Deep learning enables accurate sound redistribution via nonlocal metasurfaces

Hua Ding  
Tongji University

Xinsheng Fang  
Tongji University

Bin Jia  
Tongji University

Nengyin Wang  
Tongji University

Qian CHENG  
Tongji University  https://orcid.org/0000-0003-3621-5743

Yong Li (✉ yongli@tongji.edu.cn)  
Tongji University  https://orcid.org/0000-0001-8049-9128

Article

Keywords: Sound Propagation Control, Mutual Coupling of Subunits, Incident Wave, Multi-channel Reflections, Acoustic Functional Devices

DOI: https://doi.org/10.21203/rs.3.rs-639770/v1

License: ☺️ ☀️ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Deep learning enables accurate sound redistribution via nonlocal metasurfaces

Hua Ding,1 Xinsheng Fang,1 Bin Jia,1 Nengyin Wang,1 Qian Cheng,1,* and Yong Li1, †

1 Institute of Acoustics, School of Physics Science and Engineering, Tongji University, Shanghai 200092, China
(Dated: June 19, 2021)

Abstract

Conventional acoustic metasurfaces are constructed with gradiently “local” phase shift profiles provided by subunits. The local strategy implies the ignorance of the mutual coupling between subunits, which limits the efficiency of targeted sound manipulation, especially in complex environments. By taking into account the “nonlocal” interaction among subunits, nonlocal metasurface offers an opportunity for accurate control of sound propagation, but the requirement of the consideration of gathering coupling among all subunits, not just the nearest-neighbor coupling, greatly increases the complexity of the system and therefore hinders the explorations of functionalities of nonlocal metasurfaces. In this work, empowered by deep learning algorithms, the complex gathering coupling can be learned efficiently from the preset dataset so that the functionalities of nonlocal metasurfaces can be significantly uncovered. As an example, we demonstrate that nonlocal metasurfaces, which can redirect an incident wave into multi-channel reflections with arbitrary energy ratios, can be accurately predicted by deep learning algorithms. Compared to the theory, the relative error of the energy ratios is less than 1%. Furthermore, experiments witness three-channel reflection with three types of energy ratios of (1, 0, 0), (1/2, 0, 1/2), and (1/3, 1/3, 1/3), proving the validity of the deep learning enabled nonlocal metasurfaces. Our work might blaze a new trail in the design of acoustic functional devices, especially for the cases containing complex wave-matter interactions.
The emergence of acoustic metamaterials, which are artificial materials with targeted functionalities, have significantly broaden the fields of physical and material science. Novel acoustic manipulations based on metamaterials have been realized, such as acoustic focusing, acoustic cloaking, acoustic holography, and asymmetric acoustic transmission. In the last decade, acoustic metasurfaces, as 2D acoustic metamaterials, stand out as distinct choices to manipulate sound for their vanishing thickness. By employing “locally” designed phase shift provided by subunits of metasurfaces, fascinating acoustic manipulations are extensively explored recently, including acoustic cloaking, acoustic holography, and compact sound absorbers. However, in the most of previous designs, the subunits are designed individually according to the discretized phase shift profile. In this way, the mutual coupling among the subunits which also contributes the wave-field forming is neglected. It therefore turns out that the designed metasurfaces suffer from the lower efficiency, especially in the cases where incident/reflected/transmitted angles are relatively large. Furthermore, to reshape wave field with fine resolution, a large number of subunits are necessary which inevitably introduce undesired loss due to the viscou-thermal effect induced by small apertures in small subunits. Recently, the concept of “nonlocal” metasurface is introduced to overcome these limitations. In these works, the nonlocal coupling effects among all subunits are taken into accounted to collect the contribution of nonlocality to the wave-field forming and therefore give rise to more accurate wave field manipulations. However, it also makes the nonlocal system quite complex and the relationship between desired wave field and the geometrical parameters unclear. In this situation, optimization methods are generally employed to inversely design the functionality of nonlocal metasurface. However, for different wave manipulation aims, the objective function needs to be revised correspondingly to search for the best value in the entire parameter space, resulting in the drawbacks of time-wasting and computational resource-consuming.

The development of artificial intelligence algorithms provides new insights into the inverse design of nonlocal metasurfaces. In view of its extreme success in domains correlated to the detection of diseases, speech recognition, computer vision, and natural language processing, deep learning has aroused wide attention. The excellent learning and inference capabilities of deep learning enable it to achieve the similar effect of logical calculations with physical formulas. Many works about combining deep learning with physics have been attempted, including optimization of photonic structures, high-resolution imaging, and acoustic source imaging. In this work, by harnessing the power of deep learning, we explore the capabilities of reshap-
FIG. 1. The schematic diagram of nonlocal acoustic metasurface. The metasurface is composed of repeated supercell along $x$ direction with a period of $a$. Three grooves are etched in each supercell. $t_l$ and $h_l$ represent the width and the depth of the $l$th groove, and $dx_l$ refers to the distance between adjacent grooves ($l = 1, 2, 3$). The incident wave along the $-y$ direction could be reflected to three allowed diffracted beams.
velocity field \(v(x, y)\) in the region above the structure (Region I) can be expressed by

\[
p_1(x, y) = A_0^+ e^{j k_0 y} + \sum_n A_n^- e^{-j k_{x,n} x} e^{-j k_{y,n} y},
\]

\[
v_1(x, y) = -\frac{A_0^+}{\rho \omega} k_0 e^{-j k_0 y} + \sum_n \frac{A_n^-}{\rho \omega} k_{y,n} e^{-j k_{x,n} x} e^{-j k_{y,n} y},
\]

where \(A_0^+, A_n^-\) are the amplitude of the incident wave and the \(n\)-th order of the reflected diffraction beams. \(k_{x,n} = k_0 \sin \theta_n + G_n\), and \(k_{y,n} = \sqrt{k_0^2 - k_{x,n}^2}\) are the \(x\) and \(y\) components of the \(n\)-th order diffraction wave vector with \(k_0\) being the wave number of incident waves and \(G_n = n 2\pi / a\) representing the reciprocal lattice vector. \(\rho\) and \(\omega\) refer to the mass density of the medium and the angular frequency. From the above formulas, the period \(a\) of the metasurface related to \(k_{x,n}\) decides the propagation characteristics of different diffraction modes. It means that by selecting a suitable \(a\), the numbers of the allowed propagating diffraction orders can be tuned.

To illustrate the nonlocal coupling effect among etched grooves, we need first get the pressure and velocity field inside each grooves. Considering the fact that the bottom of the grooves are acoustically hard where the normal velocity should be zero, the \(p(x, y)\) and \(v(x, y)\) in grooves (Region II) could be represented as

\[
p_{II}^l(x, y) = \sum_k a_{kl} \cos[\alpha_{kl}(x - x_l)](e^{j \beta_{kl} y} + e^{-j \beta_{kl}(y+2h_l)}),
\]

\[
v_{II}^l(x, y) = \sum_k -\frac{a_{kl}}{\rho \omega} \beta_{kl} \cos[\alpha_{kl}(x - x_l)](e^{j \beta_{kl} y} - e^{-j \beta_{kl}(y+2h_l)}),
\]

where \(a_{kl}\) represents the amplitude of the \(k\)-th order of the waveguide mode, \(x_l\) is the beginning \(x\) coordinate of the \(l\)-th groove, \(\alpha_{kl} = k\pi / t_l\) and \(\beta_{kl} = \sqrt{k_0^2 - \alpha_{kl}^2}\) are the \(x\) and \(y\) components of the wave vector for the \(k\)-th order mode in the groove, respectively. Conforming to the conditions at the air-metasurface interface, sound pressure and volume velocity should be continuous at \(y = 0\).

By plugging Eqs. (1)-(4) into continuum conditions and integrating the equations with orthogonal relationship, we can get the following equations:

\[
\begin{pmatrix}
  P_1^1 \\
  P_2^2 \\
  P_3^3
\end{pmatrix}
A_0^+ + \begin{pmatrix}
  P_1^1 \\
  P_2^2 \\
  P_3^3
\end{pmatrix} A_n^- = \begin{pmatrix}
  P_3^1 & 0 & 0 \\
  0 & P_3^2 & 0 \\
  0 & 0 & P_3^3
\end{pmatrix}
\begin{pmatrix}
  a_{kl}^1 \\
  a_{kl}^2 \\
  a_{kl}^3
\end{pmatrix},
\]

\[
V_1 A_0^+ + V_2 A_n^- = \begin{pmatrix}
  V_1^1 & V_2^2 & V_3^3
\end{pmatrix}
\begin{pmatrix}
  a_{kl}^1 \\
  a_{kl}^2 \\
  a_{kl}^3
\end{pmatrix},
\]
where

\[ P_1^l(k) = \frac{1}{l_l} \int_{x_l}^{x_l+l_l} \cos \alpha_{kl}(x-x_l) \, dx, \]  

(7)

\[ P_2^l(k,n) = \frac{1}{l_l} \int_{x_l}^{x_l+l_l} e^{-jk_x x} \cos \alpha_{kl}(x-x_l) \, dx, \]  

(8)

\[ P_3^l(k_1,k_2) = (1+e^{-j2\beta k_1 t_h}) \frac{1}{l_l} \int_{x_l}^{x_l+l_l} \cos \alpha_{k_1l}(x-x_l) \cos \alpha_{k_2l}(x-x_l) \, dx, \]  

(9)

\[ V_1(m) = -\frac{1}{a} \int_0^a k_{x,0} e^{jk_x mx} \, dx, \]  

(10)

\[ V_2(m,n) = \frac{1}{a} \int_0^a k_{x,n} e^{-j(k_{x,n}-k_{x,m})x} \, dx, \]  

(11)

and

\[ V_3^l(m,k) = -(1+e^{-j2\beta k_1 t_h}) \frac{B_{kl}}{a} \int_{x_l}^{x_l+l_l} \cos \alpha_{kl}(x-x_l)e^{jk_x mx} \, dx, \]  

(12)

with \( m,n = 0, \pm 1, \cdots, \pm (N-1)/2 \) and \( k,k_1,k_2 = 1, \cdots, K \). \( P_1^l, P_2^l, P_3^l \) are respectively related to the the sound pressure of incident wave, reflected wave in Region I and the sound wave in Region II. And their dimensions are known as \( LK \times 1 \), \( LK \times N \) and \( LK \times LK \). \( L \) is the total number of grooves in a period, \( K \) is the total number of modes inside the metasurface, and \( N \) is the total number of diffraction modes. In fact, the superscript and subscript of \( V \) have the similar meaning as those of \( P \), but \( V \) is related to the volume velocity of sound waves. The dimensions of \( V_1, V_2, V_3^l \) are \( N \times 1 \), \( N \times N \), \( N \times LK \) respectively. We simplify the above equations with matrixes and rewrite them as:

\[
\begin{pmatrix}
  P_1 \\
  V_1
\end{pmatrix} A_0^+ =
\begin{pmatrix}
  -P_2 & P_3 \\
  -V_2 & V_3
\end{pmatrix}
\begin{pmatrix}
  A^+_\mu \\
  a_{kl}
\end{pmatrix},
\]

(13)

From Eq. (13), the reflected energy of each diffraction order could be calculated. In order to show energy distribution intuitively, we utilize \( I = \left| \frac{A_{\mu}}{A_0} \right|^2 \cos \theta_r \) to denote the power flow ratio of each mode. Here, we choose -1st, 0th and +1st order diffractive components as reflected channels, and the corresponding diffraction angles for them are \( \theta_r = -60^\circ, 0^\circ, 60^\circ \). The working frequency is set to \( f_0 = 8000 \text{Hz} \). In real situations, the harmonic modes of the metasurface and its internal modes are very complicated, so we choose \( K=20 \), \( N=51 \), \( L=3 \) to be as close to the actual situations as possible. Obviously, the consideration of non-locality makes the dimension of the matrixes large which means that the computational complexity is higher. This brings more troubles to inverse design.

**Deep learning model.** The special operating mechanism of deep learning will help to explore the relationship between metasurface and its response to acoustic waves. Here, we establish a
FIG. 2. Schematic representation of data acquisition and the tandem network. This network is composed of pre-trained forward network $A_{net}$ (in the red dotted box) and inverse network $B_{net}$ (in the blue dotted box). The intermediate layer where the blue nodes are located is the output layer of the inverse design.

To realize the intelligent design of metasurfaces, we firstly should establish a data set. We introduce a vector $S = (h_1, h_2, h_3, t_1, t_2, t_3, dx_1, dx_2)^T$ which is composed of the structure parameters of metasurfaces. Meanwhile, reflected power flow ratio in three channels is expressed as a state vector $I = (I_1, I_2, I_3)^T$, where $I_1, I_2, I_3$ respectively correspond to power flow of $-1$st, 0th and $+1$st order diffractive components. The dataset based on diffraction theory for metasurfaces contains many arrays $\{S, I\}$, including 270,000 training samples and 30000 test samples. Specifically, as shown in Fig. 2, we randomly generate many different sets of $S$ and calculate their corresponding $I$ by MATLAB. Here, we construct two network models ($A_{net}$ and $B_{net}$), one of which takes $S$ as input data and $I$ as label data while the other is the opposite. Obviously, there is a correspondence between these two vectors, $I = A_{net}(S), S = B_{net}(I)$, where $A_{net}$ is the forward deep learning network and $B_{net}$ is the inverse neural network. It should be noted that $B_{net}$ can hardly converge when it is trained in a similar way as $A_{net}$ since the same power-flow distribution could be accounted for.
different arrangement of meta-units.

To tackle this multi-value problem, we exploit tandem neural network introduced in Ref.47. As Fig. 2 shows, the tandem network model is composed of $A_{\text{net}}$ and $B_{\text{net}}$, which connects the trained $A_{\text{net}}$ served as a backend behind the $B_{\text{net}}$. This connected network is similar to an autoencoder network (consists of two networks: encoder and decoder) in terms of structures, in which the input data in the first layer has the same physical meaning as the data in the final layer and structural parameters exist as the encoded feature data in the intermediate layer. The entire network trained successfully is actually a transient reactor. In the construction of it, the forward network with dataset \{S, I\} is supposed to be trained as the pre-trained network firstly. Then, when training the connected model, I is both the input data and the target value that needs to be compared with the output data. Different from the general autoencoder training mechanism which is often used for unsupervised learning, we have a prepared dataset including label data, so we can train the “decoder” ($A_{\text{net}}$) first. Training part of the network in advance transforms the surjective relation between S and I into bijective relation, and in this way, the root cause of the failure to converge inverse network is resolved.

For all training processes, Adaptive Moment Estimation (Adam) optimizer\(^5\) is chosen as the global optimizer. Forward training is done by minimizing the cost function, $C_1 = \frac{1}{N} \sum |I_{\text{true}} - I_{\text{pred}}|$, in which $N$ is the batch size of training dataset, $I_{\text{pred}}$ is the possible power flow ratio predicted by $A_{\text{net}}$ and $I_{\text{true}}$ is the target value in dataset. The value of cost function is inclined to be steady after 1500 epochs of forward training. The successful training of $A_{\text{net}}$ informs that deep learning model could acquire the relationship between physical quantities. During the secondary training procedure, every layer in pre-trained forward model is frozen, that is, weights and bias of each node are not updated in layers of $A_{\text{net}}$ but these values in $B_{\text{net}}$ are optimized all the time. The cost function of this combination model can be expressed by $C_2 = \frac{1}{N} \sum |I_{\text{true}} - A_{\text{net}}(S)|$. In a way, the “decoder” converts the multi-value problem into a single-value problem without complicated data filtering. As the corresponding relation between the design S and the response I changes, the training of connected network converges easily (nearly after 1500 epochs).

With the entire model trained successfully, the transient reactor for three-channel metasurfaces is completely constructed. When we take the middle layer as input and the last layer as the output, it could realize the forward data prediction. While we use the first layer as the input and the middle layer as the output, the desired structure parameters would be predicted by the middle layer, i.e., the output layer of inverse network, as long as we input the target power flow distribution.
Actually, what we are concerned more about is the function implemented by the latter, because nonlocal metasurfaces do not have a clear inverse theoretical relationship between structures and their responses, but the realization of the model allows this theoretical relationship to be simulated. Moreover, it avoids the troubles of conventional optimization methods that require to iterate many times to search for the best value and easily fall into local optimality.

**Nonlocal metasurfaces predicted by deep learning.** The target structures of acoustic metasurfaces could be predicted quickly and intelligently by the transient reactor. We randomly select several groups of power flow distributions with different ratios to verify the accuracy and efficiency of this reactor. Part of the prediction results is shown in Table I. The values on the left indicate the target ratios fed to the model and the theoretical values in middle represent the real energy distributions calculated by the predicted geometric parameters as shown in the right column. All the ratios agree well with each other, which reveals the accuracy of this reactor.

For further verification, we choose three sets of predicted data for simulation and experiment. These three target ratios are set to be $\mathbf{I}_I = [1, 0, 0]$, $\mathbf{I}_II = \left[ \frac{1}{2}, 0, \frac{1}{2} \right]$, $\mathbf{I}_III = \left[ \frac{1}{4}, \frac{1}{2}, \frac{1}{4} \right]$. For these cases, the structures designed by the transient reactor are shown as follows (unit: mm):

$$
\mathbf{S}_I = [6.80, 23.52, 11.74, 8.19, 6.41, 7.06, 8.09, 8.19],
$$
$$
\mathbf{S}_II = [10.10, 23.01, 10.12, 8.30, 7.99, 8.46, 8.37, 8.37],
$$
$$
\mathbf{S}_III = [11.53, 19.96, 11.53, 8.44, 7.80, 8.44, 8.46, 8.47].
$$

Corresponding power flow ratios in theory are $\mathbf{I}'_I = [0.9910, 0.0026, 0.0063]$, $\mathbf{I}'_II = 0.4986, 0.0017, 0.4998]$, $\mathbf{I}'_III = [0.3341, 0.3315, 0.3344]$, which fit well with the target values. The finite
FIG. 3. Intensity distribution in the total field. (a)-(c) Power flow value for $S_I$, $S_{II}$, and $S_{III}$ along $x$ direction at the surface of metasurface. The size and direction of white arrows in (d)-(f) respectively indicate the strength and direction of the power flow here.

As shown in Fig. 1 and each matrix $S$, along with these rectangular pits, which are attached randomly in one period, comes the non-local effect from metasurfaces. Compared with traditional gradient metasurface, additional degree of freedom for sound regulation could be utilized in these metasurfaces, which supports the lateral transport of energy flow. To demonstrate more details of non-local coupling, the energy flow at $y = 0$ for these three metasurfaces in simulation has been extracted in Figs. 3a-c. The local intensity vector $I_x$ could be obtained by $I_x = \frac{1}{2} \text{Re}[\rho(v_x)^*]$, where $\rho$ means pressure and the $v_x$ is the local velocity in the $x$ direction. The sign of $I_x/I_0$ refers to the direction of energy flow in $x$-axis. From Fig. 3a, $I_x/I_0$ are almost negative, which suggests the sound energy is concentrated and reflected in the left channel. In contrast, the curves in Figs. 3b-c float above and below zero, which indicates that the structures $S_{II}$, and $S_{III}$ not only allocate energy to the left channel but also allocate to the others. Figs. 3d-f confirm this point more intuitively, corresponding to Figs. 3a-c, respectively, which reveal the strong lateral energy exchange between grooves. In any case, metasurfaces via non-local effects realize efficient wave manipulation by using lateral energy transport in the form of etched grooves. Although the
efficiency of non-local metasurfaces could be perfectly exerted in sound manipulation, the inverse calculation is complicated due to the randomness of its geometrical parameters. Herein, deep learning vastly simplifies it and facilitates the mass production for complex meta-structures. Not only that, it enables designers without theoretical foundations to quickly implement on-demand design.

**Experimental demonstrations.** When a plane wave fed perpendicularly into the metagratings, scattering field with different ratios are represented in Figs. 4a-c, corresponding to $S_{I}-S_{III}$ respectively. The red arrows in the figure indicate the position and direction of the incident wave whilst white arrows represent reflection. The reflected sound waves are obviously emitted from three channels we set, which follow the target energy distribution.

The experimental samples with 20 periods are fabricated by 3D-printing technology using the stereo lithography apparatus with photosensitive resin. Each measurement area is marked with a red dashed frame in the upper panels of Figs. 4a-c and the size is $200 \times 200 \text{ mm}^2$. The experimental results shown in the middle of Fig. 4 correspond to the scattering fields with target ratios ($I_{I}-I_{III}$). All the scattering fields in experiment have been normalized by the incident pressure. In addition, we select a line in each measurement area and plot the data extracted from the line in lower panels, defining $R(\theta) = \text{Re}(p_\theta/p_{\text{max}})$ as the abscissa. Experimental results agree well with the simulation, which demonstrates the fantastic functionality of the metasurfaces designed by deep learning algorithm. In the last experimental field, the pressure field of the middle channel is weaker than the two others, which may be caused by the multiple reflection between loudspeaker array and the metasurface.

**Discussion**

In this work, we investigate the beam splitting function of the grooved metasurfaces based on nonlocal coupling mechanism and realize the arbitrary regulation of the energy ratio for the multi-channel metasurfaces. We successfully conducted model training on $S$ and $I$ taking into account the nonlocality among subunits, so that finally the transient reactor can predict the target structures. Meanwhile, the efficiency and accuracy of the network model are validated by simulations and experiments. This intelligent method fairly shortens the optimization time for acoustic device design and this way could be extended to other forms of device structures. Our work provides new design solutions for non-local acoustic functional devices and promotes explorations of more fancy properties of meta-materials. These results might inspire more combinations of acoustics and artificial intelligence algorithms to assist in solving acoustic problems.
FIG. 4. Finite element simulation and experimental results for scattered pressure fields. (a)-(c) Scattered fields corresponding to $S_I$, $S_{II}$, $S_{III}$. The upper panels are the results of the simulation. To show the reflected fields more clearly, incident waves are not shown in the figures. Middle panels are experimental measurement fields, in which each measured area is a square with a size of $200 \times 200$ mm$^2$. The insets in middle panels are 3D-printed sample photos (one period is shown). The values at black dashed lines in the middle panels are extracted and plotted in the lower panels, where the abscissa $R(\theta)$ is the ratio of the real part for $p_\theta$ and $p_{\text{max}}$. $p_\theta$ represents the pressure value at different positions on the black dotted lines.

Method

Neural network training mechanism. Given that one network model has $D$ layers, then the first layer gets the input data $X^0$ and the layer $D$ outputs the data $X^D$ which has been activated many times by nonlinear activation functions. $f$ represents the activation function. The complete neural network training process is mainly composed of two parts. In the former process, neural network parameters $w$ and $b$ are used to calculate the output values in conjunction with the activation function. Forward equations are as follows:

$$z^d = w^d X^{d-1} + b^d \quad \text{for} \quad d = 1, 2, \cdots, D,$$

$$X^d = f(z^d) \quad \text{for} \quad d = 1, 2, \cdots, D,$$

11
where \( z \) called weighted input will play an important role in the back propagation. Cost function \( C \) evaluates the error between predicted values and expected values. Actually, the goal of training given data is to seek suitable model parameters for which cost function return as small value as possible. Whereas this purpose depends on the later process to complete. In the back propagation, parameters \( w \) and \( b \) could be updated by gradient calculation and the global optimizer to improve the predictive performance of models. Referring to the back propagation algorithm proposed in Ref. \(^{58}\), error gradient of each layer can be expressed as:

\[
\delta^D = \frac{\partial C}{\partial X^D} f'(z^D),
\]

\[
\delta^d = [(w^{d+1})^T \delta^{d+1}] \odot f'(z^d) \quad for \quad d = 1, 2, \cdots, D-1,
\]

Then the partial differential of the cost function to the internal parameters of the network can be written as:

\[
\frac{\partial C}{\partial w^d} = a^{d-1} \delta^d,
\]

\[
\frac{\partial C}{\partial b^d} = \delta^d.
\]

Note that every neuron will go through these calculation cycles. Considering that the calculation relationship between two neurons located in adjacent layers is similar, for the sake of simplicity, we just use the superscript representing the number of layer for these parameters in above equations instead of marking the specific position of each neuron in a certain layer.

Network model structure and hyperparameter settings. In order to balance the training speed and prediction accuracy of the network model, we adopt the batch training method, that is, part of the data is sent for training each time. Batch size is the number of samples sent for each training. In all our training process, batch size is set to 256. We use the Leaky Relu as the activation function for each layer. The initial learning rates of forward training and tandem training are 0.0014 and 0.0006, respectively. The architecture of \( A_{net} \) is selected as 8-512-256-256-128-3 and \( B_{net} \) is 3-256-256-128-128-8. The 8 nodes match the number of structure parameters \( S \), and the 3 nodes correspond to the number of elements in the vector \( I \). Although pre-trained model and connected model could converge well after training over a thousand times, we terminated all the model training until epochs reached 5000 times to make the models more stable.

Finite element simulation. COMSOL MULTIPHYSICS is used as a preliminary verification tool for preliminary simulation verification of the predicted results. The pressure acoustic module
in the frequency domain is used to simulate the effect of the metasurface. A fixed-length plane
incident wave as the sound source is incident on metasurfaces.

**Experimental setup.** The phase shift and amplitude of pressure fields are collected by a 1/8-inch
microphone (Brüel & Kjær type 2670) and the data-acquisition hardware (NI PXI-4461 in NI
PXIe-1071). The microphone as a probe scans fields in a step of 5mm. In the automatic scanning
platform, we disposed 7 loudspeakers with a total width of 245 mm as the source. Samples are
put in an organic glass plate waveguide about 2 cm high for experimental measurements. Foams
are distributed at the edge of the measurement field to absorb external sound waves and reduce the
influence of internal multiple reflection waves. We first measured the empty field with no sample
placed, and then measured the total field with the placed sample. In this way, the difference
between the two measurements is the result of the reflected field.

**Data availability**

The data that support the plots within this paper and other findings of this study are available from
the corresponding author upon reasonable request.

---

* q.cheng@tongji.edu.cn
† yongli@tongji.edu.cn

1 Liu, Z. *et al.* Locally resonant sonic materials. *Science* **289**, 1734–1736 (2000).
2 Fang, N. *et al.* Ultrasonic metamaterials with negative modulus. *Nat. Mater.* **5**, 452–456 (2006).
3 Yang, Z., Mei, J., Yang, M., Chan, N. H. & Sheng, P. Membrane-type acoustic metamaterial with
    negative dynamic mass. *Phys. Rev. Lett.* **101**, 204301 (2008).
4 Liang, Z. & Li, J. Extreme acoustic metamaterial by coiling up space. *Phys. Rev. Lett.* **108**, 114301
    (2012).
5 Kaina, N., Lemoult, F., Fink, M. & Lerosey, G. Negative refractive index and acoustic superlens from
    multiple scattering in single negative metamaterials. *Nature* **525**, 77–81 (2015).
6 Cummer, S. A., Christensen, J. & Alu, A. Controlling sound with acoustic metamaterials. *Nat. Rev.
    Mater.* **1**, 16001 (2016).
7 Li, J., Fok, L., Yin, X., Bartal, G. & Zhang, X. Experimental demonstration of an acoustic magnifying
    hyperlens. *Nat. Mater.* **8**, 931–934 (2009).
8 Zhu, J. *et al.* A holey-structured metamaterial for acoustic deep-subwavelength imaging. *Nat. Phys.* **7**,
Chen, J., Xiao, J., Lisevych, D., Shakouri, A. & Fan, Z. Deep-subwavelength control of acoustic waves in an ultra-compact metasurface lens. *Nat. Commun.* **9**, 4920 (2018).

Popa, B. I., Zigoneanu, L. & Cummer, S. A. Experimental acoustic ground cloak in air. *Phys. Rev. Lett.* **106**, 253901 (2011).

Zhu, X., Liang, B., Kan, W., Zou, X. & Cheng, J. Acoustic cloaking by a superlens with single-negative materials. *Phys. Rev. Lett.* **106**, 014301 (2011).

Zhang, S., Xia, C. & Fang, N. Broadband acoustic cloak for ultrasound waves. *Phys. Rev. Lett.* **106**, 024301 (2011).

Marzo, A. *et al.* Holographic acoustic elements for manipulation of levitated objects. *Nat. Commun.* **6**, 8661 (2015).

Melde, K., Mark, A. G., Qiu, T. & Fischer, P. Holograms for acoustics. *Nature* **537**, 518–522 (2016).

Zhu, Y. *et al.* Fine manipulation of sound via lossy metamaterials with independent and arbitrary reflection amplitude and phase. *Nat. Commun.* **9**, 1632 (2018).

Liang, B., Yuan, B. & Cheng, J. C. Acoustic diode: rectification of acoustic energy flux in one-dimensional systems. *Phys. Rev. Lett.* **103**, 104301 (2009).

Liang, B., Guo, X. S., Tu, J., Zhang, D. & Cheng, J. C. An acoustic rectifier. *Nat. Mater.* **9**, 989–992 (2010).

Fleury, R., Sounas, D. L., Sieck, C. F., Haberman, M. R. & Alu, A. Sound isolation and giant linear nonreciprocity in a compact acoustic circulator. *Science* **343**, 516–519 (2014).

Devaux, T., Tournat, V., Richoux, O. & Pagneux, V. Asymmetric acoustic propagation of wave packets via the self-demodulation effect. *Phys. Rev. Lett.* **115**, 234301 (2015).

Li, Y., Liang, B., Gu, Z. M., Zou, X. Y. & Cheng, J. C. Reflected wavefront manipulation based on ultrathin planar acoustic metasurfaces. *Sci. Rep.* **3**, 2546 (2013).

Zhao, J., Li, B., Chen, Z. & Qiu, C. W. Manipulating acoustic wavefront by inhomogeneous impedance and steerable extraordinary reflection. *Sci. Rep.* **3**, 2537 (2013).

Tang, K. *et al.* Anomalous refraction of airborne sound through ultrathin metasurfaces. *Sci. Rep.* **4**, 6517 (2014).

Li, Y. *et al.* Experimental realization of full control of reflected waves with subwavelength acoustic metasurfaces. *Phys. Rev. Applied* **2**, 064002 (2014).

Xie, Y. *et al.* Wavefront modulation and subwavelength diffractive acoustics with an acoustic metasur-
face. Nat. Commun. 5, 5553 (2014).

25 Mei, J. & Wu, Y. Controllable transmission and total reflection through an impedance-matched acoustic metasurface. New J. Phys. 16, 123007 (2014).

26 Ma, G., Yang, M., Xiao, S., Yang, Z. & Sheng, P. Acoustic metasurface with hybrid resonances. Nat. Mater. 13, 873–878 (2014).

27 Li, Y., Jiang, X., Liang, B., Cheng, J. C. & Zhang, L. K. Metascreen-based acoustic passive phased array. Phys. Rev. Applied 4, 024003 (2015).

28 Li, Y. et al. Tunable asymmetric transmission via lossy acoustic metasurfaces. Phys. Rev. Lett. 119, 035501 (2017).

29 Zhu, Y. F., Fan, X. D., Liang, B., Cheng, J. C. & Jing, Y. Ultrathin acoustic metasurface-based Schroeder diffuser. Phys Rev X 7, 021034 (2017).

30 Li, J., Shen, C., DÃ­az-Rubio, A., Tretyakov, S. A. & Cummer, S. A. Systematic design and experimental demonstration of bianisotropic metasurfaces for scattering-free manipulation of acoustic wavefronts. Nat. Commun. 9, 1342 (2018).

31 Assouar, B. et al. Acoustic metasurfaces. Nat. Rev. Mater. 3, 460–472 (2018).

32 Fu, Y. et al. Sound vortex diffraction via topological charge in phase gradient metagratings. Sci. Adv. 6, eaba9876 (2020).

33 Jin, Y., Fang, X., Li, Y. & Torrent, D. Engineered diffraction gratings for acoustic cloaking. Phys. Rev. Appl. 11, 011004 (2019).

34 Mei, J. et al. Dark acoustic metamaterials as super absorbers for low-frequency sound. Nat. Commun. 3, 756 (2012).

35 Yang, M., Chen, S., Fu, C. & Sheng, P. Optimal sound-absorbing structures. Mater. Horizons 4, 673–680 (2017).

36 Huang, S. et al. Compact broadband acoustic sink with coherently coupled weak resonances. Sci. Bull. 65, 373–379 (2020).

37 Ra’di, Y., Sounas, D. L. & Alu, A. Metagratings: Beyond the limits of graded metasurfaces for wavefront control. Phys. Rev. Lett. 119, 067404 (2017).

38 Asadchy, V. et al. Flat engineered multichannel reflectors. Phys. Rev. X 7 (2017).

39 Hou, Z., Fang, X., Li, Y. & Assouar, B. Highly efficient acoustic metagrating with strongly coupled surface grooves. Phys. Rev. Applied 12, 034021 (2019).

40 Ni, H., Fang, X., Hou, Z., Li, Y. & Assouar, B. High-efficiency anomalous splitter by acoustic meta-
339 grating. *Phys. Rev. B* **100**, 104104 (2019).

340 Ribli, D., Horvath, A., Unger, Z., Pollner, P. & Csabai, I. Detecting and classifying lesions in mammo-
341 grams with deep learning. *Sci. Rep.* **8**, 4165 (2018).

342 Naseer, A. *et al.* Refining parkinsonâ€™s neurological disorder identification through deep transfer learning. *Neural. Comput. & Applic.* **32**, 839–854 (2019).

344 Nassif, A. B., Shahin, I., Attili, I., Azzeh, M. & Shaalan, K. Speech recognition using deep neural networks: A systematic review. *IEEE Access* **7**, 19143 – 19165 (2019).

346 Voulodimos, A., Doulamis, N., Doulamis, A. & Protopapadakis, E. Deep learning for computer vision: A brief review. *Comput. Intell. Neurosci.* **2018**, 7068349 (2018).

348 Young, T., Hazarika, D., Poria, S. & Cambria, E. Recent trends in deep learning based natural language processing. *IEEE Comput. Intell. Mag.* **13**, 55–75 (2018).

350 Wu, S. *et al.* Deep learning in clinical natural language processing: a methodical review. *J. Am. Med. Inf. Assoc.* **27**, 457–470 (2019).

352 Liu, D., Tan, Y., Khoram, E. & Yu, Z. Training deep neural networks for the inverse design of nanopho-
353 tonic structures. *ACS Photonics* **5**, 1365–1369 (2018).

354 Sajedian, I., Badloe, T. & Rho, J. Optimisation of colour generation from dielectric nanostructures using reinforcement learning. *Opt. Express* **27**, 5874–5883 (2019).

356 Asano, T. & Noda, S. Optimization of photonic crystal nanocavities based on deep learning. *Opt. Express* **26**, 32704–32717 (2018).

358 Long, Y., Ren, J., Li, Y. & Chen, H. Inverse design of photonic topological state via machine learning. *Appl. Phys. Lett.* **114**, 181105 (2019).

360 Tahersima, M. H. *et al.* Deep neural network inverse design of integrated photonic power splitters. *Sci. Rep.* **9**, 1368 (2019).

362 Pilozzi, L., Farrelly, F. A., Marcucci, G. & Conti, C. Machine learning inverse problem for topological photonics. *Commun. Phys.* **1**, 57 (2018).

364 John, P. *et al.* Nanophotonic particle simulation and inverse design using artificial neural networks. *Sci. Adv.* **4**, eaar4206– (2018).

366 Liu, Q., Hang, R., Song, H. & Li, Z. Learning multiscale deep features for high-resolution satellite image scene classification. *IEEE Trans. Geosci. Remote Sensing* **56**, 117–126 (2018).

368 Orazbayev, B. & Fleury, R. Far-field subwavelength acoustic imaging by deep learning. *Phys. Rev. X* **10**, 031029 (2020).
Xu, P., Arcondoulis, E. J. G. & Liu, Y. Acoustic source imaging using densely connected convolutional networks. Mech. Syst. Signal Proc. 151, 107370 (2021).

Kingma, D. & Ba, J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014).

Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. Nature 323, 533–536 (1986).

Acknowledgements

This work is supported by the National Key R & D Program of China under Grant Nos. 2020YFA0211400 and 2020YFA0211402, the National Natural Science Foundation of China under Grant No. 12074286, and the Shanghai Science and Technology Committee under Grant No. 20ZR1460900.

Author contributions

Y.L. conceived the idea. H.D. implemented model construction and training. X.F., and N.W. provides support for basic physical model. H.D., B.J., and N.W. performed the experiments. Y.L. and Q.C. supervised the whole project. H.D., X.F. and Y.L. wrote the paper with contribution from all authors.

Competing interests

The authors declare no competing interests.