Title

Recovery of spatially varying acoustical properties via automated partial differential equation identification.

Journal

The Journal of the Acoustical Society of America, 153(6)

ISSN

0001-4966

Authors

Liu, Ruixian
Gerstoft, Peter
Bianco, Michael J
et al.

Publication Date

2023-06-01

DOI

10.1121/10.0019592

JUNE 02 2023

Recovery of spatially varying acoustical properties via automated partial differential equation identification

Ruixian Liu ; Peter Gerstoft ; Michael J Bianco ; Bhaskar D. Rao

J Acoust Soc Am 153, 3169 (2023)
https://doi.org/10.1121/10.0019592

JASA ARTICLE

Recovery of spatially varying acoustical properties via automated partial differential equation identification

Ruixian Liu,[1,a] Peter Gerstoft,[2] Michael J Bianco,[2] and Bhaskar D. Rao[1]
[1]Department of Electrical and Computer Engineering, University of California, San Diego, California 92161, USA
[2]Scripps Institution of Oceanography, University of California, San Diego, California 92037, USA

ABSTRACT:
Observable dynamics, such as waves propagating on a surface, are generally governed by partial differential equations (PDEs), which are determined by the physical properties of the propagation media. The spatial variations of these properties lead to spatially dependent PDEs. It is useful in many fields to recover the variations from the observations of dynamical behaviors on the material. A method is proposed to form a map of the physical properties' spatial variations for a material via data-driven spatially dependent PDE identification and applied to recover acoustical properties (viscosity, attenuation, and phase speeds) for propagating waves. The proposed data-driven PDE identification scheme is based on ℓ_1-norm minimization. It does not require any PDE term that is assumed active from the prior knowledge and is the first approach that is capable of identifying spatially dependent PDEs from measurements of phenomena. In addition, the method is efficient as a result of its non-iterative nature and can be robust against noise if used with an integration transformation technique. It is demonstrated in multiple experimental settings, including real laser measurements of a vibrating aluminum plate. Codes and data are available online at https://tinyurl.com/4wza8vxs. https://doi.org/10.1121/10.0019592
(Received 21 November 2022; revised 16 April 2023; accepted 11 May 2023; published online 2 June 2023)
[Editor: Efren Fernandez-Grande] Pages: 3169–13

I. INTRODUCTION

Natural phenomena are, in general, caused by partial differential equations (PDEs), where the PDE coefficients are derived from medium properties. For example, the observed waves propagating on a plate are governed by wave equations with the PDE coefficients determined by phase speeds and attenuation factors, which are further decided by the elastic properties and density of the plate's material.

The focus of this paper is to solve the inverse problem, i.e., we invert for the active PDE terms from the observations, determine the coefficients of each active PDE term, and then use the coefficients to recover physical properties. We highlight the use case of the proposed method in recovering various properties for the medium of propagating waves in acoustical scenarios and, thus, the relative PDEs, like the wave equation and Burgers's equation, are used as examples.

This area has been an active focus of applied mathematics research[1–10] with a few applications.[11–13] Considering the abundance of sensor-collected measurements and wide range of use cases of material properties recovery (e.g., materials diagnostics and fatigue detection), this inversion technique would be broadly applicable.

Superior to previous data-driven PDE identification approaches,[1–13] our method can recognize spatially dependent PDEs and, hence, recover one-dimensional (1D) or two-dimensional (2D) maps for the spatial variations of material properties from measurements of the phenomena. Unlike classic spatially dependent parameter estimation or tomography methods which require the PDE form known *a priori* and employ the domain knowledge only pertinent to such a particular PDE,[14–20] the proposed method can identify multiple kinds of unknown PDEs using a same formulation with no PDE-dependent knowledge and, thus, is widely applicable and can recover various properties in more scenarios with fewer assumptions.

We start with the spatially 1D case in the theory and include 2D examples for experiments. Consider a physical system, $U(x,t)$, that describes the spatiotemporal dynamics. Within the system $U(x,t)$, suppose we have N_x evenly spatially distributed sensors collecting measurements (e.g., displacement, pressure, etc.) at M_t evenly separated time steps and, therefore, we obtain measurements, $\mathbf{U} \in \mathbb{R}^{N_x \times M_t}$, at discrete spatial-temporal coordinates of $U(x,t)$. The system properties are recovered by identifying its governing PDE,

$$\mathsf{N}[U] = 0, \tag{1}$$

from the measurements. Here, $\mathsf{N}[U]$ is a linear combination of PDE terms involving partial derivatives of U, e.g., $\mathsf{N}[U]$ describing the 1D wave propagation at speed c with the attenuation factor α is

$$\mathsf{N}[U] = U_{tt} + \alpha U_t - c^2 U_{xx}, \tag{2}$$

where U_t, U_{tt} are the first- and second-order temporal derivative, respectively, and U_{xx} is the second-order spatial derivative of U. Recently, there are many developments focusing

[a]Electronic mail: rul188@ucsd.edu

on identifying PDEs directly from observed data.[1–13] However, they have two limitations: (i) the need for prior knowledge of the active PDE terms and (ii) the inability to recover spatially dependent parameters.

Regarding (i), most previous developments require one or more active PDE terms to be known *a priori* (e.g., the first-order time derivative U_t,[1–3,5–7] the second-order time derivative U_{tt},[13] one term with predefined order,[8] or multiple PDE terms[4]). Then, they derive other contributing PDE terms and their coefficients and, thus, only parts of the PDE are inferred from data. This is problematic when the knowledge of which term should be assumed active is uncertain or unknown, e.g., to identify the governing PDE for a wave which may either be an inviscid Burgers's equation ($U_t + UU_x = 0$) or a non-attenuating wave equation ($U_{tt} - c^2\nabla^2 U = 0$) with no PDE terms in common, one must specify the correct active term from more prior information. The few methods that do not require active term assumption use a sparse Bayesian learning (SBL) based approach[9] or a cross-validation (CV) based method.[10] They both iteratively assume one active term from a library of terms, identify the PDE for each assumption using SBL[9] or sparsity penalized CV,[10] and, finally, select the best assumption by comparing the posterior confidence[9] or minimal fitting error.[10] They are time-consuming as the identification process is repeated for every assumption.

Regarding (ii), in reality, the PDEs governing the observed system can have spatially dependent coefficients, e.g., the coefficient for U_{xx} in Eq. (2) can vary across space, which indicates the phase speed, c, is spatially dependent. The spatially dependent coefficients are due to the spatial variation of the materials and, thus, the recovery of spatially dependent PDEs can unveil the spatial properties of the underlying materials. The above methods, however, can only identify PDEs that are constant across space. The current spatially dependent coefficients recovery schemes are limited to a few specific PDEs,[15,16] and they cannot be used for PDE identification as they require the type of PDE to be known. No methods can identify unknown PDEs which are potentially spatially dependent.

Suppose there is no information about the spatial variation of the PDEs available and, therefore, we must identify the PDE for every location. The challenges in tackling this task include: (a) fewer measurements available for one location comparing to the whole field and (b) longer central processing unit (CPU) time because the process is repeated for all of the locations. Thus, a viable method should be robust in selecting the correct PDE from limited measurements and also computationally efficient.

To address these limitations, we propose an ℓ_1-norm minimization based data-driven method that identifies unknown PDEs for every spatial location without any assumption of the assumed active PDE term. No information about the spatial variation of the PDEs is assumed. An auxiliary vector is introduced to robustify the identification from limited measurements and enable the recovery of all of the active PDE terms without iterative assumptions. The method is applied pixel-wise and recovers the spatial variation of the PDEs to the highest resolution. For noisy measurements, we extend the integration transformation approach[7] to our spatially dependent PDE identification scheme to make it more robust against noise.

This paper is extended from the work[21] to integrate the noise resistance technique and perform extensive experimental validations, including the validations using data collected from real physical settings. It is organized as follows: Sec. II A presents the spatially dependent PDE identification method, and Sec. II B briefly introduces the integration transformation approach to alleviate the impact of noise. Section III shows synthetic experimental results, including an experiment for noise-robustness in Sec. III C. The experiments for a real vibrating aluminum plate using clean and noisy measurements are described in Sec. IV. In Sec. V, we emphasize the efficiency of the proposed approach by comparing the required CPU times to identify the PDEs for three datasets using the proposed method, the methods in Refs. 9 and 10, and exhaustive search. Finally, we conclude our work in Sec. VI.

The notation is as follows:

(a) For the measurements $\mathbf{U} \in \mathbb{R}^{N_x \times M_t}$, $\mathbf{U}(i_x, i_t)$ is $\mathbf{U}$ sampled at the coordinate $(i_x\Delta x, i_t\Delta t)$, where $0 \le i_x \le N_x - 1$, $0 \le i_t \le M_t - 1$, and Δx, Δt are sampling intervals;
(b) sets I_x and I_t contain N spatial and M temporal coordinates within the region of interest (ROI), respectively. Use $n \in [1, N]$ as the index of the elements in I_x. The temporal coordinates in I_t are indexed by $m \in [1, M]$;
(c) for any matrix $\mathbf{A}$ other than $\mathbf{U}$, the entry at its ith row and jth column, where $i \ge 1$, $j \ge 1$, is denoted by $\mathbf{A}(i, j)$. $\mathbf{A}(i, :)$ denotes the ith row, and $\mathbf{A}(:, j)$ denotes the jth column;
(d) the subscript/superscript of a matrix denotes the properties of the whole matrix and *not* its entries. For example, $\mathbf{A}_n$ can indicate that all of the entries of $\mathbf{A}$ are computed from the measurements at location n, and $\mathbf{A}_n(i, j)$ is for the entry at the ith row and jth column of $\mathbf{A}_n$;
(e) for vector $\mathbf{a}$, its ith entry ($i \ge 1$) is denoted by either $\mathbf{a}(i)$ or a_i; and
(f) the variable superscripted by "$\hat{\ }$" denotes its estimation.

II. THEORY

In this section, we propose the method to efficiently identify spatially dependent PDEs in Sec. II A and then introduce an integration transformation technique that increases the approach's robustness against measurement noise in Sec. II B.

A. PDE identification

First, we formulate the PDE identification problem mathematically in Sec. II A 1 and then solve it in Sec. II A 2.

1. Problem formulation

We select N spatial locations with M time steps for each location from the measurements $\mathbf{U} \in \mathbb{R}^{N_x \times M_t}$ as the ROI. In this work, suppose we know the range of PDEs governing the dynamics in the ROI and to be specific, they may be the (attenuating) wave equation, the (viscous) Burgers's equation with a nonlinear term UU_x (product of U and U_x), and

https://doi.org/10.1121/10.0019592

the sine-Gordon equation with a nonderivative term $\sin(U)$.[22,23] They can model various fluid dynamics.

For the nth spatial location within the ROI, we use $\mathbf{u} \in \mathbb{R}^M$ to denote its measurements at all of the M time steps, i.e., from $\mathbf{U}(I_x(n), I_t(1))$ to $\mathbf{U}(I_x(n), I_t(M))$. We build a dictionary, $\mathbf{\Phi}_n$, containing all of the $D = 6$ PDE terms potentially appearing in the PDEs mentioned above such that

$$\mathbf{\Phi}_n = [\mathbf{u}_t,\ \mathbf{u}_{tt},\ \mathbf{u} \circ \mathbf{u}_x,\ \mathbf{u}_{xx},\ \mathbf{u}_{tx},\ \sin(\mathbf{u})] \in \mathbb{R}^{M \times D}, \quad (3)$$

where each term is an M-length vector evaluated at all of the M time steps, and the derivatives are computed numerically by finite difference,[38] e.g., the mth entry of $\mathbf{u}_t$ is $[\mathbf{U}(I_x(n), I_t(m)+1) - \mathbf{U}(I_x(n), I_t(m)-1)]/(2\Delta t)$. The "$\circ$" denotes element-wise production, e.g., the mth entry of $\mathbf{u} \circ \mathbf{u}_x$ is $\mathbf{U}(I_x(n), I_t(m)) \times \{[\mathbf{U}(I_x(n)+1, I_t(m)) - \mathbf{U}(I_x(n)-1, I_t(m))]/(2\Delta x)\}$. Measurements outside of the ROI boundaries should exist such that the spatial derivatives for $n = 1$ or N and the temporal derivatives for $m = 1$ or M can be computed by finite difference, which requires $N < N_x - 1$ and $M < M_t - 1$.

Initially, we can treat the problem as recovering the coefficient $\mathbf{a}_n = [\mathbf{a}_n(1) \cdots \mathbf{a}_n(D)]^{\mathrm{T}} \in \mathbb{R}^D$ such that

$$\mathbf{\Phi}_n \mathbf{a}_n \approx \mathbf{0}, \quad ||\mathbf{a}_n||_1 > 0, \quad (4)$$

where $||\mathbf{a}_n||_1 > 0$ is to avoid the trivial solution $\mathbf{a}_n = \mathbf{0}$, and the approximation is due to the assumption $M > D$ and noise in the measurements or generated by numerical differentiation.

An exhaustive search can find the set of active PDE terms by iterating through all of the combinations of atoms in the dictionary and minimizing the fitting error with the constraint that not too many columns of $\mathbf{\Phi}_n$ are chosen (sparsity constraint). The cardinality of all of the possible sets ranges from two to D. For each hypothetical set, assume that the coefficient for one term in $\mathbf{a}_n$ is one and fit the other terms by least squares regression.

Given an assumed active term with coefficient one, the difference between $\mathbf{\Phi}_n \mathbf{a}_n$ and $\mathbf{0}$, i.e., $||\mathbf{\Phi}_n \mathbf{a}_n||_2$, is monotonically nonincreasing as more columns of $\mathbf{\Phi}_n$ are chosen and decreasing for most cases. Using all columns of $\mathbf{\Phi}_n$ will minimize the difference, but it indicates that the PDE has all of the D active terms, which is typically not true. Thus, a sparsity constraint should be introduced to avoid selecting all of the terms in $\mathbf{\Phi}_n$ as active terms. One way to impose this constraint is to minimize an augmented loss function, which is the sum of $||\mathbf{\Phi}_n \mathbf{a}_n||_2$ and a penalty proportional to $||\mathbf{a}_n||_0$.[10]

This process requires combinatorial complexity because we cannot directly select all of the columns as discussed above and need to explore all of the possible sets of selected columns with cardinality ranging from two to D. For a dictionary with only a few columns, such as in Eq. (3), this is feasible, but for larger dictionaries, the CPU time increases quickly. For example, for Eq. (3), there are 57 kinds of sets in total, whereas for Eq. (24) with 9 terms, there are 502 cases. A demonstration of such a transition is included in Sec. V.

2. Solving PDE coefficients by ℓ_1-norm minimization

Instead of an exhaustive search, we introduce a normalization matrix, $\mathbf{W}_n \in \mathbb{R}^{D \times D}$, which is diagonal with $\mathbf{W}_n(i,i) = ||\mathbf{\Phi}_n(:,i)||_2$, and by finding the coefficients, $\bar{\mathbf{a}}_n = [\bar{\mathbf{a}}_n(1) \ldots \bar{\mathbf{a}}_n(D)]^{\mathrm{T}}$, that make the columns in the normalized dictionary, $\bar{\mathbf{\Phi}}_n = \mathbf{\Phi}_n \mathbf{W}_n^{-1}$, fit well under a fitting error, tol_n,

$$||\bar{\mathbf{\Phi}}_n \bar{\mathbf{a}}_n||_2^2 \le \mathrm{tol}_n \quad \text{such that } ||\bar{\mathbf{a}}_n||_1 = 1, \quad (5)$$

we acquire $\mathbf{a}_n = \mathbf{W}_n^{-1} \bar{\mathbf{a}}_n$.

We use the normalized dictionary such that the variation of magnitudes for columns in $\mathbf{\Phi}_n$ does not affect the column selection and, thus, the selection is only based on the dynamical characters (i.e., the variation of the entries within each column). In addition, we use $||\bar{\mathbf{a}}_n||_1 = 1$ in Eq. (5) instead of $||\bar{\mathbf{a}}_n||_1 > 0$ as in Eq. (4), otherwise, the magnitudes of the nonzero elements in $\bar{\mathbf{a}}_n$ can be arbitrarily small to encourage $\mathbf{\Phi}_n \mathbf{W}_n^{-1} \bar{\mathbf{a}}_n \approx \mathbf{0}$.

The limitation of Eq. (5) is that $||\bar{\mathbf{a}}_n||_1 = 1$ is not a convex set and, therefore, Eq. (5) is not solvable via efficient convex optimization tools. To make Eq. (5) amenable to convex optimization, we specify the positive/negative signs in $||\bar{\mathbf{a}}_n||_1$ and, thus, reduce $||\bar{\mathbf{a}}_n||_1 = 1$ to an affine constraint as detailed below.

We use physical information to reduce $||\bar{\mathbf{a}}_n||_1 = 1$ to $\sum_{i=1}^{D} s_i \bar{\mathbf{a}}_n(i) = 1$, where $s_i \in \{-1, 1\}$ is decided by the information from potential PDE forms, e.g., s_1 and s_2 are the same because coefficients for U_t and U_{tt} are of the same sign for all of the considered PDEs involving them. Although there are various PDEs being taken into consideration, the relations of the coefficient signs for the shared terms among different PDEs do not conflict, e.g., the signs for U_t and U_{xx} are always different for the viscous Burgers's equation and the attenuating wave equation (which indicates s_1 and s_4 are opposite). Setting $s_1 = 1$, we obtain the auxiliary vector, $\mathbf{s} = [s_1 \cdots s_D]^{\mathrm{T}} \in \mathbb{R}^D$, for the dictionary [Eq. (3)] that is consistent with all of the potential PDEs considered in this work as

$$\mathbf{s} = [1, 1, 1, -1, -1, 1]^{\mathrm{T}} \in \mathbb{R}^D, \quad (6)$$

and $||\bar{\mathbf{a}}_n||_1 = 1$ is reduced to $\mathbf{s}^{\mathrm{T}} \bar{\mathbf{a}}_n = 1$, which is affine.

The dictionary [Eq. (3)] includes the terms for all of the potential PDEs, and the true governing PDE only involves a few of them by experience. Therefore, it is preferable for the identified PDE to have fewer active terms under the data fitting constraint. Thus, we use ℓ_1-norm minimization due to its ability to promote sparsity:[24–30]

$$\bar{\mathbf{\Phi}}_n = \mathbf{\Phi}_n \mathbf{W}_n^{-1}, \quad \bar{\mathbf{\Phi}}_n^s = \begin{bmatrix} \bar{\mathbf{\Phi}}_n \\ \mathbf{s}^{\mathrm{T}} \end{bmatrix} \in \mathbb{R}^{(M+1) \times D}, \quad (7a)$$

$$\hat{\bar{\mathbf{a}}}_n = \arg\min ||\bar{\mathbf{a}}_n||_1 \text{ such that } ||\bar{\mathbf{\Phi}}_n^s \bar{\mathbf{a}}_n - \mathbf{e}||_\infty \le \tau_n, \quad (7b)$$

$$\hat{\mathbf{a}}_n = \mathbf{W}_n^{-1} \hat{\bar{\mathbf{a}}}_n, \quad (7c)$$

where $\mathbf{e} \in \mathbb{R}^{M+1}$ has all zero entries except $\mathbf{e}(M+1) = 1$ and τ_n is a predefined toleration. The constraint $||\bar{\mathbf{\Phi}}_n^s \bar{\mathbf{a}}_n$

$-\mathbf{e}||_\infty \le \tau_n$ in Eq. (7b) enforces $\bar{\mathbf{\Phi}}_n\bar{\mathbf{a}}_n \approx \mathbf{0}$ and $\mathbf{s}^{\mathrm{T}}\bar{\mathbf{a}}_n \approx 1$. The optimization problem in Eq. (7) enables identifying all of the active PDE terms simultaneously and, thus, is more efficient than the previous methods in Refs. 9 and 10, relying on iterations, which at first iteratively assume that $\mathbf{a}_n(i) = 1$ for all $1 \le i \le D$ to solve Eq. (4), and then select the best assumption among these D cases.

The $\mathbf{s}$ in Eq. (6) is an example used for the considered PDEs in this paper for the spatially 1D case. Fixing the six entries in Eq. (6), it can be extended by adding more ± 1 entries to accommodate more PDE terms in the dictionary, e.g., Eq. (24) for the 2D extension. $\mathbf{\Phi}_n$ and $\mathbf{s}$ can also be extended to include more terms and, therefore, encompass most PDEs for the dynamics of acoustical waves like the Korteweg-deVries (KdV) equation, Stokes's wave equation,[31] van Wijngaarden's equation,[31] etc. Because the sign relation in Eq. (6) is suitable for most acoustical PDEs, the $\mathbf{s}$ conveys less information in the potential kinds of PDEs compared to knowing at least one active term and, thus, represents a weaker prior knowledge.

We only put the various dynamical patterns of U, which are denoted by the PDE terms, into the dictionary, $\mathbf{\Phi}_n$, and leave the information of the magnitudes and signs for these terms appearing in the PDE into the coefficient $\hat{\mathbf{a}}_n$ for better physical interpretability. For example, for the wave equation, we prefer the recovered PDE to have the form of Eq. (2) instead of $\mathsf{N}[U] = U_{tt} + \alpha U_t + c^2(-U_{xx})$. This is achieved by adding the negative signs onto entries of $\mathbf{s}$ instead of onto the corresponding columns in $\mathbf{\Phi}_n$ while keeping $\mathbf{s}$ an all-one vector.

The information in $\mathbf{s}$ enables identifying PDEs from limited data because it encourages Eq. (7b) to select $\bar{\mathbf{a}}_n$, whose nonzero entries have the same signs as their corresponding entries in $\mathbf{s}$ and, hence, filter out the potential combinations of columns in $\bar{\mathbf{\Phi}}_n$, which are better fitted but have no physical meaning. To see this, suppose there are two vectors, $\mathbf{p}, \mathbf{q} \in \mathbb{R}^D$, which satisfy the requirement for $\bar{\mathbf{a}}_n$ in Eq. (7b) for a small τ_n, i.e.,

$$\begin{aligned} &||\bar{\mathbf{\Phi}}_n\mathbf{p}||_\infty \le \tau_n, \quad ||\bar{\mathbf{\Phi}}_n\mathbf{q}||_\infty \le \tau_n, \\ &\left|\sum_{i=1}^{D} s_i p_i - 1\right| \le \tau_n, \quad \left|\sum_{i=1}^{D} s_i q_i - 1\right| \le \tau_n, \end{aligned} \tag{8}$$

and, thus, using the triangle inequality,

$$\left|\sum_{i=1}^{D} s_i p_i - \sum_{i=1}^{D} s_i q_i\right| \le 2\tau_n. \tag{9}$$

If the signs of nonzero entries in $\mathbf{p}$ are the same as corresponding entries in $\mathbf{s}$, while for $\mathbf{q}$, one entry, q_{i_0}, has the opposite sign of s_{i_0}, resulting in $s_{i_0}q_{i_0} < 0$, then when $-s_{i_0}q_{i_0} > \tau_n$, which is likely for a small τ_n, the method will choose $\mathbf{p}$ over $\mathbf{q}$ because of a smaller ℓ_1 norm:

$$||\mathbf{p}||_1 = \sum_{i=1}^{D} s_i p_i < \sum_{i=1}^{D} s_i q_i - 2s_{i_0}q_{i_0} = ||\mathbf{q}||_1. \tag{10}$$

An example of the failed identification due to the replacement of the informative $\mathbf{s}$ by $\mathbf{1}$ is given in Sec. IV A.

In addition to encouraging selecting the $\bar{\mathbf{a}}_n$ whose entries have the correct signs, the incorporation of the nonzero $\mathbf{s}$ also aims to avoid the trivial solution $\bar{\mathbf{a}}_n = \mathbf{0}$ and, thus, the sign information $\mathbf{s}$ (even the incorrect $\mathbf{s}$, which leads to a wrong solution as discussed above) must be provided to make the method work. Meanwhile, to ensure that every column of $\bar{\mathbf{\Phi}}_n^s$ in Eq. (7a) has the same ℓ_2-norm such that the optimization is not influenced by magnitudes of dictionary atoms, the entries in $\mathbf{s}$ should have the same magnitude. Here, the magnitude of one is used for simplicity.

We repeat Eq. (7) for all of the N locations in the ROI and, hence, recover the physical properties described by spatially dependent PDEs. Note that for different locations, the set of recovered active PDE terms can be different. For example, in a wavefield, the attenuation of the waves can be negligible in some regions and, thus, the wave equation (2) does not have U_t term, but for other regions where the attenuation is obvious, the wave equation contains the term U_t.

To accelerate searching the PDE terms, we use the equivalence of the ℓ_2-norm and ℓ_∞-norm for a vector (for any $\mathbf{v} \in \mathbb{R}^k$, $\|\mathbf{v}\|_\infty \le \|\mathbf{v}\|_2 \le \sqrt{k}\|\mathbf{v}\|_\infty$), replace the ℓ_∞-norm constraint in Eq. (7b) with ℓ_2-norm, which is $||\bar{\mathbf{\Phi}}_n^s\bar{\mathbf{a}}_n - \mathbf{e}||_2^2 \le \tau_n$, and solve it using its Lagrangian (i.e., Lasso[32,33]),

$$\begin{aligned} \hat{\bar{\mathbf{a}}}_n &= \arg\min_{\bar{\mathbf{a}}_n} ||\bar{\mathbf{a}}_n||_1 + \lambda(||\bar{\mathbf{\Phi}}_n^s\bar{\mathbf{a}}_n - \mathbf{e}||_2^2 - \tau_n) \\ &= \arg\min_{\bar{\mathbf{a}}_n} ||\bar{\mathbf{\Phi}}_n^s\bar{\mathbf{a}}_n - \mathbf{e}||_2^2 + \lambda_n||\bar{\mathbf{a}}_n||_1, \end{aligned} \tag{11}$$

where $\lambda_n = 1/\lambda = 0.2\lambda_0$ is chosen empirically with $\lambda_0 = 2||\bar{\mathbf{\Phi}}_n^{s\mathrm{T}}\mathbf{e}||_\infty = 2$, the boundary parameter above which the output of Eq. (11) is $\mathbf{0}$, according to the Lasso path.[34–36] Equation (11) can be efficiently solved by coordinate descent, where a complete iteration of updating all of the D entries in $\bar{\mathbf{a}}_n$ costs $O((M+1)D)$ operations,[37] and the number of iterations to reach convergence is often small.

Due to the noise from numerical computation, the $\bar{\mathbf{a}}_n$ minimizing Eq. (11) may not be sparse enough. To further promote sparsity, we threshold entries of $\bar{\mathbf{a}}_n$ using an adaptive threshold proportional to $||\bar{\mathbf{a}}_n||_\infty$. Then, use least squares regression to solve the coefficients only in the T kept entries [denoted by $\tilde{\mathbf{a}}_n(\Lambda)$, where Λ with cardinality T is the set of indices for the T kept nonzero entries] and assign zero to the other entries. Thus, Eq. (7c) is replaced by

$$\Lambda = \{\forall i,\ |\bar{\mathbf{a}}_n(i)| \ge \epsilon||\bar{\mathbf{a}}_n||_\infty\}, \tag{12a}$$

$$\tilde{\mathbf{\Phi}}_n^s = \left[\mathbf{\Phi}_n(:,\Lambda)^{\mathrm{T}}\ \mathbf{s}(\Lambda)\right]^{\mathrm{T}} \in \mathbb{R}^{(M+1)\times T}, \tag{12b}$$

$$\hat{\tilde{\mathbf{a}}}_n(\Lambda) = \tilde{\mathbf{\Phi}}_n^{s^\dagger}\mathbf{e}, \tag{12c}$$

where ϵ is the threshold to be tuned. It can be tuned according to prior knowledge or from grid search and cross-validation if training data are available.

https://doi.org/10.1121/10.0019592

B. Denoising by integration

In this section, we extend the technique of integration transformation[7] to the spatially dependent PDE identification to make it more robust to noise, see Sec. II B 1, and discuss implementation details in Sec. II B 2. Overall, the idea is to replace $\mathbf{\Phi}_n$ in Eq. (3) with a new dictionary, $\mathbf{\Phi}_n^{\text{int}}$, which is built from integration transformation, as detailed below. All of the other steps for the PDE identification are the same as those in Sec. II A.

1. Transformed dictionary by integration

The terms in $\mathbf{\Phi}_n$ defined in Eq. (3) are sensitive to noise because the noise is typically broadband and its influence is emphasized by differentiation. We extend the denoising technique by integration[7] to spatially dependent PDE identification to recover the spatial variation of properties from noisy measurements. To be specific, replacing $\mathbf{\Phi}_n$ in Eq. (3) with another dictionary, $\mathbf{\Phi}_n^{\text{int}}$, as described below, and finishing all of the following steps for PDE identification in Sec. II A using it.

The integration method uses integration by parts to transfer the derivatives of noisy measurements, $\mathbf{U}$, to the derivatives of a predefined weighting function $\mathbf{W}$, which is noise-free.

For an arbitrary region Ω in the ROI and a finite smooth function $W(x,t)$ defined on Ω,

$$\mathsf{N}[U] = 0 \;\rightarrow\; \int_\Omega \mathsf{N}[U] W \,\mathrm{d}\Omega = 0, \tag{13}$$

where $\int_\Omega \mathsf{N}[U] W \,\mathrm{d}\Omega$ is for $\int_{(x,t)\in\Omega} \mathsf{N}[U(x,t)] W(x,t) \mathrm{d}x\,\mathrm{d}t$. From Eq. (13), in an arbitrary region Ω within the ROI, the integral of the product between W and the summation of all of the active PDE terms is zero. For example, if $\mathsf{N}(U)$ is the wave equation (2), then Eq. (13) becomes

$$\int_\Omega (U_{tt} + \alpha U_t - c^2 U_{xx}) W \,\mathrm{d}\Omega = 0. \tag{14}$$

Let Ω be a square region $\{x_l \le x \le x_u, t_l \le t \le t_u\}$ with $(x_u - x_l)/\Delta x = (t_u - t_l)/\Delta t = 2a$ intervals so that it covers $(2a+1)\times(2a+1)$ spatial-temporal coordinates within the ROI, we can move the derivatives of U in Eq. (13) onto W, e.g., for component U_t,

$$\begin{aligned}\int_\Omega W U_t \mathrm{d}\Omega &= \int_\Omega (WU)_t \mathrm{d}\Omega - \int_\Omega U W_t \mathrm{d}\Omega \\ &= \int_{x_l}^{x_u} \left[(WU)\big|_{t_l}^{t_u} \right] \mathrm{d}x - \int_\Omega U W_t \mathrm{d}\Omega \\ &= -\int_\Omega U W_t \mathrm{d}\Omega,\end{aligned} \tag{15}$$

where W is defined as zero at t_l and t_u. Similarly, using a W where its $(p-1)$st-order derivatives ($p \ge 2$) vanished at its spatial-temporal boundaries, we can transfer the derivatives on U for all of the terms in Eq. (3) onto W. An eligible W is

$$W(x,t) = (\bar{x}^2 - 1)^p (\bar{t}^2 - 1)^p,$$

where

$$\begin{aligned}\bar{x} &= 2\frac{x - x_l}{x_u - x_l} - 1 \in [-1, 1], \\ \bar{t} &= 2\frac{t - t_l}{t_u - t_l} - 1 \in [-1, 1].\end{aligned} \tag{16}$$

Because the PDEs are assumed to be spatially dependent, for spatial location n_x in $\mathbf{U}$, we select M_{int} integration domains $\Omega_1, \ldots, \Omega_{M_{\text{int}}}$ centered at $(n_x, (m_0 + \delta)_t), \ldots, (n_x, (m_0 + M_{\text{int}}\delta)_t)$, where δ is the interval between temporal centers of two neighboring domains, and the spatial center of the domains is always the location n_x. All of the Ω_m with $m = 1, \ldots, M_{\text{int}}$ are of the same size as the Ω mentioned previously. Ideally $\delta = 1$, but a larger δ is chosen to reduce computations. Thus, we can construct a new library of atoms, $\mathbf{\Phi}_n^{\text{int}} \in \mathbb{C}^{M_{\text{int}} \times D}$, consisting of integration for $D = 6$ integrands on all Ω_m, $1 \le m \le M_{\text{int}}$, whose mth row is

$$\begin{aligned}\mathbf{\Phi}_n^{\text{int}}(m,:) = \int_{\Omega_m} \Big[&-UW_t^m, UW_{tt}^m, -\frac{1}{2}U^2 W_x^m, \\ &UW_{xx}^m, UW_{xt}^m, \sin(U) W^m \Big] \mathrm{d}\Omega_m,\end{aligned} \tag{17}$$

where W^m is the shifted W such that its domain exactly overlaps Ω_m.

2. Computation of integrals

To compute entries in Eq. (17) numerically, we evaluate the values of W^m at discrete coordinates and save it in $\mathbf{W}^m \in \mathbb{R}^{N\times M}$, in which

$$\mathbf{W}^m(i_x, i_t) = \begin{cases} W(i_x\Delta x, i_t\Delta t) & \text{if } (i_x, i_t) \in \Omega_m, \\ 0 & \text{otherwise,} \end{cases} \tag{18}$$

where W is predefined with an eligible choice given in Eq. (16), whose x_u, x_l, t_u, and t_l are the spatiotemporal boundaries of Ω_m. Similarly, the derivatives are also computed and saved in matrices in $\mathbb{R}^{N\times M}$, e.g.,

$$\mathbf{W}_{tt}^m(i_x, i_t) = \begin{cases} W_{tt}(i_x\Delta x, i_t\Delta x) & \text{if } (i_x, i_t) \in \Omega_m, \\ 0 & \text{otherwise,} \end{cases} \tag{19}$$

where $W_{tt}(i_x\Delta x, i_t\Delta t) = (\partial^2 W(x,t)/\partial t^2)|_{x=i_x\Delta x, t=i_t\Delta t}$ is computed analytically and, thus, noise-free.

A demonstration of the integration is shown in Fig. 1(a), where $M_{\text{int}} = 4$. The integration of the dth integrand indicated in Eq. (17) within Ω_m is $\mathbf{\Phi}_n^{\text{int}}(m,d)$. Figure 1(b) shows a $\mathbf{W}^m$ with its nonzero region in $\Omega_m = \{(x,t) | 20\Delta x \le x \le 30\Delta x, 17\Delta t \le t \le 27\Delta t\}$ enclosed by a dashed box.

By choosing W as in Eq. (16) with $p \ge 2$ and analytically computing its derivatives, all of the derivatives of W necessary for Eq. (17) are obtained, enabling the transfer of derivatives on U to derivatives on W. The integral for each term in Eq. (17) within each Ω_m is integrated along t and then along x numerically by summation approximation. Note that W and its derivatives are calculated analytically with $\bar{t}$ in $[-1, 1]$, but the

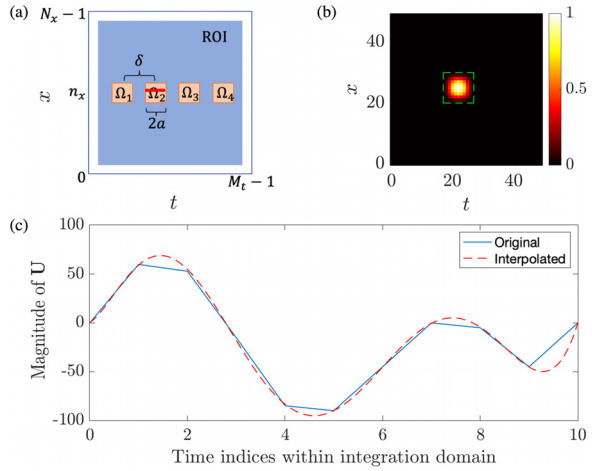

FIG. 1. (Color online) A demonstration for integration, showing (a) four integration domains in ROI with each having $2a+1$ points ($2a$ intervals) in x and t axes. It is possible to choose $\delta < 2a$, i.e., the domains can be overlapped. (b) A $\mathbf{W}^m$ whose nonzero part (in green dashed box) is centered at $(x = 25\Delta x, t = 22\Delta t)$. (c) When integrating along a slice of signal in time [e.g., red line in (a)], use polynomial interpolation and integrate on the interpolated slice (red dashed line).

true domain of W covers $[-a\Delta t, a\Delta t]$. Thus, for the p_0th-order temporal derivative of W, its real values are the analytical result divided by $(a\Delta t)^{p_0}$, and for spatial derivatives, the scaling is similar. For example, the mth row of the second column in $\mathbf{\Phi}_n^{\text{int}}$ is computed as

$$\mathbf{\Phi}_n^{\text{int}}(m,2) = \left(\frac{1}{a\Delta t}\right)^2 \sum_{\{i_x,i_t\}\in\Omega_m} \mathbf{U}(i_x,i_t)\mathbf{W}_{tt}^m(i_x,i_t)\Delta x\Delta t, \quad (20)$$

which is a scaled summation of $(2a+1)^2$ values of the element-wise product between $\mathbf{U}$ and $\mathbf{W}_{tt}^m$ within Ω_m.

To reduce the error caused by approximating the integration with a finite sum, which can dominate when the derivative of W has a high order, we use interpolation. Before integration, each time slice of the raw signal $\mathbf{U}$ within Ω_m [e.g., red line in Ω_2 in Fig. 1(a)] is interpolated using polynomial fitting. Let the raw signal be the blue line in Fig. 1(c) with $2a+1$ points (here, $a=5$), we fit it using a fifth-order polynomial and interpolate q values evenly between each neighboring point. Therefore, the interpolated signal has $2a(q+1)+1$ points. To match the interpolated signal, W^m and its derivatives are also evaluated at these evenly spaced $2a(q+1)+1$ points along the temporal direction. We then sum up values of the integrand at all of these points to be the temporal integral for this location, and $\mathbf{\Phi}_n^{\text{int}}(m,d)$ is acquired by summing up such temporal integrals for all $2a+1$ spatial locations within Ω_m. Thus, in the example for Eq. (20), $\mathbf{\Phi}_n^{\text{int}}(m,2)$ becomes a scaled summation of $(2a+1)\times(2a(q+1)+1)$ values.

Replacing $\mathbf{\Phi}_n$ in Eq. (3) with $\mathbf{\Phi}_n^{\text{int}}$ efficiently increases its robustness to noise, as shown in Secs. III C and IV B. In all of the experiments, $a=5$, $p=3$, and $q=9$.

III. NUMERICAL EXPERIMENTS

This section includes a spatially independent (not known *a priori*) and two spatially dependent PDE identification experiments. Sections III A–III B are based on clean signal, and Sec. III C experiments with noisy measurements. The datasets are generated by finite difference modeling.[38]

For the 1D case in Sec. III A, the dictionary used is Eq. (3). To simplify the demonstration, we index its columns as

$$\mathbf{\Phi}_n = \left[\mathbf{u}_t^1,\ \mathbf{u}_{tt}^2,\ \mathbf{u}\circ{}^3\mathbf{u}_x,\ \mathbf{u}_{xx}^4,\ \mathbf{u}_{tx}^5,\ \overset{6}{\sin(\mathbf{u})}\right] \in \mathbb{R}^{M\times 6}. \quad (21)$$

and use the indices for the columns (1–6) to show the results.

A. Spatially independent Burgers's equation

Here, we recover the fluid viscosity from fluid speeds. The Burgers's equation,[39,40]

$$U_t + UU_x - \nu U_{xx} = 0, \quad (22)$$

is nonlinear and can model the formation of shock waves in free turbulence, where $\nu \geq 0$ is the viscosity of the fluid, which is spatially independent in this example. The aim of this experiment is to show that the method can extract the medium property from a nonlinear dynamic system, and it works when the PDE to be identified is, in fact, spatially independent (indicating a homogeneous medium) but not known *a priori*.

A dataset $\mathbf{U}\in\mathbb{R}^{101\times 151}$, with $\Delta t = 0.05$ s and $\Delta x = 1$ m, modeling the speed of the fluid at each location along a thin pipe as time progresses governed by Eq. (22) with $\nu = 0.1$, is generated as Fig. 2(a). The initial state is a scaled probability density function (PDF) of the normal distribution, and as time goes by, the wave is moving in the positive x direction.

We choose the spatial region where the dynamics can be easily observed as the ROI and, to be specific, choose $I_x = \{i_x | 20 \leq i_x \leq 90\}$, which is bounded by the red lines in Fig. 2(a), such that $N=71$ and $n=1$ corresponds to $i_x=20$. We do not consider the temporal boundaries where the derivatives are not well defined and, thus, use $I_t = \{i_t | 1 \leq i_x \leq 149\}$ for the ROI, i.e., $M=149$.

To find the governing PDE, we build $\bar{\mathbf{\Phi}}_n^s$ according to Eqs. (3), (6), and (7a) for every $1\leq n\leq N$. From Eq. (11), the coefficients are distributed as Fig. 2(b). After thresholding using Eq. (12a) with $\epsilon = 10^{-3}$, $\{\mathbf{u}_t, \mathbf{u}\circ\mathbf{u}_x, \mathbf{u}_{xx}\}$ appearing in the Burgers's equation are selected for all of the locations in the ROI.

For every location, we build $\tilde{\mathbf{\Phi}}_n^s$ and compute $\tilde{\mathbf{a}}_n$ as in Eq. (12). The coefficients for U_t and UU_x are always nearly identical as $\sum_{n=1}^N |\tilde{\mathbf{a}}_n(3) - \tilde{\mathbf{a}}_n(1)| = 1.7\times 10^{-14}$. The estimated viscosity is $\hat{\nu}_n = -\tilde{\mathbf{a}}_n(4)/\tilde{\mathbf{a}}_n(1)$, which is equal to 0.1 for every n.

B. 2D spatially dependent wave equation

In this section, we recover the 2D maps of phase speeds and attenuation from observed propagating waves. A 2D wavefield $\mathbf{U}\in\mathbb{R}^{32\times 32\times 200}$ in which $\Delta x = \Delta y = 0.1$ m and

https://doi.org/10.1121/10.0019592

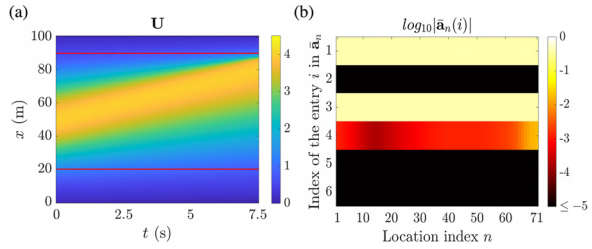

FIG. 2. (Color online) For Burgers's equation (22), (a) $\mathbf{U}$ with spatially independent $\nu = 0.1$, $\Delta x = 1$ m, $\Delta t = 0.05$ s. Therefore, $0 \le i_x \le 100$, $0 \le i_t \le 150$. The ROI is $\{i_x | 20 \le i_x \le 90\}$ between the red lines where obvious dynamics are observed. (b) $\log_{10}|\bar{\mathbf{a}}_n(i)|$, where i corresponds to the indices in Eq. (21) for all $n = 1, \ldots, 71$.

$\Delta t = 0.01$ s describing waves excited by an initial perturbation and propagating through various media is used for the experiment. The PDE governing it is the wave equation

$$U_{tt} + \alpha U_t - c^2 \nabla^2 U = 0, \tag{23}$$

where $\alpha \ge 0$ is the attenuating factor, $c > 0$ is the phase speed, and ∇^2 is the Laplacian, i.e., $U_{xx} + U_{yy}$. We adopt the Dirichlet boundary condition, where the measurements on spatial boundaries are zero. The initial perturbation is shaped as a scaled 2D normal distribution PDF, and the phase speeds $2 \le c \le 4$ m/s and attenuation $0 \le \alpha \le 0.2$ are varying across the domain as shown in Fig. 3. Some frames are shown in Fig. 4. We choose the ROI to be all of the spatial regions without the boundaries and its immediate neighboring points (i.e., $2 \le i_x \le 29$, $2 \le i_y \le 29$) and the time steps $1 \le i_t \le 198$. Therefore, $N = 28^2 = 784$, $M = 198$, and $n = 1$ corresponds to $i_x = i_y = 2$. The 2D locations within ROI are indexed from 1 to N in the row-major manner.

For this 2D case, we extend the dictionary [Eq. (3)] and $\mathbf{s}$ [Eq. (6)] to

$$\begin{aligned} \mathbf{\Phi}_n = \Big[& \mathbf{u}_t^1, \quad \mathbf{u}_{tt}^2, \quad \mathbf{u} \circ {}^3\mathbf{u}_x, \quad \mathbf{u} \circ {}^4\mathbf{u}_y, \quad \mathbf{u}_{xx}^5, \\ & \mathbf{u}_{yy}^6, \quad \mathbf{u}_{tx}^7, \quad \mathbf{u}_{ty}^8, \quad \overset{9}{\sin}(\mathbf{u}) \Big], \\ \mathbf{s} = & [1, 1, 1, 1, -1, -1, -1, -1, 1]^{\mathrm{T}}. \end{aligned} \tag{24}$$

Build dictionaries using Eq. (24) according to Eq. (7a), and from Eq. (11), the $\bar{\mathbf{a}}_1$ to $\bar{\mathbf{a}}_N$ are acquired as Fig. 5(a). After thresholding as in Eq. (12a) with $\epsilon = 10^{-3}$, the kept nonzero entries are indicated in Fig. 5(b).

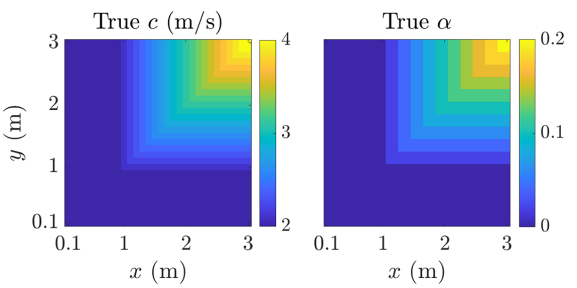

FIG. 3. (Color online) The true phase speeds c and attenuating factors α for i_x,i_y in $[1, 30]$ ($\Delta x = \Delta y = 0.1$ m). Waves cannot arrive at the places where either i_x or i_y is 0 or 31 because of the boundary condition.

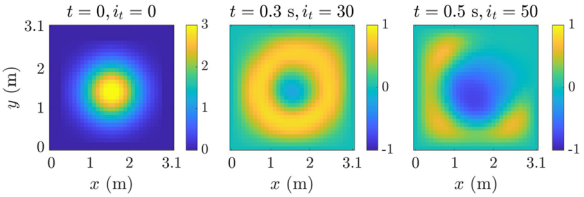

FIG. 4. (Color online) The wavefield governed by Eq. (23) with spatially dependent c and α at three selected time points. $\Delta x = \Delta y = 0.1$ m, $\Delta t = 0.01$ s. Therefore, i_x,i_y ranged from 0 to 31.

Comparing Fig. 5(b) to Eq. (23), the method successfully identifies the PDEs for all 784 locations. Figure 6 shows the number of identified active PDE terms for each location in the ROI, where the three terms contain $\{\mathbf{u}_{tt}, \mathbf{u}_{xx}, \mathbf{u}_{yy}\}$, and for four terms, $\mathbf{u}_t$ is also included. Build $\tilde{\mathbf{\Phi}}_n^s$ and compute $\tilde{\mathbf{a}}_n$ as in Eq. (12), the coefficients for U_{xx} and U_{yy} are nearly identical as $\sum_{n=1}^{N} |\tilde{\mathbf{a}}_n(5) - \tilde{\mathbf{a}}_n(6)| = 3.3 \times 10^{-12}$. The recovered $\hat{c}_n = \sqrt{-(\tilde{\mathbf{a}}_n(5) + \tilde{\mathbf{a}}_n(6))/2\tilde{\mathbf{a}}_n(2)}$ and $\hat{\alpha}_n = \tilde{\mathbf{a}}_n(1)/\tilde{\mathbf{a}}_n(2)$, which are satisfactory as the root-mean-square error $\mathrm{RMSE} = 1.1 \times 10^{-14}$ m/s for phase speeds and 1.5×10^{-14} for attenuating factors with respect to the ground truth in Fig. 3 across the whole ROI.

C. 2D spatially dependent wave equation with noise

The field is of size $\mathbf{U} \in \mathbb{R}^{100 \times 100 \times 1000}$, $\Delta x = \Delta y = 1$ m, $\Delta t = 0.2$ s. The ROI is selected to be $5 \le i_x < 95, 5 \le i_y < 95$ (8100 locations in total), and for time period use, $5 \le i_t < 995$. The field has free boundaries and is excited by two chirp sources located outside of the region at $(-6\text{ m}, 18\text{ m})$ and $(109\text{ m}, 80\text{ m})$, and governed by attenuating or non-attenuating wave equations with various coefficients, as indicated in Fig. 7(a). One frame is depicted in Fig. 8(a).

Without noise, building dictionaries for all of the locations using Eq. (24) according to Eq. (7a) and implementing Eq. (11)), we obtain the coefficients distributed as in Fig. 9. Thresholding them with $\epsilon = 10^{-2}$, the PDEs are correctly identified for all 8100 locations in the ROI. The $\hat{c}$ and $\hat{\alpha}$ are well recovered with the $\mathrm{RMSE} = 1.7 \times 10^{-15}$ m/s for phase speeds and 3.2×10^{-16} for attenuating factors compared to the ground truth.

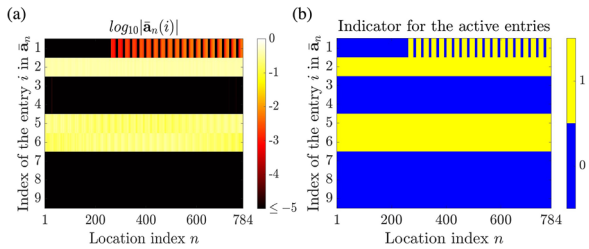

FIG. 5. (Color online) For 2D wave equation, showing (a) $\log_{10}|\bar{\mathbf{a}}_n(i)|$, where i corresponds to the indices of columns for $\mathbf{\Phi}_n$ in Eq. (24) for all $n = 1, \ldots, 28^2$ and (b) the locations of their active entries after thresholding. The 2D 28×28 locations are indexed from 1 to 784 in a row-major manner.

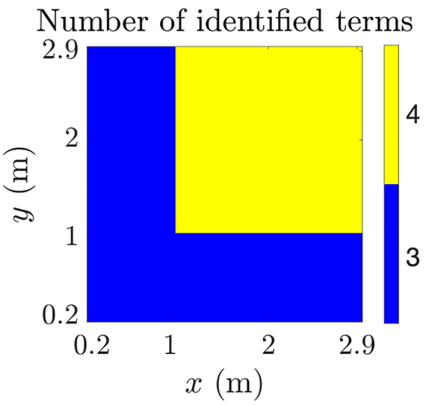

FIG. 6. (Color online) Number of identified active PDE terms within the ROI.

We add additive white Gaussian noise (AWGN) with $\sigma^2 = 2$ (arbitrarily chosen to be $\sim$1% of the signal variance in a frame) to the measurements; see Fig. 8(b). Using the same procedure as before without denoising, the $\bar{\mathbf{a}}_n$ is as in Fig. 10(a). No proper ϵ can be found to extract the active terms and, thus, the PDE identification fails. We, therefore, use the integration transformation to build $\mathbf{\Phi}_n^{\text{int}} \in \mathbb{R}^{M_{\text{int}} \times 9}$ for each location as described in Sec. II B with its mth row such that

$$\mathbf{\Phi}_n^{\text{int}}(m,:) = \int_{\Omega_m} \Big[\bar{U}\overset{1}{W}{}^m_t, \quad U\overset{2}{W}{}^m_{tt}, \quad -\frac{1}{2}U^2\overset{3}{W}{}^m_x, \\ -\frac{1}{2}U^2\overset{4}{W}{}^m_y, \quad U\overset{5}{W}{}^m_{xx}, \quad U\overset{6}{W}{}^m_{yy}, \\ U\overset{7}{W}{}^m_{xt}, \quad U\overset{8}{W}{}^m_{yt}, \quad \sin(U)\overset{9}{W}{}^m \Big] \mathrm{d}\Omega_m. \quad (25)$$

The center of the integration domains starts from $i_t = 19$ and ends at $i_t = 979$ with $\delta = 10$, hence, $M_{\text{int}} = 97$. The

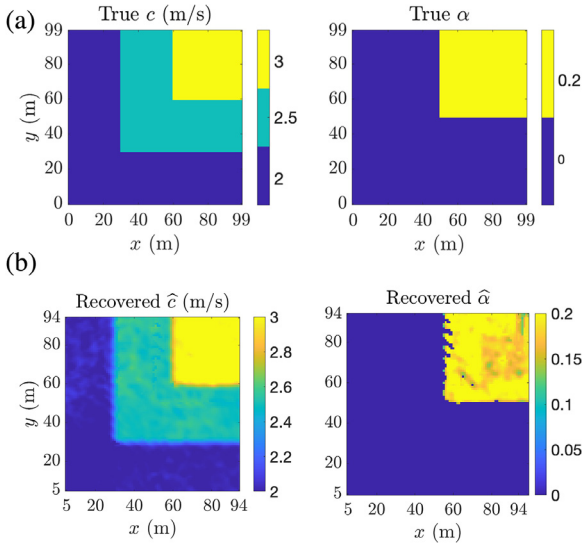

FIG. 7. (Color online) The (a) true phase speeds c and attenuating factors α for i_x,i_y in $[0, 99]$ ($\Delta x = \Delta y = 1$ m) and (b) recovered $\hat{c}$ and $\hat{\alpha}$ in the ROI for the noisy measurements using $\epsilon = 0.13$ are shown.

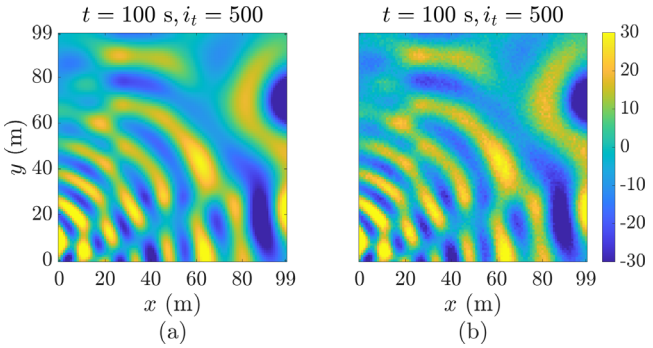

FIG. 8. (Color online) A frame of the wavefield governed by Eq. (23) with spatially dependent c and α in Fig. 7(a). The two sources are outside of the region at $(-6\text{ m}, 18\text{ m})$ and $(109\text{ m}, 80\text{ m})$, showing (a) clean measurements and (b) noisy measurements with AWGN for $\sigma^2 = 2$.

recovered $\bar{\mathbf{a}}_n$ is in Fig. 10(b), which is a clear improvement of Fig. 10(a). Now it is possible to employ thresholding to extract active terms. Using $\epsilon = 0.13$, the recovered result is shown in Fig. 7(b), where the terms corresponding to $\{U_{tt}, U_{xx}, U_{yy}\}$ or $\{U_t, U_{tt}, U_{xx}, U_{yy}\}$ are identified at 97.6% of the 8100 locations. Specifically, for the ROI with attenuation [upper right in Fig. 7(a)], the four PDE terms $\{U_t, U_{tt}, U_{xx}, U_{yy}\}$ are selected at 86.1% of the 45^2 locations. For each location of the 2.4% region, where the wave equation is not identified, the recovered $\hat{c}$ and $\hat{\alpha}$ are interpolated using the median value within a window covering 21 locations along the y axis centered at it.

For the phase speed recovery, the sharp edges between distinct true speeds are smoothed. The reason is that the integration domains centered near the edges cover the regions with different speeds, thus, the results are affected by both speeds. The recovered speed smoothly changes because the integration domain smoothly slides over the boundary. The integration domain size is important: a larger integration domain leads to more noise-robust estimation and more extensively smoothed edges. Unlike phase speed recovery, when the integration domain centers near the boundary of attenuating and non-attenuating areas, if $\bar{\mathbf{a}}_n(1)$ is kept after thresholding in Eq. (12a), then $\hat{\alpha}$ is recovered

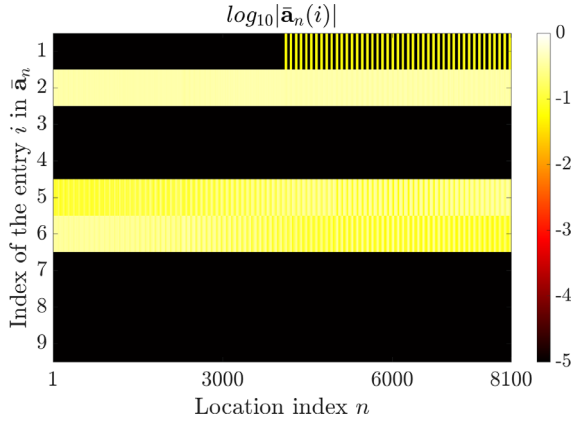

FIG. 9. (Color online) For 2D wave equation with spatially dependent c and α in Fig. 7(a), $\log_{10}|\bar{\mathbf{a}}_n(i)|$ from clean measurements, where i corresponds to the indices of columns for $\mathbf{\Phi}_n$ in Eq. (24).

https://doi.org/10.1121/10.0019592

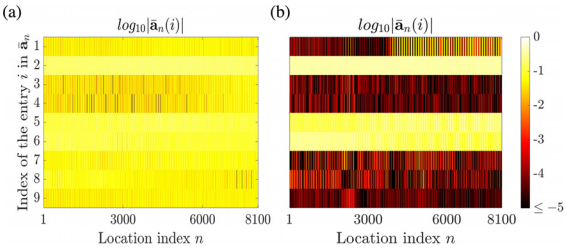

FIG. 10. (Color online) For noisy measurements of waves governed by Eq. (23) with c and α in Fig. 7(a), showing (a) $\log_{10}|\bar{\mathbf{a}}_n(i)|$ with $\bar{\mathbf{a}}_n$ recovered from $\mathbf{\Phi}_n$, where i corresponds to the indices of columns for $\mathbf{\Phi}_n$ in Eq. (24), and (b) $\log_{10}|\bar{\mathbf{a}}_n(i)|$ with $\bar{\mathbf{a}}_n$ recovered from $\mathbf{\Phi}_n^{\text{int}}$ in Eq. (25).

from least squares regression in Eq. (12a) and, therefore, near the true α, otherwise $\tilde{\mathbf{a}}_n(1)$ is set to zero, thus, $\hat{\alpha} = 0$, causing the sharp transition for $\hat{\alpha}$.

IV. EXTRACTING PDES FOR A VIBRATING PLATE

A. Identification from clean measurements

The approach is demonstrated on laser scanned measurements of vibrations of a real aluminum plate sampled at 300 kHz, which was provided by University of Utah;[41] see Fig. 11. The part taken into consideration is $\mathbf{U} \in \mathbb{R}^{100\times100\times1000}$, i.e., the measurements are collected from 10 000 sampling locations uniformly distributed on the square plate (100 rows and 100 columns) with 1000 time steps. The spatial sampling interval $\Delta x = \Delta y = 1$ mm.

The PDE governing the vibrations of the plate is the wave equation. Because the aluminum plate waves are dispersive,[42] i.e., phase speeds c varies across frequencies, we extract the narrow band signals from $\mathbf{U}$ and identify the PDE for every band. Five sixth-order Butterworth bandpass filters centered at 30–70 kHz stepped by 10 kHz are employed to extract narrowband signals with 2 kHz bandwidth, with one frame of the narrowband signal centered at 30 kHz as shown in Fig. 12(a). We drop five neighboring locations on each axis end, hence, for

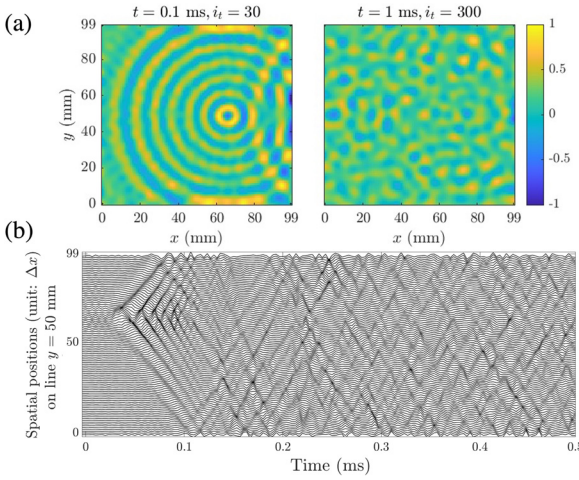

FIG. 11. (Color online) The vibrating plate, showing (a) 2 of the selected 1000 frames with magnitudes normalized and (b) the traces for locations at $y = 50$ mm in the first 0.5 ms.

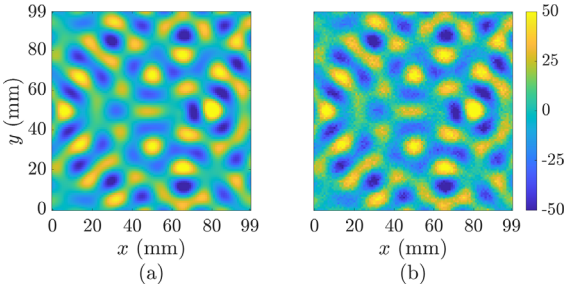

FIG. 12. (Color online) For the 300th frame of the signal with frequency band centered at 30 kHz, showing (a) clean signal and (b) signal corrupted by AWGN with $\sigma^2 = 10^3$.

the ROI $5 \le i_x < 95$, $5 \le i_y < 95$, $5 \le i_t < 995$ and, thus, $N = 90^2 = 8100$, $M = 990$.

The PDE identification results are summarized in Table I, which is explained in detail in the following. We build the dictionary $\mathbf{\Phi}_n \in \mathbb{R}^{990\times9}$ for each location as in Eq. (24), and build $\bar{\mathbf{\Phi}}_n^s$ using the terms and $\mathbf{s}$ in Eq. (24) according to Eq. (7a) for every $1 \le n \le N$. From Eq. (11), the coefficients for $\{\mathbf{u}_{tt}, \mathbf{u}_{xx}, \mathbf{u}_{yy}\}$ are found to have significant greater magnitudes, as shown in Fig. 13(a) for the 70 kHz case as an example. We threshold $\bar{\mathbf{a}}_n(i)$ as in Eq. (12a), using $\epsilon = 0.2$ at each location that gives the active PDE terms. Counting the number of locations with the PDE terms correctly identified (either $\{\mathbf{u}_{tt}, \mathbf{u}_{xx}, \mathbf{u}_{yy}\}$ or these including $\mathbf{u}_t$) and normalizing it with N gives the "success rate." The mean phase speed and mean absolute deviation (MAD, computed by $(1/N)\sum_{n=1}^{N}|\hat{c}_n - ((1/N)\sum_{n=1}^{N}\hat{c}_n)|$, where $\hat{c}_n$ is the recovered phase speed at location n) over these "success" locations, are calculated. The recovered speeds $\hat{c}_n = \sqrt{-(\tilde{\mathbf{a}}_n(5) + \tilde{\mathbf{a}}_n(6))/2\tilde{\mathbf{a}}_n(2)}$ are shown in Fig. 14. For each location in the region where the wave equation is not identified (3.2% of the ROI at most, at 70 kHz), the recovered $\hat{c}$ is interpolated using the median value within a window covering 21 locations along the y axis centered at it.

The mean phase speeds in Table I are close to the recovered phase speeds when the PDE coefficients are assumed constant across the space and recovered from one dictionary built from the measurements at all of the locations (a big dictionary $\mathbf{\Phi} = [\mathbf{\Phi}_1^{\mathrm{T}}\ \mathbf{\Phi}_2^{\mathrm{T}}\ \cdots\ \mathbf{\Phi}_N^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{NM\times D}$) in Ref. 10, as given in $\hat{c}_{\text{all}}$. The relative MAD is ∼1%, indicating the recovered speed is nearly a constant across the plate for a narrow frequency band, which coincides with our physical setting.

TABLE I. Success rate [using $\mathbf{s}$ as in Eq. (24) or $\mathbf{s} = \mathbf{1}$ in $\mathbf{\Phi}_n^s$] of the PDE identification and the recovered speeds for various frequency bands.

Center frequency (kHz)	30	40	50	60	70
Success rate using correct $\mathbf{s}$ in Eq. (24) (%)	99.5	98.1	99.1	99.4	96.8
Success rate using $\mathbf{s} = \mathbf{1}$ (%)	0	0	0	0	0
Mean speed (m/s)	467	532	586	628	663
MAD (m/s)	4.9	8.7	6.9	7.7	6.1
MAD/mean (%)	1.1	1.6	1.2	1.2	0.9
$\hat{c}_{\text{all}}$ (m/s)	463	531	587	632	668

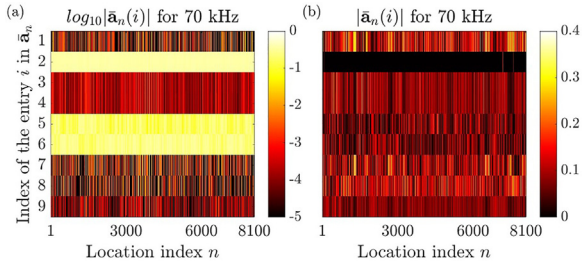

FIG. 13. (Color online) Magnitudes of $\bar{\mathbf{a}}_n(i)$, where i corresponds to the indices of columns for $\mathbf{\Phi}_n$ in Eq. (24) for all $n = 1, \ldots, 90^2$ from clean measurements, showing (a) using $\mathbf{s}$ in Eq. (24) and (b) using $\mathbf{1}$ as the $\mathbf{s}$.

The correct $\mathbf{s}$ selected from physical knowledge is a key for the successful identification. If we use $\mathbf{1}$ as $\mathbf{s}$ in Eq. (24), no wave equations are identified at any location for any frequency band. Using $\mathbf{1}$, the $\bar{\mathbf{a}}_n$ for the band centered at 70 kHz is shown in Fig. 13(b). The $\mathbf{1}$ used as $\mathbf{s}$ indicates the coefficients for U_{tt}, U_{xx}, and U_{yy} have the same sign. As this is not true, these terms are suppressed [comparing Fig. 13(b) with Fig. 13(a)], and the energy that should appear in $\bar{\mathbf{a}}_n(2)$, $\bar{\mathbf{a}}_n(5)$, and $\bar{\mathbf{a}}_n(6)$ is redistributed to other entries to make the combination of selected columns in $\bar{\mathbf{\Phi}}_n^s$ still fit $\mathbf{e}$. This cannot be remedied by tuning ϵ because the incorrect terms have larger coefficients than the correct terms.

B. Identification from noisy measurements

To test the robustness against AWGN, we add the AWGN with variance $\sigma^2 = 10^4$, 10^3, 10^2, 10, and one to the signal with frequency band centered at 30 kHz (one frame depicted in Fig. 12), and identify PDEs from the dictionary [Eq. (25)] where the center of the integration domains starts from $i_t = 19$ and ends at $i_t = 979$ with $\delta = 20$ (hence, $M_{\text{int}} = 49$). The successful identification rate is in Table II, in which the "signal variance" in the second column is the average variance across all 1000 frames of the clean signal centered at 30 kHz in the ROI. Table II shows that the integration method significantly increases the robustness against AWGN. In reality, the method should provide a high success rate for the sensors satisfying basic quality requirements. From Table II, when the noise variance is 10^2, i.e., the noise variance is 66% of the signal variance and thus the signal-to-noise ratio SNR = 1.83 dB, the success rate is larger than 98%.

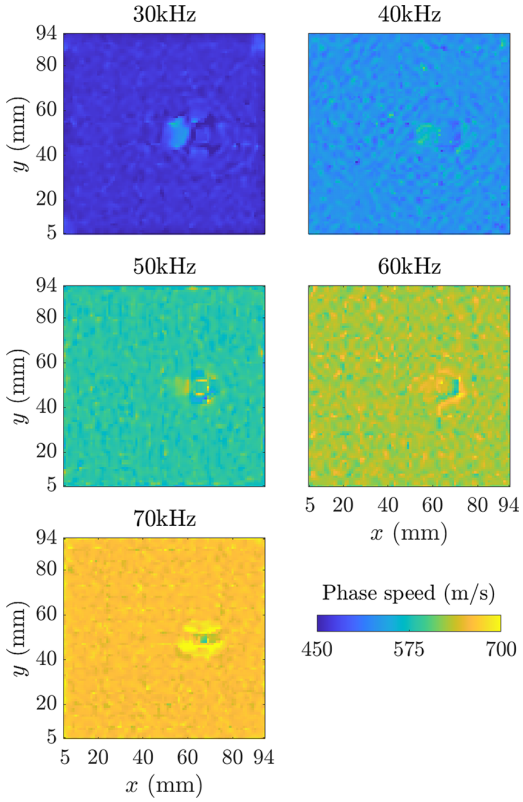

FIG. 14. (Color online) The recovered phase speeds for various frequency bands on the plate from clean measurements.

TABLE II. PDE identification success rate with or without integration transformation for the noisy narrowband signal at 30 kHz. Values in the second column are the direct quotient, not the percentage.

Noise variance	Noise/signal variance	Without integration (%)	With integration (%)
10^4	65.56	2.2	9.2
10^3	6.56	0.3	62.1
10^2	0.66	1.2	98.1
10^1	6.56×10^{-2}	25.3	99.5
10^0	6.56×10^{-3}	90.9	99.5

Add AWGN with $\sigma^2 = 100$ to all five narrowband signals and use the integration transformation to assist PDE identification, and the PDE coefficients are recovered with the example for 70 kHz as shown in Fig. 15. The coefficients for $\{U_{tt}, U_{xx}, U_{yy}\}$ have larger coefficients at most locations and, thus, can be extracted by thresholding. Use $\epsilon = 0.2$ to extract active terms and the recovered speeds $\hat{c}_n = \sqrt{-(\tilde{\mathbf{a}}_n(5) + \tilde{\mathbf{a}}_n(6))/2\tilde{\mathbf{a}}_n(2)}$ are shown in Fig. 16 and recorded in Table III. For each location in the region where the wave equation is not identified (6.6% of the ROI, at most, at 60 kHz), the recovered $\hat{c}$ is interpolated using the median value within a window covering 21 locations along the y axis centered at it.

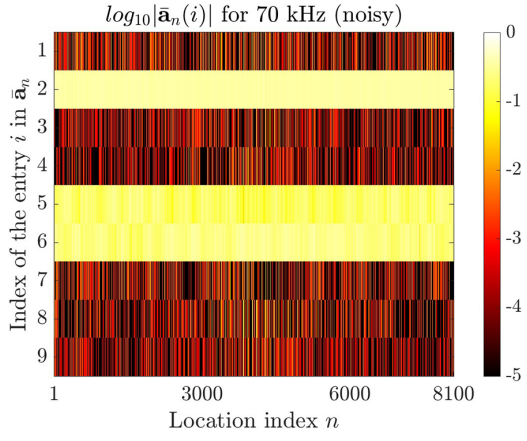

FIG. 15. (Color online) From noisy measurements of the vibrating plate centered at 70 kHz, $|\bar{\mathbf{a}}_n(i)|$, where i corresponds to the indices of columns for $\mathbf{\Phi}_n^{\text{int}}$ in Eq. (25) for all $n = 1, \ldots, 90^2$.

https://doi.org/10.1121/10.0019592

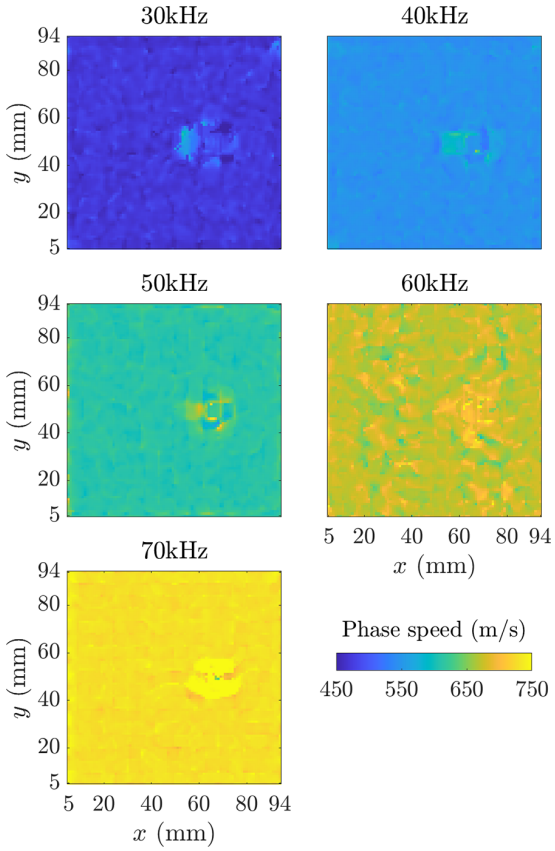

FIG. 16. (Color online) The recovered phase speeds for various frequency bands on the plate from noisy measurements.

The mean speeds in Table I for high frequencies are underestimated because of insufficient sampling along time.[10] For 70 kHz, if we assume the speed is 700 m/s, the wavelength becomes 10^{-2} m $= 10\Delta x$ while the period is 1.43×10^{-5} s $\approx 4\Delta t$. Because finite difference evaluates $\partial_{tt}\mathbf{U}(i_x, i_y, i_t)$ based on slopes of the line segments connecting $\mathbf{U}(i_x, i_y, i_t)$ with $\mathbf{U}(i_x, i_y, i_t - 1)$ and $\mathbf{U}(i_x, i_y, i_t + 1)$, when Δt is not sufficiently small, these slopes can be far from the slope of the tangent line passing $\mathbf{U}(i_x, i_y, i_t)$, causing significant bias. In comparison, the mean speeds in Table III are not underestimated because of the nine points interpolation between each neighboring time step. The result is similar to a classic phase speed estimation based on Fourier transform by first finding the primary spatial frequency $\hat{\xi}$ for each frequency f and then $\hat{c}_{cl} = f/\hat{\xi}$,[43,44] assuming the PDE is a spatially independent wave equation and the waves are isotropic. The $\hat{c}_{cl}$ computed from clean measurements is also shown in Table III, indicating that the surface wave is strongly dispersive.

TABLE III. Success rate of the PDE identification from noisy measurements aided by integration transformation and the recovered speeds for various frequency bands. The correct **s** is from Eq. (24) and other rows in the dictionary are the normalized terms of Eq. (25).

Center frequency (kHz)	30	40	50	60	70
Success rate using correct **s** (%)	98.1	99.3	99.2	93.4	96.6
Mean speed (m/s)	473	548	613	675	731
MAD (m/s)	5.8	4.2	5.5	9.6	5.4
MAD/mean (%)	1.2	0.8	0.9	1.4	0.7
$\hat{c}_{cl}$ (m/s)	476	556	610	667	722

TABLE IV. The average CPU time (s) for ten trials on a MacBook Pro to correctly identify spatially independent PDEs on three datasets using SBL (Ref. 9), the CV based method (Ref. 10), exhaustive search (exhaust), and the proposed method.

	SBL (Ref. 9)	CV (Ref. 10)	Exhaust	Proposed
Burgers's equation	0.69	0.16	0.06	0.002
Non-attenuation wave equation	1.7	15	35	0.028
Attenuation wave equation	1.0	4.3	9.4	0.010

V. EFFICIENCY

We emphasize the efficiency of the proposed method by comparing its CPU time to the PDE identification using SBL,[9] CV based method,[10] and exhaustive search (exhaust). Because the methods in Refs. 9 and 10 only work for spatially independent PDEs, we use three datasets, describing such PDEs, for the experiments: (a) dataset in Sec. III A; (b) non-attenuating 2D wave equation (23) with $\alpha = 0$, $c = 2.5$ m/s; (c) attenuating 2D wave equation (23) with $\alpha = 0.025$, $c = 2.5$ m/s. All other settings for (b) and (c) are the same as those in Sec. III B.

As in Refs. 9 and 10, as the PDE is identical for all of the locations, we concatenate $\mathbf{\Phi}_n$ [defined by Eq. (3) or (24)] for each column to build the dictionary,

$$\mathbf{\Phi} = \left[\mathbf{\Phi}_1^{\mathrm{T}}, \mathbf{\Phi}_2^{\mathrm{T}}, \ldots, \mathbf{\Phi}_N^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{NM \times D}, \tag{26}$$

for Refs. 9 and 10 and exhaustive search, and the proposed method uses normalized Eq. (26) appended by **s** as its last row.

All of the methods successfully identify the PDEs, and the CPU time shows the superior efficiency of our proposed method as demonstrated in Table IV, where the proposed method outperforms others significantly. The exhaustive search is efficient for the dictionary with six terms but time-consuming for nine terms.

In addition to efficiency, the proposed method and the baselines in Table IV have similar performance for correctly identifying the PDEs. With the same success rate, the higher efficiency makes our approach suitable for recovering spatially dependent PDEs in a large ROI, in which the PDE identification is repeated for every spatial location.

VI. CONCLUSION

We proposed a technique to efficiently recover the spatial variations of physical properties via spatially dependent PDEs identification given observations and validated it by recovering various acoustical properties for the medium of propagating waves.

The identification employs a constrained ℓ_1-norm minimization, which encourages sparsity and is solved via Lasso, to select active PDE terms from a dictionary of all of the potential terms. It is computationally efficient as it does not require iterative assumptions of active PDE terms and the implementation of a fast computing scheme for Lasso. Using an integration transformation to transfer the derivatives on the noisy measurement to a smooth predefined function, the method can also identify spatially dependent PDEs from highly noisy measurements.

[1]S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Data-driven discovery of partial differential equations," Sci. Adv. **3**(4), e1602614 (2017).
[2]S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," Proc. Natl. Acad. Sci. **113**(15), 3932–3937 (2016).
[3]H. Schaeffer, G. Tran, and R. Ward, "Extracting sparse high-dimensional dynamics from limited data," SIAM J. Appl. Math. **78**(6), 3279–3295 (2018).
[4]M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," J. Comput. Phys. **378**, 686–707 (2019).
[5]Z. Long, Y. Lu, and B. Dong, "PDE-Net 2.0: Learning PDEs from data with a numeric-symbolic hybrid deep network," J. Comput. Phys. **399**, 108925 (2019).
[6]S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control* (Cambridge University Press, Cambridge, 2019).
[7]P. A. Reinbold, D. R. Gurevich, and R. O. Grigoriev, "Using noisy or incomplete data to discover models of spatiotemporal dynamics," Phys. Rev. E **101**(1), 010203 (2020).
[8]W. Zhou, H. Zhang, and J. Wang, "Sparse Bayesian learning based on collaborative neurodynamic optimization," IEEE Trans. Cybern. **52**, 13669–13683 (2022).
[9]S. Zhang and G. Lin, "Robust data-driven discovery of governing physical laws with error bars," Proc. Math. Phys. Eng. Sci. **474**(2217), 20180305 (2018).
[10]R. Liu, M. Bianco, and P. Gerstoft, "Automated partial differential equation identification," J. Acoust. Soc. Am. **150**(4), 2364–2374 (2021).
[11]S. K. Goharoodi, P. N. Phuc, L. Dupré, and G. Crevecoeur, "Data-driven discovery of the heat equation in an induction machine via sparse regression," in IEEE International Conference on Industrial Technology (2019).
[12]D. Bhattacharya, L. K. Cheng, and W. Xu, "Sparse machine learning discovery of dynamic differential equation of an esophageal swallowing robot," IEEE Trans. Ind. Electron. **67**, 4711–4720 (2020).
[13]R. Liu, M. J. Bianco, and P. Gerstoft, "Wave equation extraction from a video using sparse modeling," in *Proceedings of the 53th Asilomar Conference on Circuits, Systems and Computers*, Pacific Grove, CA (IEEE, New York, 2019), pp. 2160–2165.
[14]M. J. Bianco and P. Gerstoft, "Travel time tomography with adaptive dictionaries," IEEE Trans. Comput. Imag. **4**(4), 499–511 (2018).
[15]M. van Berkel, G. Vandersteen, E. Geerardyn, R. Pintelon, H. Zwart, and M. de Baar, "Frequency domain sample maximum likelihood estimation for spatially dependent parameter estimation in PDEs," Automatica **50**(8), 2113–2119 (2014).
[16]S. Kramer and E. M. Bollt, "Spatially dependent parameter estimation and nonlinear data assimilation by autosynchronization of a system of partial differential equations," Chaos **23**(3), 033101 (2013).
[17]F. C. Lin, M. H. Ritzwoller, and R. Snieder, "Eikonal tomography: Surface wave tomography by phase front tracking across a regional broad-band seismic array," Geophys. J. Int. **177**(3), 1091–1110 (2009).
[18]I. D. Khurjekar and J. B. Harley, "Closing the sim-to-real gap in guided wave damage detection with adversarial training of variational autoencoders," in *Proceedings of the 2022 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, Singapore (IEEE, New York, 2022), pp. 3823–3827.
[19]C. Huang, K. Wang, R. W. Schoonover, L. V. Wang, and M. A. Anastasio, "Joint reconstruction of absorbed optical energy density and sound speed distributions in photoacoustic computed tomography: A numerical investigation," IEEE Trans. Comput. Imag. **2**(2), 136–149 (2016).
[20]P. S. Fuchs, S. Mandija, P. R. S. Stijnman, W. M. Brink, C. A. T. van den Berg, and R. F. Remis, "First-order induced current density imaging and electrical properties tomography in MRI," IEEE Trans. Comput. Imag. **4**(4), 624–631 (2018).
[21]R. Liu, M. Bianco, P. Gerstoft, and B. D. Rao, "Data-driven spatially dependent PDE identification," in *Proceedings of the 2022 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, Singapore (IEEE, New York, 2022), pp. 3383–3387.
[22]J. Rubinstein, "Sine-gordon equation," J. Math. Phys. **11**(1), 258–266 (1970).
[23]M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, "Method for solving the sine-Gordon equation," Phys. Rev. Lett. **30**(25), 1262–1264 (1973).
[24]D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," IEEE Trans. Signal Process. **53**(8), 3010–3022 (2005).
[25]D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," Proc. Natl. Acad. Sci. **100**(5), 2197–2202 (2003).
[26]A. Xenaki, P. Gerstoft, and K. Mosegaard, "Compressive beamforming," J. Acoust. Soc. Am. **136**(1), 260–271 (2014).
[27]B. D. Rao, "Signal processing with the sparseness constraint," in *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP*, Seattle, WA (IEEE, New York, 1998), pp. 1861–1864.
[28]R. G. Baraniuk, "Compressive sensing [Lecture Notes]," IEEE Signal Process. Mag. **24**(4), 118–121 (2007).
[29]R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," IEEE Trans. Inf. Theory **56**(4), 1982–2001 (2010).
[30]A. M. Salman and J. Romberg, "Sparse recovery of streaming signals using L_1-homotopy," IEEE Trans. Signal Process. **62**(16), 4209–4223 (2014).
[31]M. J. Buckingham, "On the transient solutions of three acoustic wave equations: van Wijngaarden's equation, Stokes' equation and the time-dependent diffusion equation," J. Acoust. Soc. Am. **124**(4), 1909–1920 (2008).
[32]R. Tibshirani, "Regression shrinkage and selection via the lasso," J. R. Statist. Soc. Ser. B **58**(1), 267–288 (1996).
[33]M. Wainwright, "Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1-constrained quadratic programming (Lasso)," IEEE Trans. Inf. Theory **55**(5), 2183–2202 (2009).
[34]C. F. Mecklenbräuker, P. Gerstoft, A. Panahi, and M. Viberg, "Sequential Bayesian sparse signal reconstruction using array data," IEEE Trans. Signal Process. **61**(24), 6344–6354 (2013).
[35]P. Gerstoft, A. Xenaki, and C. F. Mecklenbräuker, "Multiple and single snapshot compressive beamforming," J. Acoust. Soc. Am. **138**(4), 2003–2014 (2015).
[36]C. F. Mecklenbräuker, P. Gerstoft, and E. Zöchmann, "c–LASSO and its dual for sparse signal estimation from array data," Signal Process **130**(4), 204–216 (2017).
[37]J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," J. Stat. Softw. **33**(1), 1–22 (2010).
[38]W. F. Ames, *Numerical Methods for Partial Differential Equations* (Academic, New York, 2014).
[39]Y. C. Hon and X. Z. Mao, "An efficient numerical scheme for Burgers' equation," Appl. Math. Comput. **95**(1), 37–50 (1998).
[40]H. Mitome, "An exact solution for finite-amplitude plane sound waves in a dissipative fluid," J. Acoust. Soc. Am. **86**(6), 2334–2338 (1989).
[41]K. S. Alguri, J. Melville, and J. B. Harley, "Baseline-free guided wave damage detection with surrogate data and dictionary learning," J. Acoust. Soc. Am. **143**(6), 3807–3818 (2018).
[42]S. M. Ziola and M. R. Gorman, "Source location in thin plates using cross-correlation," J. Acoust. Soc. Am. **90**(5), 2551–2556 (1991).
[43]D. Alleyne and P. Cawley, "A two-dimensional fourier transform method for the measurement of propagating multimode signals," J. Acoust. Soc. Am. **89**(3), 1159–1168 (1991).
[44]D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques* (Simon and Schuster, Inc., New York, 1992).