp(\cdot)$ stands for the exponential function. The ML estimates of $\boldsymbol{x}$, $\boldsymbol{\tau}$ and nuisance parameter σ^2 are found by maximizing the likelihood function (21), i.e.,

$$(22) \qquad \hat{\boldsymbol{x}}, \hat{\boldsymbol{\tau}}, \hat{\sigma}^2 = \arg\max_{\boldsymbol{x}, \boldsymbol{\tau}, \sigma^2} \frac{1}{(\pi\sigma^2)^{N_r N_t}} \exp(-\|Y - K(\boldsymbol{x}, \boldsymbol{\tau})\|_F^2/\sigma^2).$$

Taking the derivative of $l_0(\boldsymbol{x}, \boldsymbol{\tau}, \sigma^2; Y)$ with respect to σ^2 and solving the estimation equation

$$(23) \qquad \frac{\partial l_0(\boldsymbol{x}, \boldsymbol{\tau}, \sigma^2; Y)}{\partial \sigma^2} = 0,$$

we get the MLE of σ^2 given $\boldsymbol{x}$ and $\boldsymbol{\tau}$ as

$$(24) \qquad \hat{\sigma}^2 = \frac{1}{N_{\mathrm{r}}N_{\mathrm{t}}} \|Y - K(\boldsymbol{x},\boldsymbol{\tau})\|_{\mathrm{F}}^2.$$

Substituting $\hat{\sigma}^2$ back into (21), we concentrate the likelihood function with respect to $\boldsymbol{x}$ and $\boldsymbol{\tau}$ as

$$(25) \qquad l_1(\boldsymbol{x},\boldsymbol{\tau};Y) = \frac{(N_{\mathrm{r}}N_{\mathrm{t}})^{N_{\mathrm{r}}N_{\mathrm{t}}}}{(\pi\|Y - K(\boldsymbol{x},\boldsymbol{\tau})\|_{\mathrm{F}}^2)^{N_{\mathrm{r}}N_{\mathrm{t}}}} \ \exp(-N_{\mathrm{r}}N_{\mathrm{t}}).$$

Now the MLE of $\boldsymbol{x}$ and $\boldsymbol{\tau}$ are found by

$$(26) \qquad \hat{\boldsymbol{x}}, \hat{\boldsymbol{\tau}} = \arg\min_{\boldsymbol{x},\boldsymbol{\tau}} \|Y - K(\boldsymbol{x},\boldsymbol{\tau})\|_{\mathrm{F}}^2.$$

To find the MLEs of $\boldsymbol{x}$ and $\boldsymbol{\tau}$ for the three-dimensional problem, we would need to solve the $5M$-dimensional nonlinear least-squares optimization problem in (26). In the following sections, we will solve this problem based on the physical models (8) and (17), respectively.

Note that if the transmit and receive arrays are co-located, the models (8) and (17) of the multistatic matrix will have a symmetric structure $K(\boldsymbol{x},\boldsymbol{\tau}) = K^T(\boldsymbol{x},\boldsymbol{\tau})$ since $A_{\mathrm{r}}(\boldsymbol{x}) = A_{\mathrm{t}}(\boldsymbol{x})$ using the reciprocity of the medium $G(\boldsymbol{r},\boldsymbol{x}',\omega) = G(\boldsymbol{x}',\boldsymbol{r},\omega)$. Then, we could replace the measurement matrix Y by its symmetric part $\tilde{Y} = (Y + Y^T)/2$ due to the following identity

$$(27) \qquad \|Y - K(\boldsymbol{x},\boldsymbol{\tau})\|_{\mathrm{F}}^2 = \|\tilde{Y} - K(\boldsymbol{x},\boldsymbol{\tau})\|_{\mathrm{F}}^2 + \|Y - \tilde{Y}\|_{\mathrm{F}}^2.$$

The proof follows from Lemma 1 (see the Appendix A for the proof).

3.1. Estimation Using the Distorted Wave Born Approximation. When the multiple scatterings among the scatterers can be assumed to be weak and negligible, we apply the physical model of the multistatic matrix using the distorted wave Born approximation. Substituting (8) into the cost function in (26), we have the following metric for estimating locations $\boldsymbol{x}$ and scattering coefficients $\boldsymbol{\tau}$

$$(28) \qquad l_2(\boldsymbol{x},\boldsymbol{\tau};Y) = \|Y - A_{\mathrm{r}}(\boldsymbol{x})T(\boldsymbol{\tau})A_{\mathrm{t}}^T(\boldsymbol{x})\|_{\mathrm{F}}^2$$

$$(29) \qquad\qquad\qquad = \|\mathrm{vec}(Y) - A_{\mathrm{t}}(\boldsymbol{x}) \odot A_{\mathrm{r}}(\boldsymbol{x})\boldsymbol{\tau}\|_{\mathrm{F}}^2,$$

where $\mathrm{vec}(\cdot)$ stacks the first to the last columns of the matrix one under another to form a long vector, $A_{\mathrm{t}}(\boldsymbol{x}) \odot A_{\mathrm{r}}(\boldsymbol{x}) = [\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}_1) \otimes \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}_1) \cdots \boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}_M) \otimes \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}_M)]$, here $\odot$ stands for the Khatri-Rao product [39], and $\otimes$ represents the Kronecker product [39]. In the second equality, we used the identity $\mathrm{vec}(AVD) = (D^T \odot A)\mathrm{vecd}(V)$ where V is diagonal and $\mathrm{vecd}(\cdot)$ forms a vector from the diagonal elements of the matrix (see T3.13 in [39]).

Observe the linearity of the vector in (29) in $\boldsymbol{\tau}$. Hence, given $\boldsymbol{x}$ we estimate $\boldsymbol{\tau}$ as a function of $\boldsymbol{x}$ and Y using the ordinary least-squares solution

$$(30) \qquad \hat{\boldsymbol{\tau}} = [A_{\odot}^H(\boldsymbol{x})A_{\odot}(\boldsymbol{x})]^{-1}A_{\odot}^H(\boldsymbol{x})\mathrm{vec}(Y),$$

where "^{H}" represents conjugate transpose and

$$(31) \qquad A_\odot(\boldsymbol{x}) \triangleq A_t(\boldsymbol{x}) \odot A_r(\boldsymbol{x}).$$

By Theorem 1 in [31], $A_r(\boldsymbol{x})$ and $A_t(\boldsymbol{x})$ are of full rank M under mild conditions if $M \leq N_r, N_t$. Then, it is not difficult to prove further that $A_\odot(\boldsymbol{x})$ is also of rank M, hence the inverse of $A_\odot^H(\boldsymbol{x})A_\odot(\boldsymbol{x})$ exists in (30), or the scattering coefficient $\boldsymbol{\tau}$ will not be identifiable anymore. We further concentrate (29) with respect to the location parameters $\boldsymbol{x}$ by substituting (30) into (29), and have

$$
\begin{aligned}
l_3(\boldsymbol{x}; Y) &= \|\mathrm{vec}(Y) - A_\odot(\boldsymbol{x})\hat{\boldsymbol{\tau}}\|_F^2 \\
&= \|\mathrm{vec}(Y) - A_\odot(\boldsymbol{x})[A_\odot^H(\boldsymbol{x})A_\odot(\boldsymbol{x})]^{-1}A_\odot^H(\boldsymbol{x})\mathrm{vec}(Y)\|_F^2 \\
(32) \qquad &= \|P_{A_\odot(\boldsymbol{x})}^\perp \mathrm{vec}(Y)\|_F^2,
\end{aligned}
$$

where $P_{A_\odot(\boldsymbol{x})}^\perp = I - A_\odot(\boldsymbol{x})[A_\odot^H(\boldsymbol{x})A_\odot(\boldsymbol{x})]^{-1}A_\odot^H(\boldsymbol{x})$ is the projection matrix that projects to the orthogonal complement subspace of the range of matrix $A_\odot(\boldsymbol{x})$, I is an identity matrix with appropriate dimension. Now the dimension of the optimization problem has been reduced from $5M$ in (29) to $3M$ in (32).

Observing (32), the likelihood-based optimization problem using the physical model (8) and measurement model (18) could be eventually interpreted as a subspace-based method: the vector $\mathrm{vec}(Y)$ representing the measured signal subspace should be orthogonal to the noise subspace $P_{A_\odot(\boldsymbol{x})}^\perp$, the orthogonal complement subspace of the range of $A_\odot(\boldsymbol{x}) = A_t(\boldsymbol{x}) \odot A_r(\boldsymbol{x})$ in $\mathbb{C}^{N_r N_t}$.

To summarize, to compute MLEs of $\boldsymbol{x}$ and $\boldsymbol{\tau}$, we

1. solve

$$(33) \qquad \hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|P_{A_\odot(\boldsymbol{x})}^\perp \mathrm{vec}(Y)\|_F^2,$$

2. compute

$$(34) \qquad \hat{\boldsymbol{\tau}} = [A_\odot^H(\hat{\boldsymbol{x}})A_\odot(\hat{\boldsymbol{x}})]^{-1}A_\odot^H(\hat{\boldsymbol{x}})\mathrm{vec}(Y),$$

3. and

$$(35) \qquad \hat{\sigma}^2 = \frac{1}{N_r N_t}\|Y - K(\hat{\boldsymbol{x}}, \hat{\boldsymbol{\tau}})\|_F^2.$$

The first step of the above MLE procedure is nonlinear, requiring a numerical approach. Starting from an initial estimate of $\boldsymbol{x}$, we solve (33) by a quasi-Newton algorithm consisting of: (i) update the approximation of the Hessian matrix and compute the search direction using the BFGS algorithm [40]-[43], and (ii) update the estimation using the mixed quadratic and cubic polynomial method. We implemented these two phases by the MATLAB optimization toolbox.

3.2. Estimation Using the Foldy-Lax Multiple Scattering Model. For the case where multiple scattering among the scatterers are non-negligible, for example when the scatterers are closely located and their scattering amplitudes are large, it is necessary to apply an appropriate physical model incorporating the underlying multiple scattering effect to obtain accurate estimation. We will then use the physical model (17), which is based on the Foldy-Lax multiple scattering model.

The ML estimate of $\boldsymbol{x}$ and $\boldsymbol{\tau}$ based on the Foldy-Lax multiple scattering model is

$$(36) \qquad \hat{\boldsymbol{x}}, \hat{\boldsymbol{\tau}} = \arg\min_{\boldsymbol{x},\boldsymbol{\tau}} \|Y - A_{\mathrm{r}}(\boldsymbol{x})[T^{-1}(\boldsymbol{\tau}) - S(\boldsymbol{x})]^{-1} A_{\mathrm{t}}^T(\boldsymbol{x})\|_{\mathrm{F}}^2.$$

Since $K(\boldsymbol{x}, \boldsymbol{\tau})$ is nonlinear with respect to both the location parameters $\boldsymbol{x}$ and scattering coefficient parameters $\boldsymbol{\tau}$, it is difficult to reduce the dimension of the optimization problem (36) by finding a concentrated metric as we did in (32), and it is inevitable to resort to an iterative algorithm or gradient-based numerical procedure. A similar quasi-Newton algorithm as in Section 3.1 could be used for solving $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{\tau}}$. Here, we propose a sub-optimal estimation method, which is a tradeoff between computation complexity and optimality.

It is shown in [27] that the MUSIC-based time-reversal imaging method (see (47) in Section 4), which was first developed using the distorted wave Born approximation model, is still applicable to the multiple scattering scenario. This conclusion is also supported by the observation that the range of $A_{\mathrm{r}}(\boldsymbol{x})[T^{-1}(\boldsymbol{\tau}) - S(\boldsymbol{x})]^{-1} A_{\mathrm{t}}^T(\boldsymbol{x})$ is equal to the range of $A_{\mathrm{r}}(\boldsymbol{x})$, which is the signal subspace using the distorted wave Born approximation model, as long as $T^{-1}(\boldsymbol{\tau}) - S(\boldsymbol{x})$ is a full rank matrix.

However, a similar conclusion does not hold in the likelihood-based method. It is interesting to note that when using the wave Born approximation, the signal subspace for the measurement $\mathrm{vec}(Y)$ is the range of $A_{\odot}(\boldsymbol{x}) = A_{\mathrm{t}}(\boldsymbol{x}) \odot A_{\mathrm{r}}(\boldsymbol{x})$, which is spanned by $\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}_m) \otimes \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}_m)$, $m = 1, 2, \ldots, M$; whereas, when the Foldy-Lax model is applied, the signal subspace becomes the range of $A_{\otimes}(\boldsymbol{x}) \triangleq A_{\mathrm{t}}(\boldsymbol{x}) \otimes A_{\mathrm{r}}(\boldsymbol{x})$, which is spanned by $\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}_m) \otimes \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}_{m'})$, $m, m' = 1, 2, \ldots, M$ due to the following equality

$$(37) \ \mathrm{vec}\big(A_{\mathrm{r}}(\boldsymbol{x})[T^{-1}(\boldsymbol{\tau}) - S(\boldsymbol{x})]^{-1} A_{\mathrm{t}}^T(\boldsymbol{x})\big) = A_{\mathrm{t}}(\boldsymbol{x}) \otimes A_{\mathrm{r}}(\boldsymbol{x}) \mathrm{vec}\big([T^{-1}(\boldsymbol{\tau}) - S(\boldsymbol{x})]^{-1}\big).$$

Here we used the identity $\mathrm{vec}(ADB) = (B^T \otimes A)\mathrm{vec}D$.

By this observation, we propose a sub-optimal estimator of $\boldsymbol{x}$ for the Foldy-Lax model that minimizes the following metric

$$l_4(\boldsymbol{x}; Y) = \|P_{A_{\otimes}(\boldsymbol{x})}^{\perp} \mathrm{vec}(Y)\|_{\mathrm{F}}^2$$
$$(38) \qquad\qquad = \|(P_{A_{\mathrm{t}}(\boldsymbol{x})} \otimes P_{A_{\mathrm{r}}(\boldsymbol{x})})^{\perp} \mathrm{vec}(Y)\|_{\mathrm{F}}^2,$$

where the second equality follows by the Corollary in Appendix B. It is easy to show that $P_{A_{\otimes}(\boldsymbol{x})}$ is of rank M^2 under mild conditions. The metric (38) is essentially a

modification of (32) by applying appropriate subspace for the multiple scattering case, but the dimension of the signal subspace increases from M to M^2.

Denoting $B = \left([T^{-1}(\boldsymbol{\tau}) - S(\boldsymbol{x})]^{-1}\right)$ and $\boldsymbol{b} = \mathrm{vec}(B)$, the statistical interpretation of this metric is that minimizing (38) will give an "ML" estimate of $\boldsymbol{x}$ for the problem in which the inner structure of $\boldsymbol{b}$ is ignored, i.e., treat $\boldsymbol{b}$ as an unconstrained vector in $\mathbb{C}^{M^2}$. Once we obtain $\hat{\boldsymbol{x}}$, we could estimate $\boldsymbol{\tau}$ as

$$(39) \qquad \hat{\boldsymbol{\tau}} = \mathrm{vecd}([S(\hat{\boldsymbol{x}}) + \hat{B}^{-1}]^{-1}),$$

where $\mathrm{vec}(\hat{B}) = \hat{\boldsymbol{b}}$ and

$$
\begin{aligned}
\hat{\boldsymbol{b}} &= [(A_\mathrm{t}(\hat{\boldsymbol{x}}) \otimes A_\mathrm{r}(\hat{\boldsymbol{x}}))^H (A_\mathrm{t}(\hat{\boldsymbol{x}}) \otimes A_\mathrm{r}(\hat{\boldsymbol{x}}))]^{-1} (A_\mathrm{t}(\hat{\boldsymbol{x}}) \otimes A_\mathrm{r}(\hat{\boldsymbol{x}}))^H \mathrm{vec}(Y) \\
(40) \qquad &= \left\{ \left([A_\mathrm{t}^H(\hat{\boldsymbol{x}})A_\mathrm{t}(\hat{\boldsymbol{x}})]^{-1} A_\mathrm{t}^H(\hat{\boldsymbol{x}})\right) \otimes \left([A_\mathrm{r}^H(\hat{\boldsymbol{x}})A_\mathrm{r}(\hat{\boldsymbol{x}})]^{-1} A_\mathrm{r}^H(\hat{\boldsymbol{x}})\right) \right\} \mathrm{vec}(Y).
\end{aligned}
$$

The second equality could be proved by the same steps as the proof of Lemma 2 in Appendix B.

4. Likelihood Time-reversal Imaging. In this section, we propose an imaging metric for the fixed-frequency time-reversal imaging methods using the likelihood function for a simplified physical model of the multistatic matrix, namely for a single-scatterer model where

$$(41) \qquad K(\boldsymbol{x}_1, \tau_1) = \tau_1 \boldsymbol{g}_\mathrm{r}(\boldsymbol{x}_1) \boldsymbol{g}_\mathrm{t}^T(\boldsymbol{x}_1).$$

That is, we apply this model even if in reality the number of scatterers is larger than one, thus reducing the optimization to a dimension of five. The resulting simplified estimation or scanning scheme is similar to [44] and is useful for example to initialize the numerical optimization of the MLE for the full physical model.

Note that this single-scatterer model is a special case of the one using the distorted wave Born approximation (8) (of course, also a special case of the one using Foldy-Lax model (17)) and location parameter $\boldsymbol{x} = \boldsymbol{x}_1$. We plug $A_\mathrm{r}(\boldsymbol{x}) = \boldsymbol{g}_\mathrm{r}(\boldsymbol{x})$ and $A_\mathrm{t}(\boldsymbol{x}) = \boldsymbol{g}_\mathrm{t}(\boldsymbol{x})$ into (30), the least-squares solution of τ_1 given $\boldsymbol{x}_1$ is found as

$$
\begin{aligned}
\hat{\tau}_1 &= \frac{[\boldsymbol{g}_\mathrm{t}^H(\boldsymbol{x}) \otimes \boldsymbol{g}_\mathrm{r}^H(\boldsymbol{x})]\mathrm{vec}(Y)}{[\boldsymbol{g}_\mathrm{t}^H(\boldsymbol{x}) \otimes \boldsymbol{g}_\mathrm{r}^H(\boldsymbol{x})][\boldsymbol{g}_\mathrm{t}(\boldsymbol{x}) \otimes \boldsymbol{g}_\mathrm{r}(\boldsymbol{x})]} \\
&= \frac{[\boldsymbol{g}_\mathrm{t}^H(\boldsymbol{x}) \otimes \boldsymbol{g}_\mathrm{r}^H(\boldsymbol{x})]\mathrm{vec}(Y)}{\boldsymbol{g}_\mathrm{t}^H(\boldsymbol{x})\boldsymbol{g}_\mathrm{t}(\boldsymbol{x}) \otimes \boldsymbol{g}_\mathrm{r}^H(\boldsymbol{x})\boldsymbol{g}_\mathrm{r}(\boldsymbol{x})} \\
(42) \qquad &= \frac{\boldsymbol{g}_\mathrm{t}^H(\boldsymbol{x}) \otimes \boldsymbol{g}_\mathrm{r}^H(\boldsymbol{x})}{[\boldsymbol{g}_\mathrm{t}^H(\boldsymbol{x})\boldsymbol{g}_\mathrm{t}(\boldsymbol{x})][\boldsymbol{g}_\mathrm{r}^H(\boldsymbol{x})\boldsymbol{g}_\mathrm{r}(\boldsymbol{x})]}\mathrm{vec}(Y),
\end{aligned}
$$

where the second equality follows by the identity $(A \otimes B)(C \otimes D) = AC \otimes BD$. Then,

the concentrated cost function for solving $\hat{\boldsymbol{x}}$ is

(43) $l_5(\boldsymbol{x}; Y) = \|P_{\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}) \otimes \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x})}^{\perp} \mathrm{vec}(Y)\|_{\mathrm{F}}^2$

$\qquad = \|\mathrm{vec}(Y) - [\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}) \otimes \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x})] \dfrac{\boldsymbol{g}_{\mathrm{t}}^H(\boldsymbol{x}) \otimes \boldsymbol{g}_{\mathrm{r}}^H(\boldsymbol{x})}{[\boldsymbol{g}_{\mathrm{t}}^H(\boldsymbol{x})\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x})][\boldsymbol{g}_{\mathrm{r}}^H(\boldsymbol{x})\boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x})]} \mathrm{vec}(Y)\|_{\mathrm{F}}^2$

$\qquad = \|[I - \dfrac{\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x})\boldsymbol{g}_{\mathrm{t}}^H(\boldsymbol{x}) \otimes \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x})\boldsymbol{g}_{\mathrm{r}}^H(\boldsymbol{x})}{[\boldsymbol{g}_{\mathrm{t}}^H(\boldsymbol{x})\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x})][\boldsymbol{g}_{\mathrm{r}}^H(\boldsymbol{x})\boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x})]}] \mathrm{vec}(Y)\|_{\mathrm{F}}^2$

(44) $\qquad = \|Y - \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x})\boldsymbol{g}_{\mathrm{r}}^H(\boldsymbol{x}) Y \boldsymbol{g}_{\mathrm{t}}^*(\boldsymbol{x})\boldsymbol{g}_{\mathrm{t}}^T(\boldsymbol{x}) / [\boldsymbol{g}_{\mathrm{t}}^H(\boldsymbol{x})\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x})] / [\boldsymbol{g}_{\mathrm{r}}^H(\boldsymbol{x})\boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x})]\|_{\mathrm{F}}^2,$

where the last equality follows by the identities $\mathrm{vec}(ADB) = (B^T \otimes A)\mathrm{vec}D$, $\mathrm{vec}(A + B) = \mathrm{vec}(A) + \mathrm{vec}(B)$, and $\|A\|_{\mathrm{F}} = \|\mathrm{vec}(A)\|_{\mathrm{F}}$.

Now, the only unknown parameter in this concentrated metric (44) is the location parameter $\boldsymbol{x}$ of one scatterer, thus the dimension of the optimization problem in (44) has been reduced to 3, rather than $3M$ searches necessary in (32) and (38). We propose to create the *likelihood time-reversal image* by evaluating the following metric over a fine grid of the probed scenario

(45) $\qquad l_{\mathrm{s}}(\boldsymbol{x}; Y) = 1/l_5(\boldsymbol{x}; Y) = \dfrac{1}{\|P_{\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}) \otimes \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x})}^{\perp} \mathrm{vec}(Y)\|_{\mathrm{F}}^2},$

and use it as the imaging metric. Note that the final concentrated likelihood function using the single scatterer model is

(46) $\qquad l_6(\boldsymbol{x}; Y) = \dfrac{(N_{\mathrm{r}}N_{\mathrm{t}})^{N_{\mathrm{r}}N_{\mathrm{t}}}}{(\pi \|P_{\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}) \otimes \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x})}^{\perp} \mathrm{vec}(Y)\|_{\mathrm{F}}^2)^{N_{\mathrm{r}}N_{\mathrm{t}}}} \ \exp(-N_{\mathrm{r}}N_{\mathrm{t}}),$

and both the concentrated metrics (45) and (46) are monotonically decreasing functions with respect to $\|P_{\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}) \otimes \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x})}^{\perp} \mathrm{vec}(Y)\|_{\mathrm{F}}^2$, thus the imaging metric (45) is a monotonically increasing function of the concentrated likelihood function (46). Therefore, the image represents the approximated relative likelihood, supported by the measurement Y, of the existence of a single scatterer at each imaging location, and the peaks in the image will indicate locations of the possible scatterers. Since this imaging is based on the simplified physical model (41), it is sub-optimal in its nature, but it useful also to initialize the non-linear optimization procedures in Section 3.1 and Section 3.2 by using the locations of the first M local maxima as the initial estimates of $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_M$.

We note that in [31], the following MUSIC-based pseudo-spectrum is used as a fixed-frequency computational time-reversal imaging metric for the co-located transmit and receive arrays case, (where $N_{\mathrm{r}} = N_{\mathrm{t}} = N$, and $\boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}) = \boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}) = \boldsymbol{g}(\boldsymbol{x})$)

(47) $\qquad D(\boldsymbol{x}) = \dfrac{1}{\sum_{m_0=M+1}^{N} |\langle \boldsymbol{\mu}_{m_0}^*, \boldsymbol{g}(\boldsymbol{x}) \rangle|^2},$

where $\boldsymbol{\mu}_{m_0}$ is the m_0-th eigenvector of time-reversal matrix $T = Y^*Y$ having zero eigenvalue. This MUSIC algorithm makes use of the fact that the time-reversal matrix

T has the same range as the subspace spanned by the complex conjugates of the Green function vectors when noiseless, which is the so-called signal subspace in [31], and the noise subspace is spanned by the eigenvectors of T corresponding to zero eigenvalues. Similar to the observation we made in Section 3.1, the likelihood-based optimization problem (43) finally becomes a subspace-based method. However, the orthogonality between the signal subspace and noise subspaces is employed differently in the likelihood-based (43) and MUSIC-based (47). In the likelihood-based (43), both the signal and noise subspaces are in $\mathbb{C}^{N_\mathrm{r}N_\mathrm{t}}$, the noise subspace represented by $P^\perp_{\boldsymbol{g}_\mathrm{t}(\boldsymbol{x})\otimes\boldsymbol{g}_\mathrm{r}(\boldsymbol{x})}$ arises from the physical model (41) and the signal subspace from the measurement $\mathrm{vec}(Y)$, whereas in the MUSIC-based (47), the signal and noise subspaces are in $\mathbb{C}^N$, the noise subspace is estimated from the measurement Y and the signal space comes from the Green function vector $\boldsymbol{g}(\boldsymbol{x})$. In addition, during the likelihood-based imaging process the noise subspace is a function $\boldsymbol{x}$, thus evaluated at every imaging position, but in the MUSIC-based imaging the signal subspace $\boldsymbol{g}(\boldsymbol{x})$ is evaluated. Comparing (47) with (32) and (38), the MUSIC-based imaging does not employ the physical model of the multistatic matrix except for the array manifold and the orthogonality between the signal and noise subspaces, thus could be less accurate if the model is accurate. Moreover, the MUSIC-based method cannot be extended to more complex models, such as unknown spatially correlated noise. The proposed statistically based estimator, instead, is more scalable in the sense that it could be extended to account for spatially correlated noise by employing a more realistic noise model as well as a more general physical model. All these extensions will finally result in more general optimization metrics, of course at the expense of higher computational complexity.

5. Multiple-frequency Likelihood Time-reversal Imaging. In this section, we generalize our fixed-frequency likelihood time-reversal imaging method to multiple frequencies by combining the imaging metrics at various frequencies. We demonstrate its usefulness in resolving the spatial ambiguity of a sparse array.

The original motivation for this generalization is to solve the dilemma in the tradeoff between the array size and resolution capability in the fixed-frequency methods by introducing a new degree of freedom. Using a homogeneous medium as an example, in order to keep the cross-range and range diffraction resolutions unchanged with the increasing of the range R, we need to increase the array aperture a linearly accordingly, due to facts that the diffraction resolution of the refocused field in the cross-range direction is $\lambda R/a$, and $\lambda(R/a)^2$ in the range direction [3], [4]. This requires a large number of antennas when R is large for arrays with antennas densely spaced. One possible solution to avoid the need for this large number of antennas is to employ a sparse array whose antenna spacing is much larger than half of the wavelength. However, it is well-known that a spatial aliasing will occur when we

undersample in the spatial domain, and the introduced spatial ambiguity cannot be resolved without any *a priori* information.

To illustrate this spatial ambiguity, we use the *coherent point spread function* (CPSF) [31] of a co-located array, which is defined as

$$(48) \qquad\qquad H(\boldsymbol{r}, \boldsymbol{x}') \triangleq \langle \boldsymbol{g}(\boldsymbol{r}), \boldsymbol{g}(\boldsymbol{x}') \rangle.$$

The CPSF of the array represents an "image" of a point source at position $\boldsymbol{x}'$ formed by the antenna array from the measurement of the outgoing wave Green function $G(\boldsymbol{r}, \boldsymbol{x}')$ at various array elements [31]. Note that if we normalize the coherent point spread function of the antenna array in by the number of antennas, the quantity $\frac{1}{N} H(\boldsymbol{r}, \boldsymbol{x}')$ is exactly a generalization of the *conventional beam pattern* [45] of an array, thus intrinsic to the array. We plot the amplitudes of the CPSF (48) for two uniform linear arrays (ULAs) with antennas uniformly distribute on the line between (-20,1) and (20,1): in Figure 2, the antenna spacing is 1/2 and in Figure 3 it is 5. Wavelength $\lambda = 1$ in both figures and we assume the point source is located at $\boldsymbol{x}' = (-5, 20)$. Compared with the dense array having no spatial aliasing in Figure 2, we could see that the CPSF of the sparse array in Figure 3 has a complicated structure that consists of ridges and valleys and local maxima appear besides the true source location, and these local maxima correspond to grating lobes. To resolve this spatial ambiguity, we propose to introduce the frequency diversity.

Many techniques dealing with wideband signal processing have been proposed for passive direction finding under the conventional array signal processing framework in the existing literature. Wideband beamforming aims at producing frequency-invariant beam-patterns for a wide signal bandwidth [46]. Nonoverlapping narrowbands are combined to obtain direction of arrival (DOA) estimates in [47], [48]. A coherent signal-subspace (CSS) method [49] decomposes the wideband signal into narrowband components, and transforms the narrowband array manifold matrices into a matrix corresponding to a selected reference frequency, then apply narrowband (DOA) estimation methods. For reviews on wideband sensor array processing, see [46], [50].

Unlike these techniques, we will make use of the distinctness embedded in the measurements at different frequencies for the purpose of resolving the spatial ambiguity of a sparse array. In Figure 4, we display the amplitude of the CPSF under the same setup as in Figure 3 except using wavelength $\lambda = 0.8$. Compared with Figure 3, we can see that the ambiguity patterns, i.e., the distribution of the grating lobes is a function of the wavelength, however, the peak at the scatterer location, which is indicated by "∘", remains. For this reason, we prefer not to treat the multistatic matrices at different frequencies as simple repeated measurements in frequency domain, but observations of the probed scenario from diverse perspectives. Using multistatic matrices measured at multiple frequencies, we generalize the fixed-frequency likelihood time-reversal imaging methods to the multiple-frequency version by combining
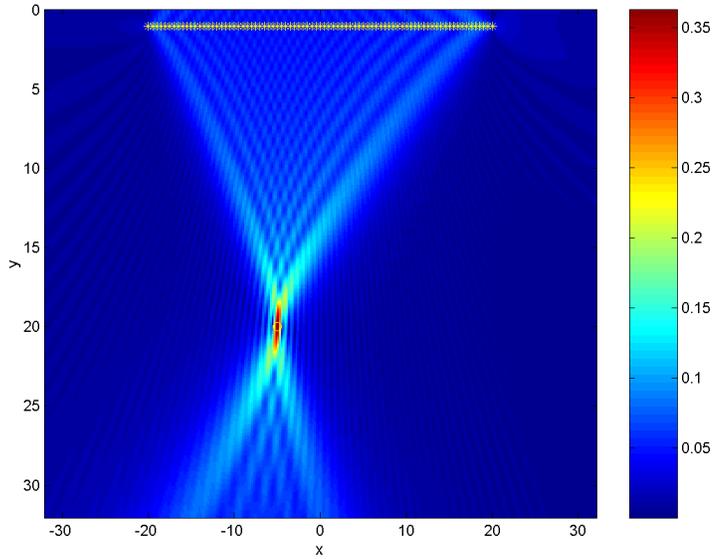
FIG. 2. *amplitude of the coherent point spread function (CPSF) for a uniform linear array (ULA) with a point source located at (-5,20), wavelength $\lambda = 1$ and antenna spacing of half wavelength. $*$: antenna, $\circ$: scatterer.*
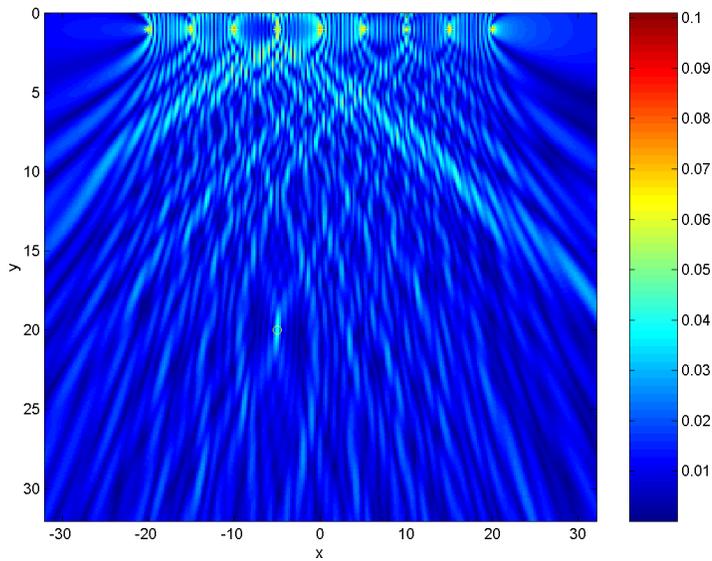
FIG. 3. *amplitude of the CPSF for a ULA with a point source located at (-5,20), wavelength $\lambda = 1$ and antenna spacing of 5.*
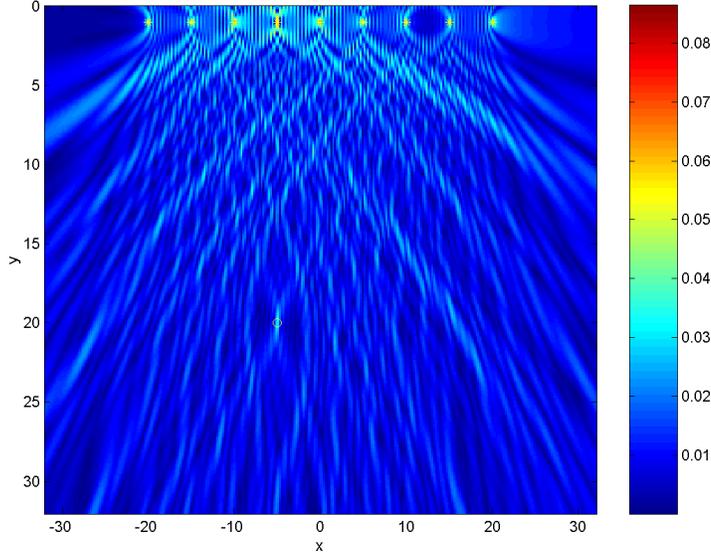
FIG. 4. *amplitude of the CPSF for a ULA with a point source located at (-5,20), wavelength* $\lambda = 0.8$ *and antenna spacing of* 5.

likelihood imaging metrics at different frequencies via multiplication, i.e., using the following metric for the multiple-frequency likelihood imaging

(49)
$$
\begin{aligned}
l_7(\boldsymbol{x}; Y_1, Y_2 \cdots Y_L) &= \prod_{l=1}^{L} l_{\mathrm{s}}(\boldsymbol{x}; \omega_l, Y_l), \\
&= \frac{1}{\prod_{l=1}^{L} \| P^{\perp}_{\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x},\omega_l) \otimes \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x},\omega_l)} \mathrm{vec}(Y_l) \|_{\mathrm{F}}^2},
\end{aligned}
$$

where Y_l is the multistatic matrix measured at the l-th frequency ω_l, $l = 1, 2, \ldots, L$, and L is the number of total frequencies used. Here we recall that $l_{\mathrm{s}}(\cdot)$ depends on the frequency through the receive Green function vector $\boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x})$ and transmit Green function vector $\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x})$. We emphasize this by changing the notations into $l_{\mathrm{s}}(\boldsymbol{x}; \omega, Y)$, $\boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}, \omega)$, and $\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}, \omega)$. Intuitively, the imaging metric $l_7(\boldsymbol{x}; Y_1, Y_2 \cdots Y_L)$ will peak at the position where $\boldsymbol{x}$ coincides with the scatterer location, since every $l_{\mathrm{s}}(\boldsymbol{x}; \omega_l, Y_l)$ has a local maximum at the true scatterer location for $l = 1, 2, \ldots, L$; on the other hand, when $l_7(\boldsymbol{x}; Y_1, Y_2 \cdots Y_L)$ is evaluated at the grating lobes for some frequency, there exists at least one frequency $\omega_{l'}$ such that $l_{\mathrm{s}}(\boldsymbol{x}; \omega_{l'}, Y_{l'})$ is small due to the different ambiguity patterns at different frequencies.

Assuming independence between different frequencies [51], the choice of the multiplication instead of an addition as the scheme of combining information from multiple-frequency data in (49) is justified by the concentrated likelihood function of the fol-

lowing estimation problem. Let Y_l be modeled as

(50) $$Y_l = \tau_1(\omega_l)\boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}_1, \omega_l)\boldsymbol{g}_{\mathrm{t}}^T(\boldsymbol{x}_1, \omega_l) + W(\omega_l),$$

where $W(\omega_1), W(\omega_2), \ldots, W(\omega_L)$ are zero-mean jointly circularly symmetric complex Gaussian distributed, $\mathrm{E}\{\mathrm{vec}(W(\omega_l))\mathrm{vec}(W(\omega_l))^H\} = \sigma_l^2 I$, and $W(\omega_l)$ and $W(\omega_{l'})$ are independent for $l \neq l'$. The concentrated likelihood with respect to $\boldsymbol{x} = \boldsymbol{x}_1$ is found as

(51) $$l_8(\boldsymbol{x}; Y_1, Y_2 \cdots Y_L) = \frac{(N_{\mathrm{r}}N_{\mathrm{t}})^{LN_{\mathrm{r}}N_{\mathrm{t}}} \exp(-LN_{\mathrm{r}}N_{\mathrm{t}})}{\pi^{LN_{\mathrm{r}}N_{\mathrm{t}}} \prod_{l=1}^{L} \|P_{\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}, \omega_l) \otimes \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}, \omega_l)}^{\perp} \mathrm{vec}(Y_l)\|_{\mathrm{F}}^{2N_{\mathrm{r}}N_{\mathrm{t}}}},$$

see Appendix C for the derivation. Since the imaging metric (49) is a monotonically increasing function of (51), the multiple-frequency imaging scheme could still be interpreted using a likelihood argument as in the fixed-frequency likelihood time-reversal imaging.

Observing that (49) is simply a product of the imaging metrics at every single frequency, the multiple-frequency image could be easily updated if more measurements are obtained at additional frequencies. In the time domain time-reversal imaging methods [3], the backpropagated waves on the whole spectrum are combined coherently using the inverse Fourier transform, which is more computationally expensive compared with the multiple-frequency methods. Furthermore, the proposed multiple-frequency imaging should not be considered as a simplification of the time domain imaging methods since it is much more flexible in the sense that the metric selected for imaging and the way of synthesizing are more diverse than those in the time domain methods.

6. Numerical and Experimental Examples. We present numerical and experimental examples to demonstrate the applicabilities and the performance of the proposed methods, namely the fixed-frequency likelihood time-reversal imaging, multiple-frequency likelihood time-reversal imaging, ML and sub-optimal estimators.

6.1. Numerical Examples. For convenience of visualization, we consider the two-dimensional case, i.e., the antenna elements and scatterers are parallel lines and they are embedded in homogeneous background; then, from a mathematical point of view the locations of targets and antennas could be represented as points in $\mathbb{R}^2$. The Green function for this two-dimensional case is

(52) $$G(\boldsymbol{r}, \boldsymbol{x}') = \frac{i}{4}H_0(2\pi|\boldsymbol{r} - \boldsymbol{x}'|/\lambda),$$

where H_0 is the zero order Hankel function of the first kind, see [33]. We will drop the unessential constant $i/4$ in the simulations. In each example, all the multistatic matrices are computed using (17), i.e., incorporating all the multiple scattering among the scatterers, and corrupted by the white Gaussian noise using (18). In all the
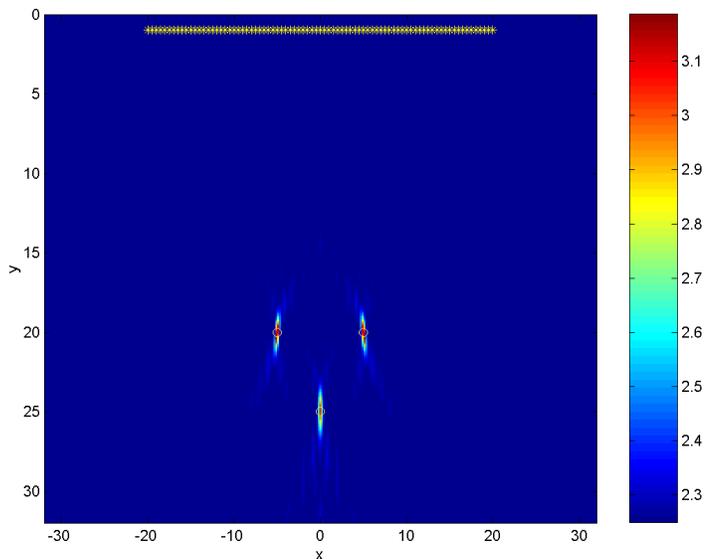
FIG. 5. *likelihood time-reversal image of point scatterers located at (-5,20), (0,25), and (5,20) using a ULA with wavelength $\lambda = 1$ and antenna spacing of half wavelength.*

numerical examples, we employ co-located ULAs as transmit and receive arrays and scatterers is assumed to have unit scattering coefficients.

In the first example, we demonstrate the time-reversal image using the likelihood imaging metric (45). The transmit and receive ULAs are located between (-20,1) and (20,1) with antenna spacing 1/2, the wavelength $\lambda = 1$, three scatterers, which are represented by "○", locate at (-5,20), (0,25), and (5,20), respectively. The image is generated over the grid of 301×301 and the antennas are denoted as "*". In Figure 5, the peaks indicate the correct scatterer locations. However, when the antenna spacing increases to 5 a spatial ambiguity appears in Figure 6. We can see a number of grating lobes in this image.

In the second example, seven frequencies are used in the multiple-frequency likelihood imaging, which correspond to the wavelength $\lambda = 0.60, 0.73, 0.87, 1, 1.13, 1.27,$ and 1.4. The simulation setup is the same as that of Figure 6, and the multiple-frequency likelihood image using the (49) is shown in Figure 7, from which we could see that the grating lobes are suppressed effectively and the true scatterer locations are resolved without ambiguities. Note that 81 antennas are used in Figure 5, whereas only 9 antennas are used in Figure 7.

In the third example, we examine the performances of the ML and three sub-optimal estimators of the location parameters in terms of the mean-squared error
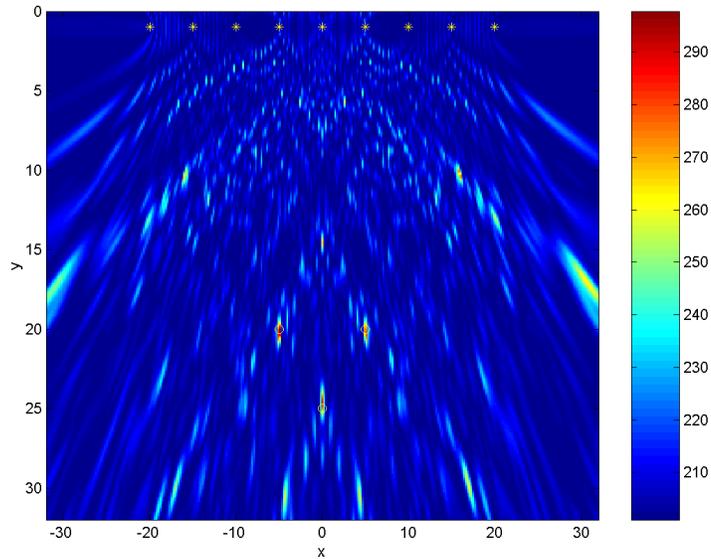
FIG. 6. *likelihood time-reversal image of point scatterers located at (-5,20), (0,25), and (5,20) using a ULA with wavelength $\lambda = 1$ and antenna spacing of* 5.
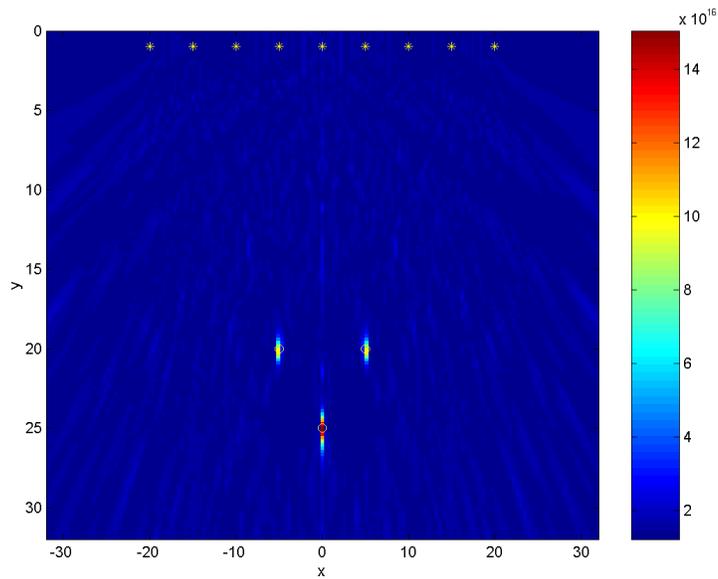
FIG. 7. *multiple-frequency likelihood time-reversal image of point scatterers located at (-5,20), (0,25), and (5,20) using a ULA with wavelengths $\lambda = 0.60, 0.73, 0.87, 1, 1.13, 1.27,$ and 1.4 and antenna spacing of* 5.
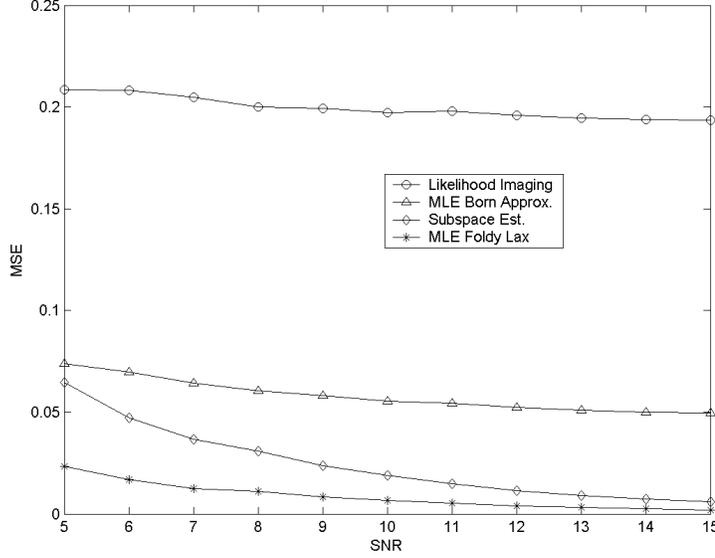
FIG. 8. *mean-squared error (MSE) of performance metrics: (43) in the likelihood imaging, (32) using the distorted wave Born approximation, (38) that partially explores the Foldy-Lax model, and (36) which uses the full Foldy-Lax model.*

(MSE) and the signal-to-noise ratio (SNR) is defined as

$$\text{(53)} \qquad \text{SNR} = \frac{\|K(\boldsymbol{x}, \boldsymbol{\tau})\|_{\text{F}}^2}{\text{E}\{\|W\|_{\text{F}}^2\}} = \frac{\|K(\boldsymbol{x}, \boldsymbol{\tau})\|_{\text{F}}^2}{N^2 \sigma^2}.$$

By 1500 Monte Carlo runs, we compute the MSEs of the MLE (36) which uses the full Foldy-Lax model, (43) in the likelihood imaging, (32) using the distorted wave Born approximation, and (38) that partially employs the Foldy-Lax model. The setups of the antennas, wavelength, and scatterers are the same as those in Figure 7. We assume knowledge of the scatterer locations in this example so that we do not have to resolve the spatial ambiguity of the sparse array when using only single frequency measurement. As expected, the likelihood imaging scheme has the largest MSE among the four estimators, and the MLE (36) based on full Foldy-Lax model performs the best. Note that the performance of the estimator based on the metric (38) improves quickly with increasing of SNR in Figure 8, whereas the estimator using (32) using the Born approximation does not improve much, thus less efficient in high SNR case. The subspace-based metric (38) that partially exploits the Foldy-Lax model is proved to be a good trade-off between performance and complexity.

6.2. Experimental Examples. In this section, we use experimental data provided by our colleagues at Carnegie Mellon University to demonstrate the applicability of the proposed methods in this paper. The experimental setup is illustrated in Figure
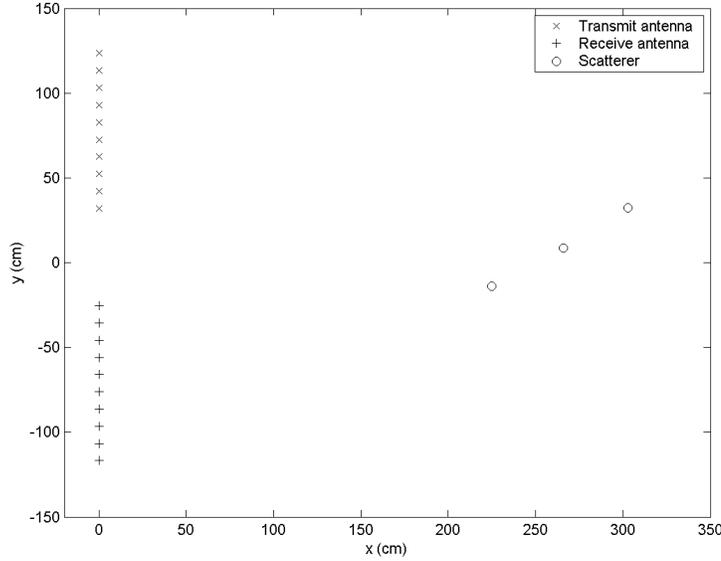
FIG. 9. *Carnegie Mellon University experiment setup illustration. x: transmit antenna, +: receive antenna, ∘: scatterer, units are in centimeter.*

9, where a transmit array (represented by "x") and a receive array (represented by "+") are spatially separated, both are located on the y-axis. In the experiment, the transmit and receive arrays are realized by one pair of transmit and receive antennas at different element positions, and $K_{j,k}$ is measured as the channel from the transmit antenna (located at the position of the k-th element of the transmit array) to the receive antenna (located at the position of the j-th element of the receive array). Cylinder pipes are used as scatterers, they are placed perpendicular to the x-y plane, and are represented by their their cross-sections (circles) in Figure 9. An absorbing wall is employed behind the test scenario preventing interference scattering from the laboratory environment. The measurements were taken at 201 frequency points ranging from 4 G Hz to 6 G Hz by a vector network analyzer.

In the first experimental example, we demonstrate the fixed-frequency and multiple-frequency likelihood imaging. We use the three scatterers data set, where the experiment was conducted using three copper pipes as scatterers located at (302.7, 32.3), (224.7, -14), and (265.9, 8.6) in the x-y plane with a distance unit of centimeter. The transmit array antennas are of 10 elements located uniformly from (0, 123.4) to (0, 32.0) with spacing 10.16 cm (4 inches), and the receive array antennas are of 10 elements located uniformly from (0, -25.4) to (0, -116.8) with spacing 10.16 cm, too. Figure 10 shows the likelihood image using the measurement of the multistatic matrix at central frequency 5 G Hz, from which we could see that the likelihood image peaks
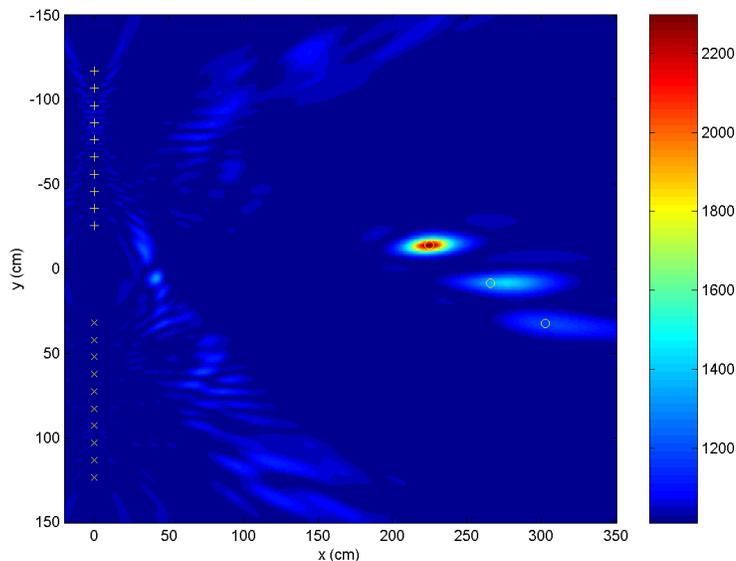
FIG. 10. *fixed-frequency likelihood image using measurement at central frequency 5 G Hz.*

around the locations of the three scatterers and grating lobes appear in the imaging domain. The multiple-frequency likelihood image is shown in Figure 11, in which we use the measurements at 4 G Hz, 4.5 G Hz, 5 G Hz, 5.5 G Hz, and 6 G Hz. Compared with the fixed-frequency image, we observe that the quality of the multiple-frequency case is improved (i.e. the grating lobes are mitigated).

In the second example, we show the scatterer location estimation by the likelihood imaging and ML estimates. Using the same setup and data set at 5 G Hz in the last example, we applied the quasi-Newton algorithm to estimate the scatterer locations by minimizing corresponding metrics. Using the single-scatterer model (41) for likelihood imaging, we find the estimates of the locations of the three scatterers as the three local maxima of the metric (45). Using the estimates from the likelihood imaging as the initial values, we found the ML estimate (33) using the Born approximation. These estimates are displayed in Figure 12, which show that the likelihood imaging provide a good initial estimation and the ML method refine the estimation further. Minimizing (38), we find the estimates based on the Foldy-Lax model, which gives estimates very close to those using the Born approximation, thus the multiple scattering is weak in this experiment.

In the third set of examples, we demonstrate the application of the likelihood time-reversal imaging in detecting a target in a rich scattering scenario. In these examples, one copper pipe is used as the target located at (238.5, 21.6) and 17 dielectric pipes as local scatterers. The scenario is probed in both cases with and without the target,
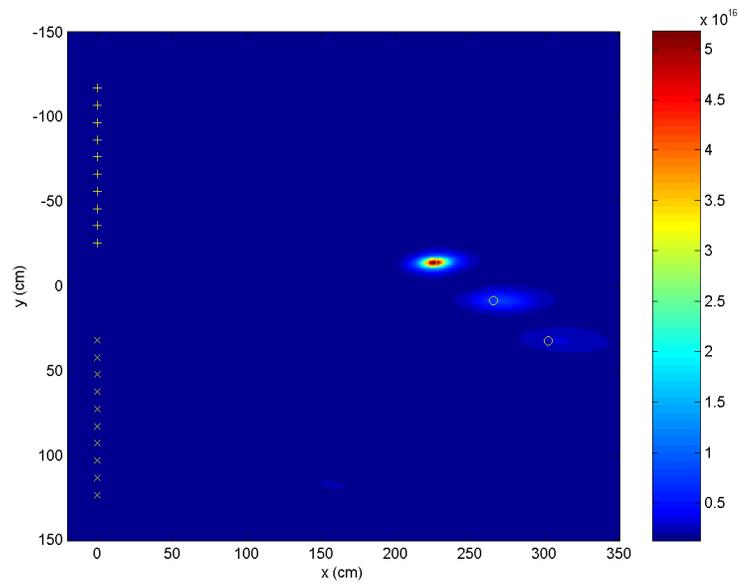
FIG. 11. *multiple-frequency likelihood image using measurements at 4 G Hz, 4.5 G Hz, 5 G Hz, 5.5 G Hz, and 6 G Hz.*
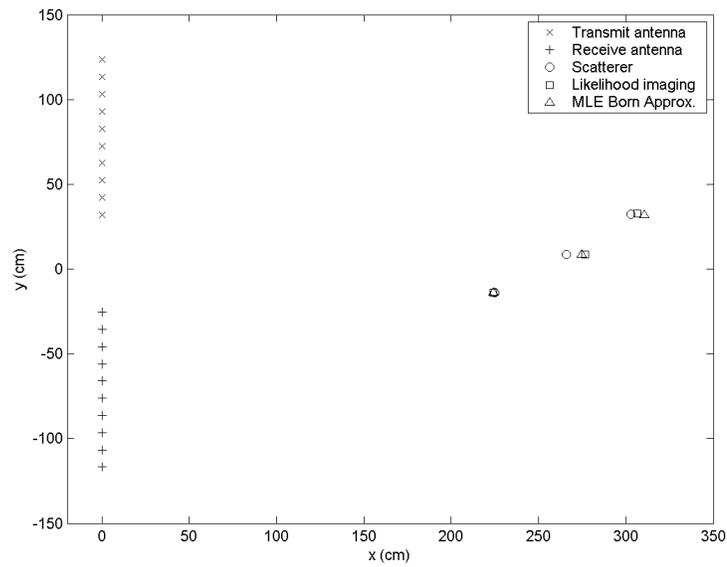
FIG. 12. *scatterer location estimations using likelihood imaging and maximum likelihood estimation. x: transmit antenna, +: receive antenna, ○: scatterer, □: likelihood imaging estimation, △: maximum likelihood estimation.*
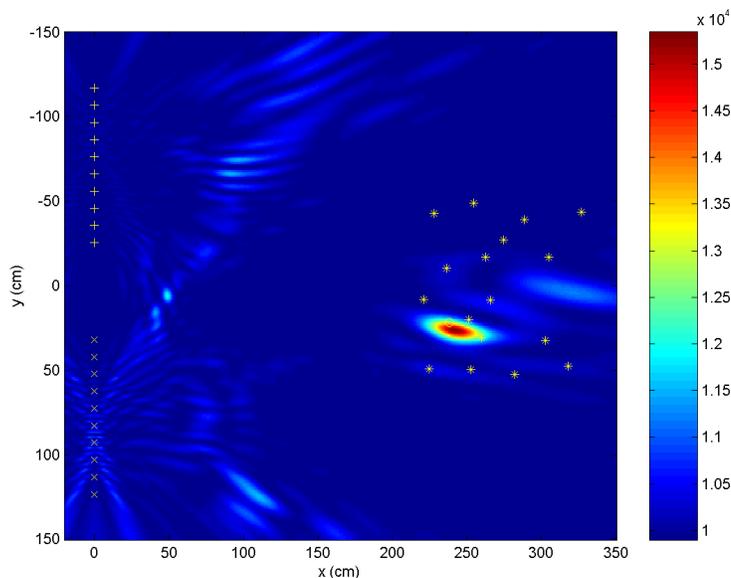
FIG. 13. *target detection by fixed-frequency likelihood image using measurement at 5 G Hz. ∗: dielectric scatterers, ◦: copper target.*

and the difference of the two measurements is used for the likelihood time-reversal imaging. The image created by the measurements at 5 G Hz is shown in Figure 13, and we can see that the image, which represents the approximated likelihood of the existence of a target at the imaging positions, has a peak around the target location, thus successfully detects the target. In Figure 14, we use five frequencies at 4 G Hz, 4.5 G Hz, 5 G Hz, 5.5 G Hz, and 6 G Hz to create the multiple-frequency likelihood image which has a significant enhanced imaging performance compared the the fixed frequency one in Figure 13.

7. Conclusions. In this paper, we presented a statistical framework for the fixed-frequency computational time-reversal imaging problem assuming point scatterers in a known background medium. Our statistical measurement models were based on the physical models of the multistatic response matrix, the distorted wave Born approximation and Foldy-Lax multiple scattering models. We developed maximum likelihood (ML) estimators of the locations and reflection parameters of the scatterers. Using a simplified single-scatterer model, we also proposed the likelihood time-reversal imaging (scanning) technique which is different than the current back-propagated field or "pseudo-spectrum" imaging strategies. We generalized the proposed fixed-frequency likelihood imaging to multiple frequencies, and demonstrated that it is effective in resolving the grating lobes of a sparse array. This make it possible to achieve high resolution by deploying a large-aperture array consisting of a small
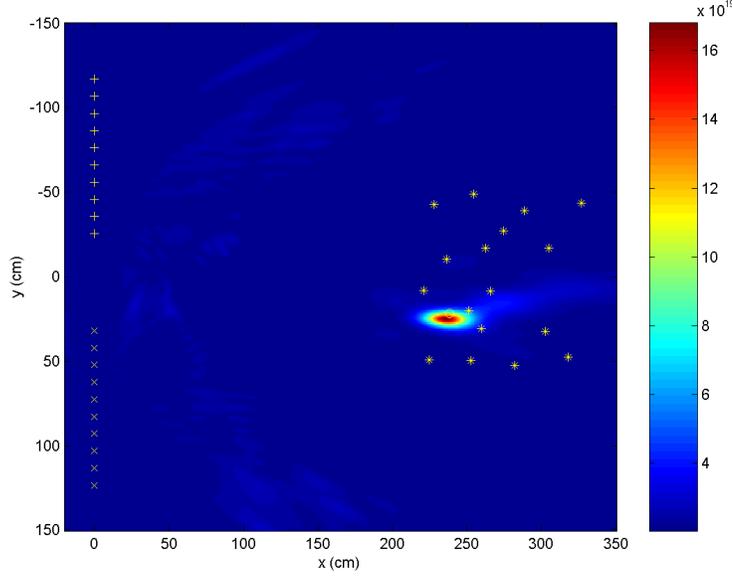
FIG. 14. *target detection by multiple-frequency likelihood image using measurements at 4 G Hz, 4.5 G Hz, 5 G Hz, 5.5 G Hz, and 6 G Hz. $*$: dielectric scatterers, $\circ$: copper target.*

number of antennas while avoiding spatial ambiguity.

In future work, we will evaluate the estimation performance by computing the Cramér-Rao lower bounds and derive statistical detectors for a target in the presence of scatterers. We will also extend the results of additive spatially white Gaussian noise to a spatially correlated case and consider the random scattering.

Appendix A.

LEMMA 1. *Let A and B be complex matrices with the same dimensions and $B = B^T$, then*

$$(A\text{-}1) \qquad \|A - B\|_{\mathrm{F}}^2 = \|\tilde{A} - B\|_{\mathrm{F}}^2 + \|A - \tilde{A}\|_{\mathrm{F}}^2,$$

where $\tilde{A} = (A + A^T)/2$ is the symmetric part of A, and $\|\cdot\|_{\mathrm{F}}$ represents the Frobenius norm of a matrix.

Proof. Denotes the Frobenius inner product $\langle A, B \rangle_{\mathrm{F}} = \operatorname{tr} A^H B$ and Frobenius norm $\|A\|_{\mathrm{F}}^2 = \langle A, A \rangle_{\mathrm{F}}$, where "tr" represents the trace of matrix, then

$$
\begin{aligned}
\|A - B\|_{\mathrm{F}}^2 &= \langle A - B, A - B \rangle_{\mathrm{F}} \\
&= \langle A - \tilde{A} + \tilde{A} - B, A - \tilde{A} + \tilde{A} - B \rangle_{\mathrm{F}} \\
&= \langle A - \tilde{A}, A - \tilde{A} \rangle_{\mathrm{F}} + \langle \tilde{A} - B, \tilde{A} - B \rangle_{\mathrm{F}} \\
&\quad + \langle A - \tilde{A}, \tilde{A} - B \rangle_{\mathrm{F}} + \langle \tilde{A} - B, A - \tilde{A} \rangle_{\mathrm{F}}.
\end{aligned}
$$

(A-2)

Note that

$$
\begin{aligned}
\langle A - \tilde{A}, \tilde{A} - B \rangle_{\mathrm{F}} &= \operatorname{tr}(A - \tilde{A})^H (\tilde{A} - B) \\
&= -\operatorname{tr}(A - \tilde{A})^* (\tilde{A} - B)^T \\
&= -\operatorname{tr}[(\tilde{A} - B)(A - \tilde{A})^H]^T \\
&= -\operatorname{tr}(\tilde{A} - B)(A - \tilde{A})^H \\
&= -\operatorname{tr}(A - \tilde{A})^H (\tilde{A} - B) \\
&= -\langle A - \tilde{A}, \tilde{A} - B \rangle_{\mathrm{F}},
\end{aligned}
$$

where the second identity follows from the fact that $\tilde{A} - B$ is symmetrical $(\tilde{A} - B)^T = \tilde{A} - B$ and $A - \tilde{A}$ is antisymmetrical $(A - \tilde{A})^T = -(A - \tilde{A})$, and the fourth identity from $\operatorname{tr} A = \operatorname{tr} A^T$. We conclude $\langle A - \tilde{A}, \tilde{A} - B \rangle_{\mathrm{F}} = 0$, similarly $\langle \tilde{A} - B, A - \tilde{A} \rangle_{\mathrm{F}} = 0$, then (A-1) is proved. $\qquad\square$

Appendix B.

LEMMA 2. *Let $P_A = A[A^H A]^{-1} A^H$ is the orthogonal projector onto the range of matrix A, A and B are of full column rank, then*

(B-1) $$ P_{A \otimes B} = P_A \otimes P_B. $$

Proof.

$$
\begin{aligned}
P_{A \otimes B} &= (A \otimes B)[(A \otimes B)^H (A \otimes B)]^{-1} (A \otimes B)^H \\
&= (A \otimes B)[(A^H \otimes B^H)(A \otimes B)]^{-1} (A^H \otimes B^H) \\
&= (A \otimes B)[A^H B \otimes A^H B]^{-1} (A^H \otimes B^H) \\
&= (A \otimes B)[(A^H B)^{-1} \otimes (A^H B)^{-1}](A^H \otimes B^H) \\
&= A(A^H A)^{-1} A^H \otimes B(B^H B)^{-1} B^H \\
&= P_A \otimes P_B,
\end{aligned}
$$

where we use $(A \otimes B)^H = A^H \otimes B^H$ in the second equality, $(A \otimes B)(D \otimes G) = AD \otimes BG$ in the third and fifth equalities, and $(N \otimes M)^{-1} = N^{-1} \otimes M^{-1}$ in the fourth identity.

$\qquad\square$

COROLLARY. $P_{A \otimes B}^{\perp} = (P_A \otimes P_B)^{\perp}$.

Proof.

$$
P_{A \otimes B}^{\perp} = I - P_{A \otimes B} = I - P_A \otimes P_B = (P_A \otimes P_B)^{\perp}.
$$

$\qquad\square$

Appendix C. In this appendix, we derive the concentrated likelihood function with respect to $\boldsymbol{x}$ based on statistical model (50). By the independence between Y_l and $Y_{l'}$ for $l \neq l'$ the joint likelihood function could be found as

$$l_9(\boldsymbol{x}, \tau_1(\omega_1), \tau_1(\omega_2), \ldots, \tau_1(\omega_L), \sigma_1^2, \sigma_2^2, \ldots, \sigma_L^2; Y_1, Y_2 \cdots Y_L)$$

(C-1) $$= \prod_{l=1}^{L} \frac{1}{(\pi \sigma_l^2)^{N_{\mathrm{r}} N_{\mathrm{t}}}} \exp(-\|Y_l - \tau_1(\omega_l) \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}, \omega_l) \boldsymbol{g}_{\mathrm{t}}^T(\boldsymbol{x}, \omega_l)\|_{\mathrm{F}}^2 / \sigma_l^2),$$

where $\boldsymbol{x} = \boldsymbol{x}_1$. Following steps similar to those in Section 3, we first solve

(C-2) $$\frac{\partial}{\partial \sigma_l^2} l_9(\boldsymbol{x}, \tau_1(\omega_1), \tau_1(\omega_2), \ldots, \tau_1(\omega_L), \sigma_1^2, \sigma_2^2, \ldots, \sigma_L^2; Y_1, Y_2 \cdots Y_L) = 0,$$
$$l = 1, 2, \ldots, L,$$

and substitute the solutions

(C-3) $$\sigma_l^2 = \frac{1}{N_{\mathrm{r}} N_{\mathrm{t}}} \|Y_l - \tau_1(\omega_l) \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}, \omega_l) \boldsymbol{g}_{\mathrm{t}}^T(\boldsymbol{x}, \omega_l)\|_{\mathrm{F}}^2$$

into (C-1). Then, we have the concentrated likelihood function

$$l_{10}(\boldsymbol{x}, \tau_1(\omega_1), \tau_1(\omega_2), \ldots, \tau_1(\omega_L); Y_1, Y_2 \cdots Y_L)$$

(C-4) $$= \frac{(N_{\mathrm{r}} N_{\mathrm{t}})^{L N_{\mathrm{r}} N_{\mathrm{t}}} \exp(-L N_{\mathrm{r}} N_{\mathrm{t}})}{\pi^{L N_{\mathrm{r}} N_{\mathrm{t}}} \prod_{l=1}^{L} \|Y_l - \tau_1(\omega_l) \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}, \omega_l) \boldsymbol{g}_{\mathrm{t}}^T(\boldsymbol{x}, \omega_l)\|_{\mathrm{F}}^{2 N_{\mathrm{r}} N_{\mathrm{t}}}},$$

which could be further concentrated with respect to $\boldsymbol{x}$ getting (51)

$$l_8(\boldsymbol{x}; Y_1, Y_2 \cdots Y_L) = \frac{(N_{\mathrm{r}} N_{\mathrm{t}})^{L N_{\mathrm{r}} N_{\mathrm{t}}} \exp(-L N_{\mathrm{r}} N_{\mathrm{t}})}{\pi^{L N_{\mathrm{r}} N_{\mathrm{t}}} \prod_{l=1}^{L} \|P_{\boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}, \omega_l) \otimes \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}, \omega_l)}^{\perp} \mathrm{vec}(Y_l)\|_{\mathrm{F}}^{2 N_{\mathrm{r}} N_{\mathrm{t}}}}$$

using the least-squares solutions

(C-4) $$\hat{\tau}_1(\omega_l) = \frac{\boldsymbol{g}_{\mathrm{t}}^H(\boldsymbol{x}, \omega_l) \otimes \boldsymbol{g}_{\mathrm{r}}^H(\boldsymbol{x}, \omega_l)}{[\boldsymbol{g}_{\mathrm{t}}^H(\boldsymbol{x}, \omega_l) \boldsymbol{g}_{\mathrm{t}}(\boldsymbol{x}, \omega_l)][\boldsymbol{g}_{\mathrm{r}}^H(\boldsymbol{x}, \omega_l) \boldsymbol{g}_{\mathrm{r}}(\boldsymbol{x}, \omega_l)]} \mathrm{vec}(Y_l), \quad l = 1, 2, \ldots, L.$$

REFERENCES

[1] B. Y. ZEL'DOVICH, N. F. PILIPETSKY, AND V. V. SHKUNOV, *Principles of Phase Conjugation* Springer-Verlag, Berlin, 1985.
[2] M. FINK, *Time-reversed acoustics*, Scientific American, Nov. 1999.
[3] L. BORCEA, C. TSOGKA, G. PAPANICOLAOU, AND J. BERRYMAN, *Imaging and time reversal in random media*, Inverse Problems, 18(2002), pp. 1247-1279.
[4] M. BORN AND E. WOLF, *Principles of Optics*, Academic Press, New York, 1970.
[5] W. A. KUPERMAN, W. S. HODGKISS, H. C. SONG, T. AKAL, C. FERLA, AND D. R. JACKSON, *Phase conjugation in the ocean: experimental demonstration of a time reversal mirror*, J. Acoust. Soc. Am., 103(1998), pp. 25-40.

[6] L. BORCEA, G. PAPANICOLAOU, AND C. TSOGKA, *Theory and applications of time reversal and interferometric imaging*, Inverse Problems, 19(2003), pp. 5139-5164.

[7] A. DERODE, P. ROUX, AND M. FINK, *Robust acoustic time reversal with high-order multiple scattering*, Phys. Rev. Let., 75:23(1995), pp. 4206-4210.

[8] M. FINK AND C. PRADA, *Acoustic time-reversal mirrors, topical review*, Inverse Problems, 17:R1-R38, 2001.

[9] H. C. SONG, W. A. KUPERMAN, AND W. S. HODGKISS, *Iterative time reversal in the ocean*, J. Acoust. Soc. Am., 105(1999), pp. 3176-3184.

[10] G. BAL, G. C. PAPANICOLAOU, AND L. RYZHIK, *Self-averaging in time reversal for the parabolic wave equation*, Stochastics and Dynamics, 2(2002), pp. 507-531.

[11] D. DOWLING AND D. JACKSON, *Phase conjugation in underwater acoustics*, J. Acoust. Soc. Am., 89(1990), pp. 171-181.

[12] M. FINK AND J. DE ROSNY, *Time-reversed acoustics in random media and in chaotic cavities*, Nonlinearity, 15(2002), R1-R18.

[13] M. FINK, D. CASSEREAU, A. DERODE, C. PRADA, P. ROUX, M. TANTER, J. L. THOMAS, AND F. WU, *Time-reversed acoustics*, Rep. Progr. Phys., 63(2000), pp. 1933-1995.

[14] W. HODGKISS, H. SONG, W. KUPERMAN, T. AKAL, C. FERLA, AND D. JACKSON, *A long-range and variable focus phase-conjugation experiment in a shallow water*, J. Acoust. Soc. Am., 105(1999), pp. 1597-1604.

[15] W. KUPERMAN, W. HODGKISS, H. SONG, T. AKAL, C. FERLA, AND D. JACKSON, *Phase-conjugation in the ocean*, J. Acoust. Soc. Am., 102(1997), pp. 1-16.

[16] J. L. THOMAS AND M. FINK, *Ultrasonic beam focusing through tissue inhomogeneities with a time reversal mirror: application to transskull therapy*, IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control, 43(1996), pp. 1122-1129.

[17] G. PAPANICOLAOU, L. RYZHIK, AND K. SOLNA, *Statistical stability in time reversal*, SIAM J. Applied Mathematics, 64:4(2004), pp. 1133-1155.

[18] P. BLOMGREN, G. PAPANICOLAOU, AND H. ZHAO, *Super-resolution in time-reversal acoustics*, J. Acoust. Soc. Am., 111(2002), pp. 228-248.

[19] D. R. DOWLING AND D. R. JACKSON, *Narrow band performance of phase conjugate arrays in dynamic random media*, J. Acoust. Soc. Am., 91(1992), pp. 3257-3277.

[20] A. KIM, P. BLOMGREN, AND G. PAPANICOLAOU, *Spatial focusing and intersymbol interference in time reversal communications*, submitted for publication 2003.

[21] S. M. EMAMI, J. HANSEN, A. D. KIM, G. PAPANICOLAOU, A. J. PAULRAJ, D. CHEUNG, AND C. PRETTIE, *Predicted time reversal performance in wireless communications using channel measurements*, submitted to IEEE Comm. Let. 2004.

[22] P. KYRITSI, G. PAPANICOLAOU, P. EGGERS, AND A. OPERA, *MISO time reversal and delay spread compression for FWA channels at 5GHz*, Antennas and Wireless Propagation Letters, No. 3, pp. 96-99, 2004.

[23] R. RANGARAJAN AND A. O. HERO III, *Analysis of a Multistatic Adaptive Target Illumination and Detection Approach (MATILDA) to Time Reversal Imaging*, MA "Hot Topics" Workshop: Adaptive Sensing and Multimode Data Inversion, June 27-30, 2004.

[24] R. RANGARAJAN, N. PATWARI, AND A. O. HERO III, *Analysis of a Multistatic Adaptive Target Illumination and Detection Approach (MATILDA) to Time Reversal Imaging*, submitted for publication.

[25] J. H. TAYLOR, *Scattering Theory*, Wiley, New York, 1972.

[26] R. K. SNIEDER AND J. A. SCALES, *Time-reversed imaging as a diagnostic of wave and particle chaos*, Physical Review E, 58(1998), pp. 5668-5675.

[27] F. K. GRUBER, E. A. MARENGO, AND A. J. DEVANEY, *Time-reversal imaging with multiple signal classification considering multiple scattering between the targets*, J. Acoust. Soc. Am., 115(2004), pp. 3042-3047.

[28] A. J. Devaney, *Time reversal imaging of obscured targets from multistatic data*, submitted to IEEE Trans. Ant. Prop.

[29] C. Prada, J. L. Thomas, and M. Fink, *The iterative time reversal process: Analysis of the convergence*, J. Acoust. Soc. Am., 97(1995), pp. 62-71.

[30] C. Prada, S. Manneville, D. Spoliansky, and M. Fink, "Decomposition of the time reversal operator: Detection and selective focusing on two scatterers," *J. Acoust. Soc. Am.* 99, pp. 2067-2076, 1996.

[31] A. J. Devaney, *Super-resolution processing of multi-static data using time reversal and MUSIC*, Journal of the Acoustical Society of America, to appear.

[32] I. Miller and M. Miller, *John E. Freund's Mathematical Statistics, Sixth Edition*, NJ: Prentice Hall, 1999.

[33] C. Tai, *Dyadic Green Functions in Electromagnetic Theory*, 2nd ed. IEEE Press, New York, 1993.

[34] D. W. Bliss, K. W. Forsythe, *Multiple-input multiple-output (MIMO) radar and imaging: degrees of freedom and resolution*, The Thrity-Seventh Asilomar Conf. on Signals, Systems & Computers, Nov. 2003, Vol. 1, pp. 54-59.

[35] A. Paulraj and C. B. Papadias, *Space-time processing for wireless communications*, IEEE Signal Processing Magazine, 14(1997), pp. 49-83.

[36] H. Lev-Ari, *Theory and application of computational time reversal progress report*, DAPAR review, Oct. 2004.

[37] H. Akaike, *Information theory as an extension of the maximum likelihood principle*, Second Intl. Symposium on Information Theory, eds. V. N. Petrov and F. Csaki, Budapest: Akailseoniai-Kiudo, pp. 267-281.

[38] G. Schwarz, *Estimating the Dimension of a Model*, The Annals of Statistics, 6(1978), pp. 461-464.

[39] J. W. Brewer, *Kronecker Products and Matrix Calculus in System Theory*, IEEE Transactions on Circuits and Systems, cas-25:9(1978), pp. 772-781.

[40] C. G. Broyden, *The convergence of a class of double-rank minimization algorithms*, J. Inst. Math. Applic., 6(1970), pp. 76-90.

[41] R. Fletcher, *A new approach to variable metric algorithms*, Computer Journal, 13(1970), pp. 317-322.

[42] D. Goldfarb, *A family of variable metric updates derived by variational means*, Mathematics of Computing, 24(1970), pp. 23-26.

[43] D. F. Shanno, *Conditioning of quasi-Newton methods for function minimization*, Mathematics of Computing, 24(1970), pp. 647-656.

[44] A. Dogandzic and A. Nehorai, *Estimating evoked dipole responses in unknown spatially correlated noise with EEG/MEG arrays*, IEEE Trans. on Signal Processing, SP-48(2000), pp. 13-25.

[45] H. L. Van Trees, *Detection, Estimation, and Linear Modulation Theory.—pt. 3. Radar-sonar Signal Processing and Gaussian Signals in Noise.—pt. 4. Optimum Array Processing*, New York, NY: Willey, 2001.

[46] T. Do-Hong and P. Russer, *Signal processing for wideband smart antenna array applications*, IEEE Microwave Magazine, 5(2004), pp. 57-67.

[47] M. Wax, T. Shan, and T. Kailath, *Spatio-temporal spectral analysis by eigenstructure methods*, IEEE Trans. Acoust. Speech, Signal Processing, 32(1984), pp. 817-827.

[48] G. Su and M. Morf, *Signal subspace approach for multiple wideband emitter location*, IEEE Trans. Acoust. Speech, Signal Processing, 31(1983), pp. 1502-1522.

[49] H. Wang and M. Kaveh, *Coherent signal location and signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources*, IEEE Trans. Acoust. Speech, Signal Processing, 33(1985), pp. 823-831.

[50] H. Krim and M. Viberg, *Two decades of array signal processing research, the parametric approach*, IEEE Signal Processing Magazine, July 1996.

[51] M. B. Priestley, *Spectral Analysis and Time Series*, London; New York: Academic Press, 1981.