PHYSICAL REVIEW APPLIED **17,** 044007 (2022)

Controlling the Spatiotemporal Response of Transient Reverberating Sound

Qiyuan Wang,[1] Philipp del Hougne,[2] and Guancong Ma[1,*]

[1] *Department of Physics, Hong Kong Baptist University, Kowloon Tong, Hong Kong, China*
[2] *Univ. Rennes, CNRS, IETR - UMR 6164, F-35000 Rennes, France*

(Received 21 December 2021; revised 29 January 2022; accepted 8 March 2022; published 5 April 2022)

Sound propagating inside a room is multiply scattered by the boundaries and other obstacles, forming a complex reverberating sound field (RSF). Such RSFs are not only spatially disordered but also temporally scrambled. Although the spatial control of steady-state RSFs at single frequencies is successfully achieved by extending the idea of wavefront shaping to acoustics, time-coherent polychromatic control of RSFs has remained out of reach. Here, we report spatiotemporal acoustic wavefield shaping by adaptively reshaping acoustic impulse responses at arbitrary positions and different instants in a reverberating room using spatial sound modulators—a reconfigurable and programmable membrane-type acoustic metasurface capable of encoding desirable phase factors over a finite bandwidth. We further analyze the mechanism and performance of spatiotemporal wavefield shaping with a simple numerical model derived from multiple scattering in the time domain. Our results expand the capabilities of RSF control to the temporal domain and demonstrate the possibilities for transient engineering of polychromatic sound in acoustic communication, imaging, and holography.

DOI: 10.1103/PhysRevApplied.17.044007

I. INTRODUCTION

The propagation of waves in multiple-scattering media is a topic of both fundamental interest and practical importance. On one hand, multiple scattering is the essential ingredient for a wide range of exotic wave phenomena, such as coherent backscattering and Anderson localization [1]. On the other hand, diverse wave phenomena (light, microwaves, sound, elastic waves, etc.) on which we rely in our daily lives for communication, sensing, imaging, energy delivery, etc. are also confronted with wave propagation in multiple-scattering media, such as biological tissue, fog, or indoor environments. Ingenious wave-control protocols enable wave engineers not only to counteract the complete scrambling of wave propagation in multiple-scattering media but even to harness its rich degrees of freedom [2,3]. Traditionally, such wave control relies on the coherent control of wave sources with spatial and/or temporal degrees of freedom, such as microwave antenna arrays, acoustic transducer arrays, or optical spatial light modulators. Prominent examples thereof are (i) acoustic time reversal, which controls the sources coherently in time (and additionally often in space) [4,5]; and (ii) optical wavefront shaping, where a (usually monochromatic) plane wave is sculpted with arrays of programmable reflectors before it enters the complex scattering medium [3,6–8]. Because complex media completely mix spatial and temporal degrees of freedom, temporally controlled sources can be used for spatial focusing at a specific time [9], and spatially controlled sources can be used for temporal focusing at a specific location [10–13]. More recently, an alternative idea emerged: controlling the complex propagation medium itself, instead of controlling the wave sources [14,15]. The required purposeful perturbations of the complex propagation medium's scattering properties can be implemented with arrays of programmable scatterers that do not actively emit waves.

An important class of complex scattering media is the reverberant chaotic cavity, an irregularly shaped enclosure, the dimensions of which are much larger than the wavelength [16]. Endowing reverberant enclosures, such as indoor environments [14,17], with electromagnetic programmability has become a very active research track in the microwave domain with applications in wireless communication [17–19], sensing [20], analog computing [21,22], and fundamental research on scattering anomalies like perfect absorption [23]. Moreover, the purely spatial control (offered by a "space-coding" programmable metasurface [24]) is leveraged to focus reverberating microwaves in space *and time* [25], enabling pulselike impulse responses despite rich scattering for low-power wireless receivers [26,27]. Indeed, control over the spatial boundaries of a transient wavefield also affects the time dependence of the impulse response because it yields partial control of all the secondary sources that interfere with the primary source to constitute the impulse response.

---

[*] phgcma@hkbu.edu.hk

The same wave physics is at play in room acoustics [28,29]: rooms are irregularly shaped and much larger than the wavelength of airborne sound in the audible regime (centimeters to meters). The audio experience of a room strongly depends on reverberation and absorption of acoustic waves therein. Rooms with programmable acoustic scattering properties constitute a new research frontier in room acoustics.

The hardware ingredient of a programmable acoustic room is the programmable acoustic metasurface. A way to implement the latter is via acoustic meta-atoms with tunable scattering properties, such as magnetic-field-controlled elastomers [30], tunable membranes [31–33], geometrically tunable resonators [34–38], or electrolysis-controlled microbubble arrays for ultrasound [39]. The interested reader is referred to Refs. [40–43] for comprehensive reviews on acoustic metamaterials. However, only a few works have, to date, reaped the abundance of ideas on how to program acoustic scattering in the context of room acoustics. In Ref. [44], a spatial sound modulator (SSM) based on a tunable membrane-type acoustic metamaterial operating in transmission mode was used to implement adaptive reverberating-sound-field (RSF) control inside a room. Specifically, the SSM is programmed such that the spatial distribution of the RSF is modified according to a predetermined goal, such as minimizing or maximizing the local acoustic amplitude at a chosen position and a chosen frequency. Because the SSM's working principle is based on resonances with a finite quality factor, the SSM makes simultaneous control of polychromatic sound conceivable. In that spirit, Ref. [45] recently explored the potential for broadband minimization of the sound intensity inside rooms in a numerical study. However, *time-coherent* polychromatic control of RSFs with the SSM has, to date, not been demonstrated, and thereby the merits and functionalities of the SSM for transient engineering of polychromatic sound have been left unexplored. Indeed, RSFs appear to be random, not only in space, but also in time, and the transient behavior of a room is much more complicated than its steady-state counterpart [29]. The acoustic impulse response (AIR) inside a room typically contains a line-of-sight contribution due to waves traveling along the direct path between the loudspeaker and microphone, followed by a long exponentially decaying coda due to multipath waves that reverberate multiple times before arriving at the microphone.

Here, by leveraging the ability of the SSM reported in Ref. [44] to modulate sound over a finite range of frequencies, we experimentally achieve the effective control of our reverberating laboratory room's AIR. We show the successful control of time-coherent transient sound at different instants and positions. We analyze the performance of the SSM using a simple numerical model derived from the properties of a diffuse sound field. The model shows that the control performance directly benefits from multiple scattering. Our results expand the known capabilities of acoustic wavefield shaping to the time-coherent regime and identify unexplored possibilities for transient acoustic engineering.

As a side note, the terminology of "spatiotemporal" used here should not be confused with "spatiotemporally modulated" acoustic metasurfaces [46]. In our work, the programmable metasurface configuration is static in time (space coding) and we change it to alter only the metasurface's functionality. In contrast, spatiotemporally modulated metasurfaces are, to date, studied in free space (where no complicated transient behavior arises by definition), and they are subject to so-called "space-time coding" [47], which changes their properties over time during any given measurement.

II. EXPERIMENTAL ENVIRONMENT

The experiments are conducted in our laboratory, which is an irregularly shaped room, as schematically shown in Fig. 1(a). The asymmetric geometry of the room, together with the furniture and equipment (not shown), give rise to a multitude of multiple-scattering events that scramble the sound field, yielding its specklelike nature. To characterize the acoustic properties of the room, we measure the AIR of the room using a loudspeaker and an omnidirectional microphone. The speaker and microphone are separated by a fixed distance, $d_0 = 4$ m, while their positions in the room are randomly chosen. Figure 1(b) shows the average of 320 AIR envelopes. Here, the zero point for time is chosen to be the emission time, and $t_0 = 0.0475$ s corresponds to the instant of direct-pulse arrival at the microphone's position. From the averaged AIR, we obtain the reverberation time for a 60-dB decay (RT_{60}, Ref. [29]) to be around 0.42 s. Based on RT_{60}, we deduce several important room parameters [29]: the decay constant, $\mu = 17.86$ s^{-1}; the average absorption coefficient, $\alpha = 0.22$; and the Schroeder frequency, $f_S = 122$ Hz. In the Schroeder regime, i.e., when considering frequencies above the Schroeder frequency [29,48,49], many resonances overlap at any given frequency, whereas, below the Schroeder frequency, the spectrum displays discrete and isolated resonances. Incidentally, the Schroeder regime is also known as the Hill regime in electromagnetic compatibility [50] and the Ericson regime in nuclear scattering [51]. Our experiments performed in the frequency range of 400–600 Hz are thus well above the Schroeder frequency of our room, such that we operate well within the regime of strong modal overlap.

III. SPATIAL SOUND MODULATORS

The modulation of reverberating sound is achieved via the SSM, which is the same as the one reported in Ref. [44]. For the sake of completeness, we present a brief description here. A photograph of one SSM panel is

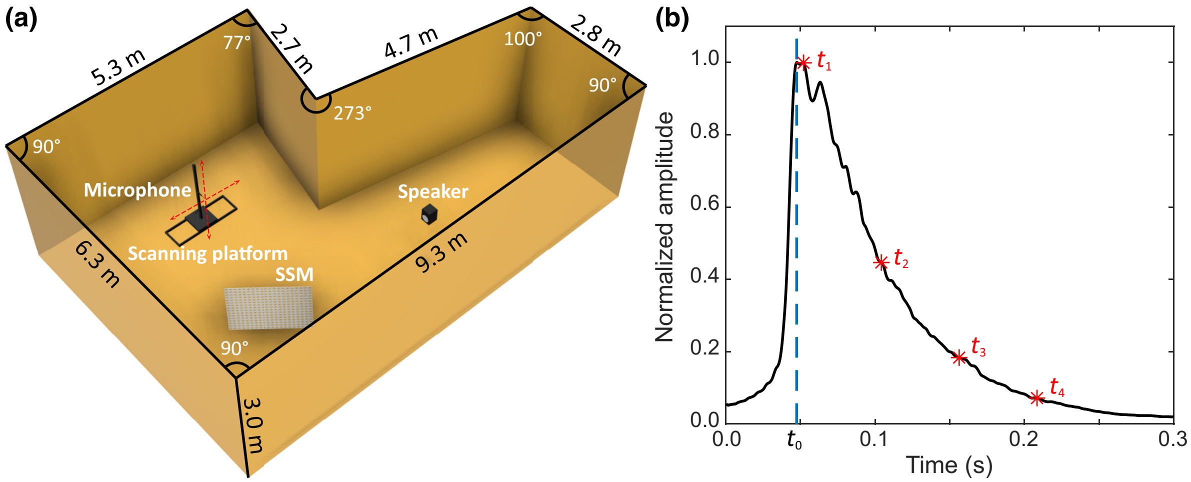

FIG. 1. (a) Schematic drawing of the experimental environment and setup. The SSM, a commercial speaker, and a microphone mounted on a two-dimensional scanning platform are placed in the room. (b) Average AIR envelope of the room. Blue dashed line at $t = t_0 = 0.0475$ s marks the time of arrival of the direct sound, and red asterisks mark the chosen optimization time, t_{opt}.

shown in Fig. 2(a). The SSM is comprised of 324 electrically switchable membrane resonators, each with a radius of 27 mm. Each unit is switchable between two states, denoted "*on*" and "*off*," respectively. The corresponding vibration profiles of the two states are shown in Fig. 2(b). Figures 2(c) and 2(d) show the transmission spectra of a single unit in its two possible states, which are measured using an acoustic impedance tube. It is seen that, in the frequency range of 400–600 Hz [highlighted in orange in Figs. 2(c) and 2(d)], the two states can produce a phase difference of about 120°, as well as an amplitude modulation for transmitted sound. This is consequently the working-frequency range of the SSM. We note that, because of relaxation of the prestressed membrane over time, the working-frequency regime is lower than that reported four years ago for the same device in Ref. [44] (see the Appendix for a more detailed analysis of this effect). Since the membrane resonators have subwavelength dimensions, we pair $2 \times 2 = 4$ units together as one controllable pixel to enhance its capability of controlling far fields. The SSM has 81 pixels in total.

IV. EXPERIMENTAL PROCEDURES

Unlike the results reported in Ref. [44], which demonstrated the spatial control of steady-state RSFs at a single frequency, our goal here is *time-coherent polychromatic* RSF control to reshape the transient response of the environment. Specifically, we aim to increase or suppress the transient AIR amplitude at a chosen time (designated as the optimization time, t_{opt}) and a chosen location in the room by optimally configuring the SSM. We subsequently repeat this experiment for different choices of optimization time and location. We use a pulse-compression technique [52–55] to synthesize the AIR. Specifically, a loudspeaker sends a linear chirp covering 400–600 Hz through a computer audio interface (MOTU 16A) for a duration of 0.4 s. This frequency range is well within the Schroeder regime and coincides with the frequency interval within which the SSM efficiently modulates sound. The acoustic signal is picked up by an omnidirectional microphone at a distance d_0 away from the speaker. The recorded signal is then cross-correlated with the emitted chirp in a MATLAB program, and the result is equivalent to the AIR of a short pulse covering the same frequency regime. The average background AIR of the room is shown in Fig. 1(b).

We choose four different optimization times, $t_1 = 0.052$ s, $t_2 = 0.104$ s, $t_3 = 0.156$ s, and $t_4 = 0.208$ s. They are marked by red asterisks in Fig. 1(b). A single realization of our experiment would not yield results that are representative of our room's average behavior but be very specific to the exact geometry and location of the loudspeaker and microphone. Statistically significant results in complex scattering media are typically obtained by averaging over many realizations. We implement different realizations by randomly changing the placement of the speaker, microphone, and SSM in the room without any preference for the relative positions of the SSM, the speaker, or the microphone. However, as stated above, we need to impose a constraint by fixing the speaker-microphone distance, d_0, to ensure that the duration of the direct line-of-sight path is always the same.

The experiment of a single realization is carried out as follows. First, we determine the optimization objective, which is to increase (maximize) or suppress (minimize) the AIR envelope at a specific location (chosen as the origin, $x = 0$) as well as one of the four chosen optimization times, t_i. The microphone is placed at the

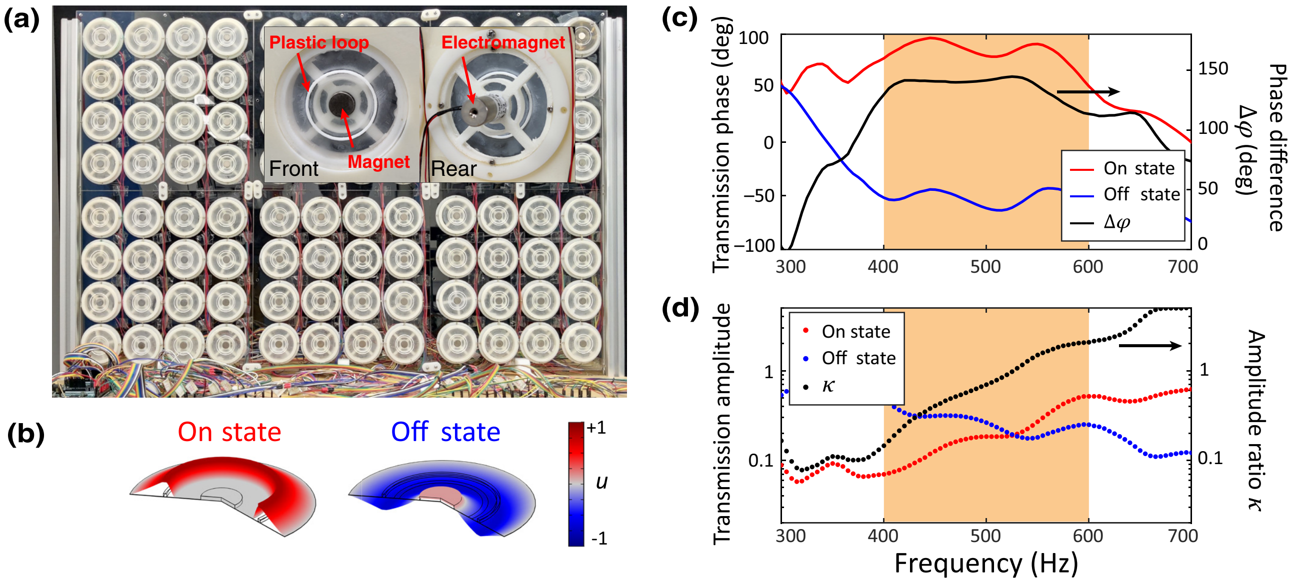

FIG. 2. (a) Photograph of one SSM panel. One pixel contains $2 \times 2 = 4$ membrane resonators. (b) Schematic drawings of the membrane resonator's vibration profiles in *on* and *off* states. Here, u is the out-of-plane displacement. (c) Transmission phases of a membrane resonator in *on* and *off* states, and black curve shows the phase difference, $\Delta\varphi$, between the two states (right axis). (d) Transmission amplitudes of the two states and the amplitude ratio, κ, between the two states (right axis).

chosen location to measure the local AIR. We then obtain the upper envelope of the measured AIR (the upper and lower envelopes are identical in our case), denoted as $\mathcal{E}(0, t_i)$, which is used as the input for a "continuous-pattern" optimization algorithm [56], by which the optimized states of all the pixels are sequentially determined. Specifically, for one pixel after another, we compare the measured AIR envelope at $(0, t_i)$ for both *on* and *off* states and keep the state resulting in an $\mathcal{E}(0, t_i)$ that is closer to our optimization objective. We find that, after this process is performed once for each SSM pixel, the algorithm converges to a local optimum, producing an $\mathcal{E}'(0, t_i)$ as the final outcome. This indicates that the optimal configurations of the SSM pixels are approximately uncorrelated. In other words, the effect of one pixel will not be affected by that of the others. Before and after each realization, one-dimensional linear scans are performed around $x = 0$ to visualize the difference in the spatiotemporal distribution of the AIR.

V. EXPERIMENTAL RESULTS

Figure 3 shows the contrast field, $\eta(x, t)$, of the spatiotemporal patterns averaged over 40 uncorrelated realizations. Here, $\eta(x,t) = \overline{\mathcal{E}'(x,t)/\mathcal{E}^0(x,t)}$, where $\mathcal{E}^0(x,t)$ denotes the preoptimization AIR envelope and the overbar denotes averaging over 40 different realizations. Figures 3(a)–3(d) and 3(e)–3(h) show the results of maximization and minimization, respectively. We observe that both optimization objectives are successfully achieved for the optimization times t_2, t_3, and t_4. In Figs. 3(b)–3(d), bright spots are clearly visible centered at the intersections of the optimization time, t_i, and the optimization location ($x = 0$). Likewise, dark spots are observed in Figs. 3(f)–3(h). All the bright and dark spots are similar in size, irrespective of the chosen optimization time. In the temporal dimension, the extension of the bright and dark spots is around 10 ms. This duration is roughly equivalent to that of a Gaussian pulse containing the same Fourier components as the signal emitted by our loudspeaker. In the spatial dimension, their sizes approach the diffraction limit, $\lambda_0/2$, where $\lambda_0 \cong 0.69$ m corresponds to the wavelength at the center frequency of the emitted pulse. This is also roughly the spatial correlation length corresponding to the average spatial size of speckle grains in disordered wavefields [2,3]. These results indicate that shaping the temporal AIR affects not only the transient response but also the spatial wavefield pattern, resulting in spatiotemporal control [25]. This is also why our purely spatial modulation via the SSM enables spatiotemporal RSF shaping. In comparison, hardly any optimization effect can be identified for t_1 in Figs. 3(a) and 3(e). This is because t_1 is very close to t_0, which is the arrival time of direct sound. In other words, at t_1, hardly any waves can be affected by the SSM linking the speaker to the microphone, and hence, the SSM cannot impact the AIR at t_1. A more detailed analysis will be presented below.

We further examine in Fig. 4 how the contrast at the chosen optimization location, $\eta(0, t_i)$, depends on the chosen optimization time. Here, $\eta(0, t_i)$ can be used as a parameter to characterize the optimization's performance. Both situations (maximization and minimization) can be fit using a numerical model, which is detailed in the next

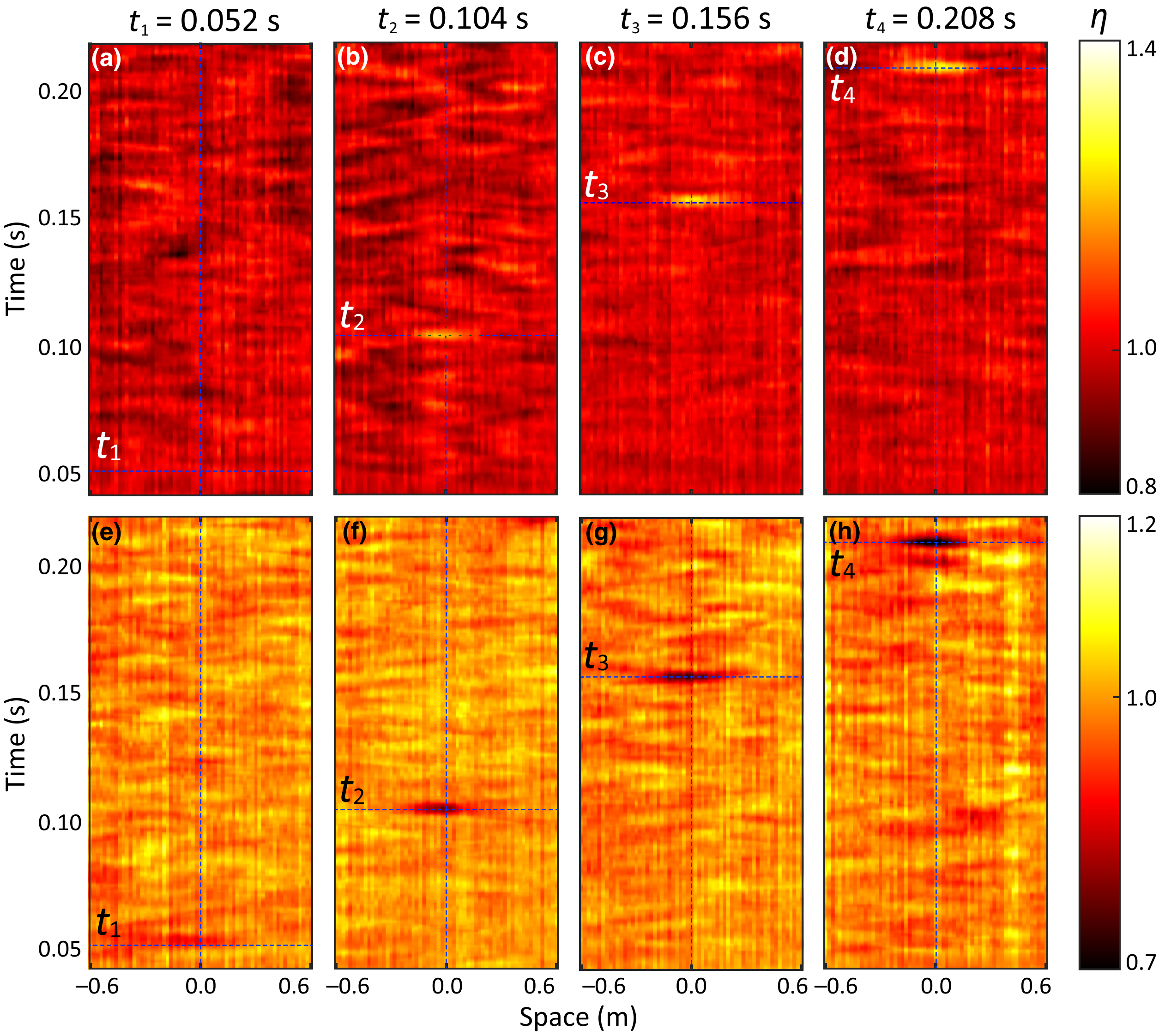

FIG. 3. Spatiotemporal maps of the AIR contrast field, $\eta(x, t)$, at different optimization times for maximization (a)–(d) and minimization (e)–(h). Results in each panel are averaged over 40 uncorrelated experimental realizations.

section. Interestingly, the numerical results suggest that the performances first increase with optimization time and reach optima near 0.18 s. By $t_4 = 0.208$ s, the performance has already deteriorated. This result can be qualitatively understood as follows. The initial performance increase is because the sound encounters more scattering events for a later optimization time, which gives the sound a higher chance of being controlled by the SSM. However, due to the finite quality factor of the room, the sound intensity eventually diminishes, and the background noise gradually takes over. Hence, the performance must inevitably drop.

VI. NUMERICAL MODELING AND ANALYSIS

While rich literature exists on modeling waves in complex scattering media in the frequency domain, extensions to the time domain are challenging and still under development [57]. Moreover, modeling *partially controlled* disorder is an additional open challenge. In this section, we develop a simple numerical model, which we call the transient-scattering model, that captures the time-domain multiple scattering of sound in the SSM-parametrized reverberating room. We remark that our experimental setup, especially the room, is not of a regular configuration, which hinders the development of an analytical model. Our goal here is not to develop a rigorous multiple-scattering theory, but rather to numerically predict the optimization performance.

A schematic of the numerical model is shown in Fig. 5. Our starting point is to consider the physical characteristics of the RSF. Because the experimental frequencies are much higher than the Schroeder frequency, the RSF

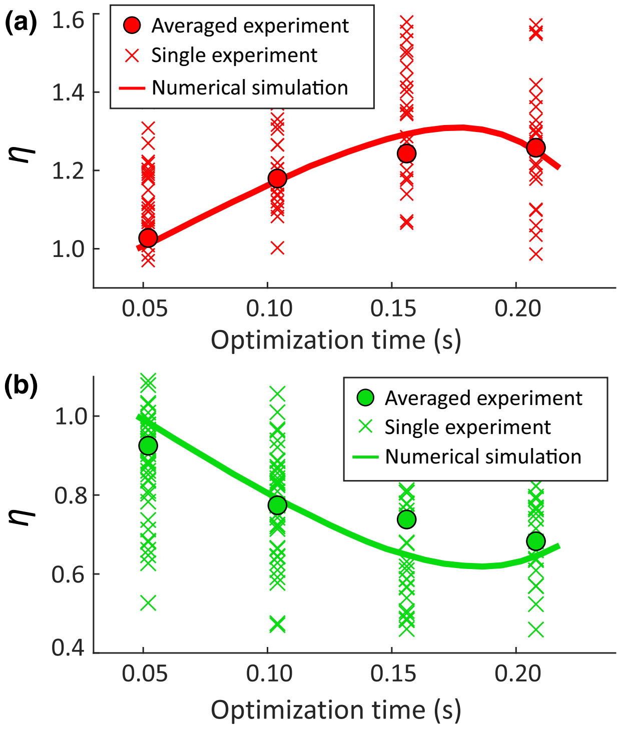

FIG. 4. Optimization performances, $\eta(0, t_i)$, at different optimization times, t_{opt}, for (a) maximization and (b) minimization. Circles indicate the averaged performance, $\eta(0,t_i) = \overline{\mathcal{E}'}(0,t_i)/\overline{\mathcal{E}^0}(0,t_i)$, over 40 experimental realizations, whereas crosses represent the performances of each experimental realization. Continuous curves are obtained from numerical optimizations averaged over 4000 independent configurations for each optimization objective (see Sec. VI).

is approximately a diffuse field [29,58], wherein the propagating sound waves can be treated as acoustic rays. Different from the conventional acoustic rays in the literature of geometrical acoustics, which operate at a single frequency and ignore phases [59], the rays here carry phases and frequencies. This is because our SSM functions mainly as an array of phase modulators within a finite frequency range, and interference is the key physical effect for wave-field shaping. Under these assumptions, we express sound at the optimization location as a linear combination of independent rays, $P_n e^{i\varphi_n}$, where P_n and φ_n are, respectively, the amplitude and phase of the nth ray, and the subscript labels the rays. We build a vector with the rays as elements:

$$R_0 = \begin{pmatrix} P_1 e^{i\varphi_1} \\ P_2 e^{i\varphi_2} \\ \vdots \\ P_N e^{i\varphi_N} \end{pmatrix}, \tag{1}$$

where N is the total number of rays. The local pressure, $\mathcal{P}$, is then

$$\mathcal{P} = \sum_{n=1}^{N} P_n e^{i\varphi_n}. \tag{2}$$

To determine N, we consider that all rays emitted by the loudspeaker must be scattered by the room boundaries (walls and the SSM), which have a total surface area of $A = A_{\text{room}} + A_{\text{SSM}} \cong 190$ m^2. When different rays have distinguishable contributions to the sound field, they can be regarded as independent. Such a condition is satisfied when (i) the rays' (secondary) sources are spatially separated by the correlation length, l_{cor}, or equivalently each ray has a cross section of l_{cor}^2; or (ii) the rays have a frequency difference that exceeds the correlation frequency, f_{cor}. Hence, the number of independent rays within an interval $f_{\text{cor}} \cong 5$ Hz centered on f_i is $N_i = A/[l_{\text{cor}}^2(f_i)]$, where $l_{\text{cor}}(f_i) \cong \lambda_i/2$ is the correlation length. Therefore, the total number, N, of independent rays is obtained by summing N_i for all $f_i \in [400, 600]$ Hz, which is the experimental frequency range. This yields $N = 67171$.

Notably, not all N rays have a contribution to the local field. Hence, many entries in R_0 are zero. The number of nonzero entries is estimated as follows. In a diffuse sound field, the instantaneous positions of all rays are evenly distributed in the room. Only rays with instantaneous distances within the correlation length from the optimization position can contribute to the local field $\mathcal{P}$. In other words, the rays with positions within a sphere centered at the optimization position with radius $dr = \lambda_0/2 \cong 0.35$ m contribute to the local field. (For simplicity, λ_0 is chosen to be the wavelength at the center frequency, 500 Hz.) The number of contributing rays is then $N_r = N(dV/V_{\text{room}}) \cong 95$, where $V_{\text{room}} \cong 121$ m^3 is the volume of our laboratory and $dV = (4/3)\pi(dr)^3 = 0.172$ m^3 is the sphere's volume. Consequently, R_0 is computer generated as a 67171×1 vector with 95 random nonzero entries. These nonzero entries are randomly chosen, and they are assigned with a random complex number, $\xi e^{i\epsilon}$, where $\xi = \text{rand}(0, 1]$, $\epsilon = 2\pi\,\text{rand}[0, 1]$.

Next, we consider the multiple scattering undergone by the rays traveling inside the room. In a diffuse sound field, the average propagation distance between two scattering events is given by the scattering mean free path, l_s [29]. We use a commercial finite-element solver (COMSOL Multiphysics, geometrical acoustics module) to evaluate $l_s = 2.65$ m for our room. The time between two consecutive scattering events can be defined as $\delta t = l_s/c = 0.0077$ s, where $c = 343$ m/s is the speed of sound in air. Within each time interval, δt, the number of rays scattered by the SSM is $N_\chi = \chi N = 648$, where $\chi = A_{\text{SSM}}/A = 0.0096$ is the ratio of the SSM's surface area and the total surface area. In other words, χ can be regarded as the probability of sound being scattered by the SSM. Then, the number

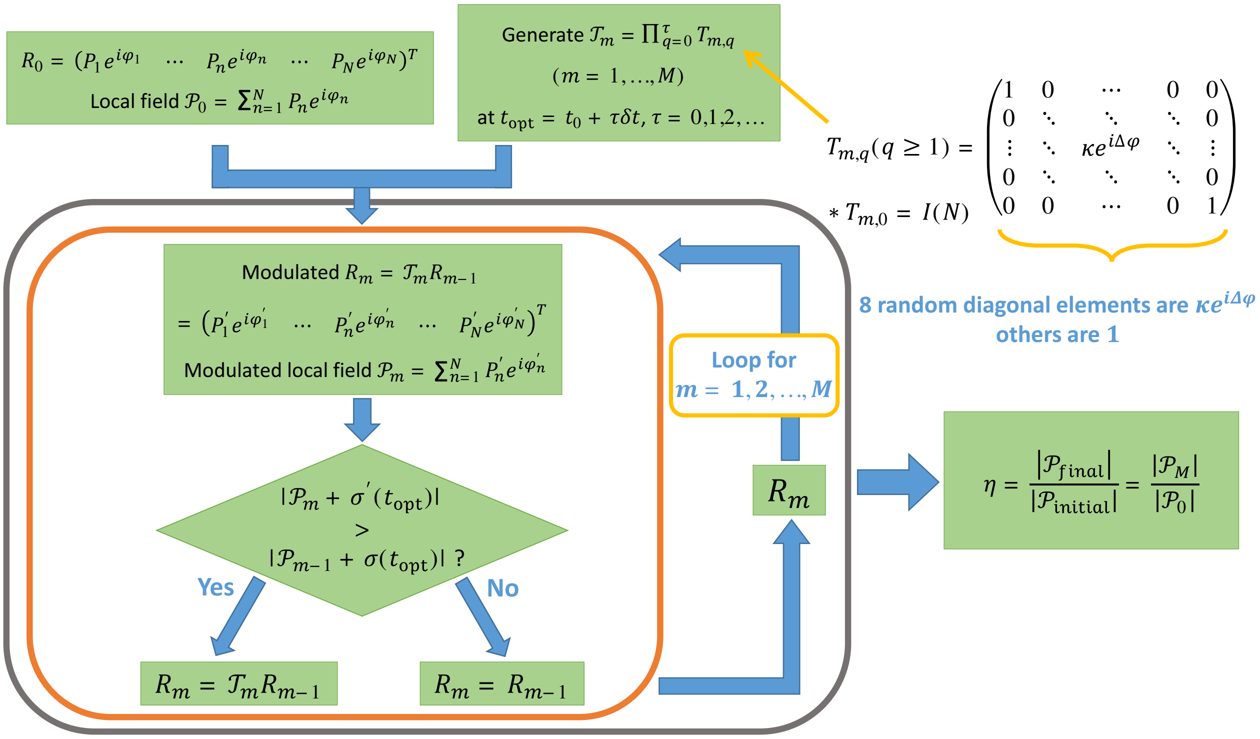

FIG. 5. Algorithmic summary of the transient-scattering model for spatiotemporal maximization of the local field amplitude, $|\mathcal{P}|$.

of scattering events experienced by each ray at time t is estimated to be

$$\tau = \left\lfloor \frac{t - t_0}{\delta t} \right\rfloor, \tag{3}$$

where "$\lfloor \ \rfloor$" denotes taking the integer part, and $t_0 = 0.0475$ s is the arrival time of direct sound. In our experiment, $\tau = 0$, 7, 14, and 21 for optimization time t_1, t_2, t_3, and t_4, respectively.

We are now ready to model the SSM's effect. The SSM can control only the rays that travel through it. In other words, the optimization and its performance are determined by those multiple scattering events that involve the SSM. Then, in each time interval, δt, each pixel can intersect $N_\chi/M \cong 8$ rays ($M = 81$ is the number of pixels) and imprint a modulation, $\kappa e^{i\Delta\varphi}$, on each ray. The amplitude and phase modulation factors, $\kappa = \kappa(f)$ and $\Delta\varphi = \Delta\varphi(f)$, are frequency dependent, and they are directly obtained from the experimental results shown in Figs. 2(c) and 2(d). The scattering of each pixel within one interval, δt, can be captured by an $N \times N$ diagonal matrix, denoted as $T_{m,q}$, where m labels the pixel and q labels time intervals. For $q = 0$, $T_{m,0}$ is an identity matrix because no scattering has occurred. For $q > 0$, eight diagonal entries of $T_{m,q}$ are randomly selected to imprint the modulation $\kappa e^{i\Delta\varphi}$, while the rest are $e^{i0} = 1$. In each subsequent interval, δt, another eight random rays are scattered by each pixel. Therefore, at a chosen optimization time, $t = t_{\rm opt}$, the time-accumulated effect of the mth pixel is

$$R_m = \left(\prod_{q=0}^{\tau} T_{m,q} \right) R_{m-1} = \mathcal{T}_m R_{m-1}. \tag{4}$$

Equation (4) captures the fact that the rays undergo more scattering at later times, which, in turn, gives a better chance of encountering an SSM pixel. Then, a total of 81 different $\mathcal{T}_m$ can be straightforwardly obtained via Eq. (4), one for each pixel. Based on Eq. (4), we can numerically characterize the optimization outcome.

To implement the element-wise iterative optimization for the mth pixel, first, we calculate the local-pressure magnitude using Eq. (2). We then evaluate the modulation effect of the mth pixel by comparing $|\mathcal{P}|$ for R_{m-1} and $R_m = \mathcal{T}_m R_{m-1}$. The final component in our model is to include the effect of background noise, which is modeled as a small perturbation:

$$\sigma(t) = \sigma_0 H(t - t_0) e^{i\epsilon} e^{\mu(t - t_0)}. \tag{5}$$

Here, $\sigma_0 = 0.024$ is a fitting parameter accounting for the experimental noise level, $\epsilon \in [0, 2\pi]$ is a random phase, and H is the Heaviside step function. Due to the exponential decay of the reverberating sound, the effect of background noise is equivalently exponentially

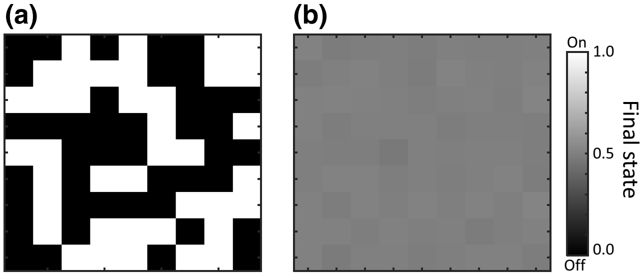

FIG. 6. Final-state patterns of all 81 pixels. (a) State pattern after one single numerical optimization. (b) State pattern averaged over 4000 numerical optimizations.

growing in time, which is represented by the last term, $e^{\mu(t-t_0)}$, with $\mu = 17.86\ \mathrm{s}^{-1}$ being the measured decay constant. The noise term described by Eq. (5) is independently added to $\mathcal{P}$ before evaluating its amplitude at each step, and then their amplitudes are compared. In other words, $|\mathcal{P}_m + \sigma'(t_{\mathrm{opt}})|$ and $|\mathcal{P}_{m-1} + \sigma(t_{\mathrm{opt}})|$ are compared to determine if the modulation yields a more desirable outcome, i.e., closer to the optimization objective. If so, the pixel's state is kept so that $R_m = \mathcal{T}_m R_{m-1}$. Otherwise, the modulation is discarded by reverting the pixel's state, i.e., $R_m = R_{m-1}$.

The optimization procedure is repeated for all $M = 81$ pixels, and we obtain the optimization performance as $\eta = |\mathcal{P}_{\mathrm{final}}|/|\mathcal{P}_{\mathrm{initial}}|$. We perform 4000 independent realizations of this numerical optimization; the averaged performances are plotted as continuous curves in Figs. 4(a) and 4(b).

In Figure 4, good agreement with our experimental results is seen. For example, the numerical results of maximization show that the averaged η first increases with t_{opt} and then drops, indicating that the optimization performance is indeed a trade-off between accumulated scattering events and sound dissipation. We remark that the optimization performance here is equivalent to the experimental performance, defined as the ratio of the AIR envelopes, because both of them are computed using the local-pressure amplitude. Interestingly, although our model is based on diffuse sound, its prediction for t_1 also matches our experimental results well, despite the fact that t_1 lies in the early reflection regime [29]. This is because, at t_1, the number of scattering events $\tau \cong 0$, so our model can still capture the fact that the SSM has almost no control over $|\mathcal{P}|$.

The final-state patterns of the pixels are plotted in Fig. 6. Figure 6(a) shows the final state after one numerical optimization, wherein the states are in a scrambled pattern. Figure 6(b) plots the averaged final state. The averaged result of multiple optimizations converges to the median (0.5), indicating no preference of the pixel's states. Both results show the disordered nature of the complex sound field.

VII. CONCLUSION

We experimentally demonstrate the effective control of transient acoustic responses of an RSF inside a room, the acoustic scattering properties of which can be partially programmed with the SSM. Our results show that the capabilities of acoustic wavefield shaping go beyond the spatial control of steady-state RSFs reported in Ref. [44] and extend to time-coherent polychromatic manipulations. The optimization performance is numerically reproducible by a transient-scattering model, which is based on the consideration of multiple scattering and diffuse sound fields. The model can serve as a guide for real-life applications. We expect the presented spatiotemporal acoustic wavefield shaping to lay the foundation for more advanced adaptive control of complex sound fields. It may also benefit an abundance of applications in acoustic engineering, such as noise suppression, real-time soundscaping, and even temporal acoustic illusions. Recent developments on "spatial ultrasound modulators" [39] enable the achievement of functionalities similar to those presented here for audible sound also for ultrasonic imaging and therapies.

ACKNOWLEDGMENTS

Q.W. and G.M. thank Zhao-Qing Zhang for helpful discussions and Zhen Dong for assisting with the experiments. This work is supported by the National Natural Science Foundation of China (Grants No. 11922416), Hong Kong Research Grants Council (Grants No. 22302718, No. 12302420), ANR/RGC A-HKUST601/18, and Hong Kong Baptist University (Grant No. RC-SGT2/18-19/SCI/006).

APPENDIX: UNIT-CELL DESIGN OF THE SSM

The SSM panels are the same as those reported in Ref. [44], and a photograph is shown in Fig. 2(a). Briefly summarized, the SSM is comprised of 324 units of membrane resonators, which are grouped into 81 pixels. For each membrane resonator, a polyurethane membrane (radius of 27 mm, thickness of 0.1 mm) is uniformly stretched and then glued to a plastic frame. A neodymium magnet disk (radius of 6 mm, mass of 0.9 g) and a plastic loop [inner (outer) radius of 16 mm (17 mm), mass of 0.15 g] are both attached to the membrane. By applying a dc voltage of 10 V, the suspended electromagnet can snap the magnet, so the inner part of the membrane is pinned, which effectively turns the membrane into an annular shape. We denote this as the *on* state. The original state is the *off* state. Their mode shapes are shown in Fig. 2(b).

The working-frequency regime of the membrane resonator is lower than the reported one in Ref. [44]. This is due to relaxation of the prestressed membrane over time and aging of the membrane. From numerical simulations using COMSOL Multiphysics, we identify that the prestress decreases from 0.53 to 0.30 MPa. As a result, the resonant frequency of the *off* (*on*) state drops from around 450 (855) to 349 Hz (706 Hz). However, the transmission properties, including the phase difference between the two states, remain largely unchanged, as plotted in Fig. 2(c).

Lastly, we remark that there is some loss within the membrane resonators, which dissipate acoustic energy, thereby weakening the modulation effectiveness. However, the loss also lowers the sharp resonant peaks in the membrane's transmission spectra, thereby lessening the contrast in the transmission amplitudes between *on* and *off* states. It also leads to a flatter profile of the phase difference. These aspects are beneficial for the polychromatic wavefield shaping demonstrated here. On the other hand, the membrane's loss inevitably reduces the room's Q factor. Nevertheless, because the SSM panels cover only a very small portion of the room's total boundary (<1%), the membrane's loss does not have a significant impact on the room's Q factor.

[1] P. Sheng, *Introduction to Wave Scattering, Localization and Mesoscopic Phenomena*, 2nd ed. (Springer, New York, 2006).

[2] A. P. Mosk, A. Lagendijk, G. Lerosey, and M. Fink, Controlling waves in space and time for imaging and focusing in complex media, Nat. Photonics **6**, 283 (2012).

[3] S. Rotter and S. Gigan, Light fields in complex media: Mesoscopic scattering meets wave control, Rev. Mod. Phys. **89**, 57 (2017).

[4] M. Fink, Time reversed acoustics, Phys. Today **50**, 34 (1997).

[5] C. Draeger and M. Fink, One-channel Time Reversal of Elastic Waves in a Chaotic 2D-Silicon Cavity, Phys. Rev. Lett. **79**, 407 (1997).

[6] I. M. Vellekoop and A. P. Mosk, Focusing coherent light through opaque strongly scattering media, Opt. Lett. **32**, 2309 (2007).

[7] S. M. Popoff, G. Lerosey, R. Carminati, M. Fink, A. C. Boccara, and S. Gigan, Measuring the Transmission Matrix in Optics: An Approach to the Study and Control of Light Propagation in Disordered Media, Phys. Rev. Lett. **104**, 10 (2010).

[8] S. Yoon, M. Kim, M. Jang, Y. Choi, W. Choi, S. Kang, and W. Choi, Deep optical imaging within complex scattering media, Nat. Rev. Phys. **2**, 141 (2020).

[9] F. Lemoult, G. Lerosey, J. de Rosny, and M. Fink, Manipulating Spatiotemporal Degrees of Freedom of Waves in Random Media, Phys. Rev. Lett. **103**, 17 (2009).

[10] J. Aulbach, B. Gjonaj, P. M. Johnson, A. P. Mosk, and A. Lagendijk, Control of Light Transmission Through Opaque Scattering Media in Space and Time, Phys. Rev. Lett. **106**, 103901 (2011).

[11] O. Katz, E. Small, Y. Bromberg, and Y. Silberberg, Focusing and compression of ultrashort pulses through scattering media, Nat. Photonics **5**, 372 (2011).

[12] D. J. McCabe, A. Tajalli, D. R. Austin, P. Bondareff, I. A. Walmsley, S. Gigan, and B. Chatel, Spatio-Temporal focusing of an ultrafast pulse through a multiply scattering medium, Nat. Commun. **2**, 447 (2011).

[13] M. Mounaix, D. Andreoli, H. Defienne, G. Volpe, O. Katz, S. Grésillon, and S. Gigan, Spatiotemporal Coherent Control of Light Through a Multiple Scattering Medium with the Multispectral Transmission Matrix, Phys. Rev. Lett. **116**, 253901 (2016).

[14] N. Kaina, M. Dupré, G. Lerosey, and M. Fink, Shaping complex microwave fields in reverberating media with binary tunable metasurfaces, Sci. Rep. **4**, 6693 (2014).

[15] M. Dupré, P. del Hougne, M. Fink, F. Lemoult, and G. Lerosey, Wave-field Shaping in Cavities: Waves Trapped in a box with Controllable Boundaries, Phys. Rev. Lett. **115**, 017701 (2015).

[16] H. J. Stöckmann, *Quantum Chaos: An Introduction* (Cambridge University Press, Cambridge, 1999).

[17] P. del Hougne, M. Fink, and G. Lerosey, Optimally diverse communication channels in disordered environments with tuned randomness, Nat. Electron. **2**, 36 (2019).

[18] M. D. Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin, J. de Rosny, A. Bounceur, G. Lerosey, and M. Fink, Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come, EURASIP J. Wirel. Commun. Netw. **2019**, 129 (2019).

[19] G. C. Alexandropoulos, N. Shlezinger, and P. del Hougne, Reconfigurable intelligent surfaces for rich scattering wireless communications: Recent experiments, challenges, and opportunities, IEEE Commun. Mag. **59**, 28 (2021).

[20] M. del Hougne, S. Gigan, and P. del Hougne, Deeply Subwavelength Localization with Reverberation-Coded Aperture, Phys. Rev. Lett. **127**, 043903 (2021).

[21] P. del Hougne and G. Lerosey, Leveraging Chaos for Wave-Based Analog Computation: Demonstration with Indoor Wireless Communication Signals, Phys. Rev. X **8**, 041037 (2018).

[22] J. Sol, D. R. Smith, and P. del Hougne, Meta-programmable analog differentiator, Nat. Commun. arXiv:2108.06178 (2021).

[23] M. F. Imani, D. R. Smith, and P. del Hougne, Perfect absorption in a disordered medium with programmable meta-atom inclusions, Adv. Funct. Mater. **30**, 2005310 (2020).

[24] T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, Coding metamaterials, digital metamaterials and programmable metamaterials, Light Sci. Appl. **3**, e218 (2014).

[25] P. del Hougne, F. Lemoult, M. Fink, and G. Lerosey, Spatiotemporal Wave Front Shaping in a Microwave Cavity, Phys. Rev. Lett. **117**, 134302 (2016).

[26] M. F. Imani, S. Abadal, and P. del Hougne, Metasurface-programmable wireless network-on-chip, arXiv:2109.03284 (2021).

[27] E. Arslan, I. Yildirim, F. Kilinc, and E. Basar, Over-the-Air Equalization with Reconfigurable Intelligent Surfaces, arXiv:2106.07996 (2021).

[28] F. Mortessagne, O. Legrand, and D. Sornette, Transient chaos in room acoustics, Chaos **3**, 529 (1993).

[29] H. Kuttruff, *Room Acoustics*, 6th ed. (CRC Press, New York, 2017).

[30] X. Chen, X. Xu, S. Ai, H. Chen, Y. Pei, and X. Zhou, Active acoustic metamaterials with tunable effective mass density by gradient magnetic fields, Appl. Phys. Lett. **105**, 071913 (2014).

[31] S. Xiao, G. Ma, Y. Li, Z. Yang, and P. Sheng, Active control of membrane-type acoustic metamaterial by electric field, Appl. Phys. Lett. **106**, 091904 (2015).

[32] Z. Chen, C. Xue, L. Fan, S. Zhang, X. Li, H. Zhang, and J. Ding, A tunable acoustic metamaterial with double-negativity driven by electromagnets, Sci. Rep. **6**, 30254 (2016).

[33] W. Ao, J. Ding, L. Fan, and S. Zhang, A robust actively-tunable perfect sound absorber, Appl. Phys. Lett. **115**, 193506 (2019).

[34] Z. Tian, C. Shen, J. Li, E. Reit, Y. Gu, H. Fu, S. A. Cummer, and T. J. Huang, Programmable acoustic metasurfaces, Adv. Funct. Mater. **29**, 1808489 (2019).

[35] W. K. Cao, C. Zhang, L. T. Wu, K. Q. Guo, J. C. Ke, T. J. Cui, and Q. Cheng, Tunable Acoustic Metasurface for Three-Dimensional Wave Manipulations, Phys. Rev. Appl. **15**, 024026 (2021).

[36] C. Zhang, W. K. Cao, L. T. Wu, J. C. Ke, Y. Jing, T. J. Cui, and Q. Cheng, A reconfigurable active acoustic metalens, Appl. Phys. Lett. **118**, 133502 (2021).

[37] S.-D. Zhao, A.-L. Chen, Y.-S. Wang, and C. Zhang, Continuously Tunable Acoustic Metasurface for Transmitted Wavefront Modulation, Phys. Rev. Appl. **10**, 054066 (2018).

[38] S.-W. Fan, S.-D. Zhao, A.-L. Chen, Y.-F. Wang, B. Assouar, and Y.-S. Wang, Tunable Broadband Reflective Acoustic Metasurface, Phys. Rev. Appl. **11**, 044038 (2019).

[39] Z. Ma, K. Melde, A. G. Athanassiadis, M. Schau, H. Richter, T. Qiu, and P. Fischer, Spatial ultrasound modulation by digitally controlling microbubble arrays, Nat. Commun. **11**, 4537 (2020).

[40] G. Ma and P. Sheng, Acoustic metamaterials: From local resonances to broad horizons, Sci. Adv. **2**, e1501595 (2016).

[41] S. A. Cummer, J. Christensen, and A. Alù, Controlling sound with acoustic metamaterials, Nat. Rev. Mater. **1**, 16001 (2016).

[42] B. Assouar, B. Liang, Y. Wu, Y. Li, J.-C. Cheng, and Y. Jing, Acoustic metasurfaces, Nat. Rev. Mater. **3**, 460 (2018).

[43] F. Zangeneh-Nejad and R. Fleury, Active times for acoustic metamaterials, Rev. Phys. **4**, 100031 (2019).

[44] G. Ma, X. Fan, P. Sheng, and M. Fink, Shaping reverberating sound fields with an actively tunable metasurface, Proc. Natl. Acad. Sci. **115**, 6638 (2018).

[45] S. Qu and P. Sheng, Minimizing Indoor Sound Energy with Tunable Metamaterial Surfaces, Phys. Rev. Appl. **14**, 034060 (2020).

[46] M. H. Fakheri, H. Rajabalipanah, and A. Abdolali, Spatiotemporal Binary Acoustic Metasurfaces, Phys. Rev. Appl. **16**, 024062 (2021).

[47] L. Zhang, X. Q. Chen, S. Liu, Q. Zhang, J. Zhao, J. Y. Dai, G. D. Bai, X. Wan, Q. Cheng, G. Castaldi, V. Galdi, and T. J. Cui, Space-Time-Coding digital metasurfaces, Nat. Commun. **9**, 4334 (2018).

[48] M. R. Schroeder, Frequency-correlation functions of frequency responses in rooms, J. Acoust. Soc. Am. **34**, 1819 (1962).

[49] M. R. Schroeder and K. H. Kuttruff, On frequency response curves in rooms. comparison of experimental, theoretical, and monte carlo results for the average frequency spacing between maxima, J. Acoust. Soc. Am. **34**, 76 (1962).

[50] D. A. Hill, *Electromagnetic Fields in Cavities: Deterministic and Statistical Theories*, 35th ed. (John Wiley & Sons, Hoboken, 2009).

[51] T. Ericson, Fluctuations of Nuclear Cross Sections in the “Continuum” Region, Phys. Rev. Lett. **5**, 430 (1960).

[52] D. Strickland and G. Mourou, Compression of amplified chirped optical pulses, Opt. Commun. **55**, 447 (1985).

[53] P. Maine, D. Strickland, P. Bado, M. Pessot, and G. Mourou, Generation of ultrahigh peak power pulses by chirped pulse amplification, IEEE J. Quantum Electron. **24**, 398 (1988).

[54] A. Derode, A. Tourin, and M. Fink, Ultrasonic pulse compression with one-bit time reversal through multiple scattering, J. Appl. Phys. **85**, 6343 (1999).

[55] X. Li, P. Li, M.-H. Lu, M. Fink, and G. Ma, Negative Transient Flux in the Near Field of a Subwavelength Source, Phys. Rev. Appl. **16**, L011004 (2021).

[56] I. M. Vellekoop and A. P. Mosk, Phase control algorithms for focusing light through turbid media, Opt. Commun. **281**, 3071 (2008).

[57] S. Ma, T. Antonsen, E. Ott, and S. M. Anlage, in *Proc. PIERS* (2021), p. 2.

[58] F. Jacobsen and P. M. Juhl, *Fundamentals of General Linear Acoustics* (Wiley, West Sussex, 2013).

[59] M. Vorländer, *Auralization* (Springer-Verlag Berlin Heidelberg, Berlin/Heidelberg, 2008).