Solution to the cocktail party problem: A time-reversal active metasurface for multipoint focusing

Constant Bourdeloux, Mathias Fink, Fabrice Lemoult

Editors' Suggestion

Solution to the cocktail party problem: A time-reversal active metasurface for multipoint focusing

Constant Bourdeloux, Mathias Fink, and Fabrice Lemoult*
Institut Langevin, ESPCI Paris, Université PSL, CNRS, 1 rue Jussieu, Paris 75005, France

(Received 30 November 2023; revised 26 January 2024; accepted 15 March 2024; published 21 May 2024)

The cocktail party effect is the capability to focus one's auditory attention on particular audio sources while ignoring other audio sources. We propose an experimental strategy to reproduce this ability by designing a time-dependent metasurface composed of independent active mirrors. Each active mirror is a programmable acoustic unit cell capable of hearing, computing, and re-emitting acoustic signals: each of them acts as a convolution filter. The proper configuration of the metasurface temporal filters allows one to establish an acoustic communication link between groups of individuals immersed in the noisy environment: a multiple-user multiple-input, multiple-output acoustic system is built.

DOI: 10.1103/PhysRevApplied.21.054039

Since the advent of wireless communication, a myriad of techniques has emerged to satiate the ever-growing demand for data, all rooted in wave propagation. Firstly developed in free space [1,2], the beam-forming strategies were extended to disordered media [3]. As the medium becomes more and more complex, free-space propagation cannot be considered anymore and smarter strategies are required. Thanks to their relatively ease of use and their outstanding ability to focus waves in complex environments, time-reversal methods have been extensively studied since the 1990s, first in acoustics [4,5] and then in electromagnetics [6,7]. Slightly later, this degree of control of the wave field was achieved in the optical regime thanks to spatial light modulators [8,9]. Recently, reconfigurable intelligent surfaces (RISs) [10–18] have been proposed to enable wave control in multiple scattering environments in the microwave regime. This opens a new era of improvements for telecommunications, and especially in the context of multiple-user multiple-input, multiple-output (MU-MIMO) telecommunication schemes. Recent work [19] generalized this concept to the acoustic regime too. Nevertheless, all these RISs have been studied in a relatively narrowband regime. We propose here to extend the results to broadband acoustics for speech and the cocktail party effect, which are in essence ultra-wide-band communication scenarios.

In this study, we develop the concept of a broadband RIS by studying a reconfigurable time-dependent acoustic metasurface. With such a tool, we are able to perform real-time temporal convolution for a multitude of independent pixels with reconfigurable reflection laws. Such an acoustic RIS is capable of shaping the broadband acoustic wave field in a custom and efficient way. We show that the RIS representing a fraction of a room boundary enables one to control all the wave field contained by the said room in a broadband regime. By performing experiments in a multiple-scattering medium, we demonstrate that MU-MIMO scenarios over a wide spectrum can be created. The latter benefit from the spatiotemporal focusing qualities of time reversal.

Before going to the idea at the core of this paper, let us come back to a very simple acoustic design. Indeed, the ideal environment to maximize the amplitude received on receiver R for a given emitted energy from emitter E would be a perfectly reflecting ellipsoidal room with E and R being at the two foci [Fig. 1(a)]. Such an idea emerged in the 17th century with the design of religious rooms for confessions by lepers without priests being infected [20,21]. The role of the ellipsoidal room is to put in phase all the rays originating from point E reflected by the room boundary at position M and received at point R. Mathematically, all possible paths have the same length:

$$EM + MR = 2a, \tag{1}$$

where $2a$ corresponds to the long axis of the ellipse.

Nevertheless, two limitations arise: (i) usually we cannot control the room shape; and (ii) even though an ellipsoidal room is a nice solution, it is still very restrictive since the emitter and the receiver have to be at the foci: we can define only one couple E and R.

*Corresponding author: fabrice.lemoult@espci.psl.eu

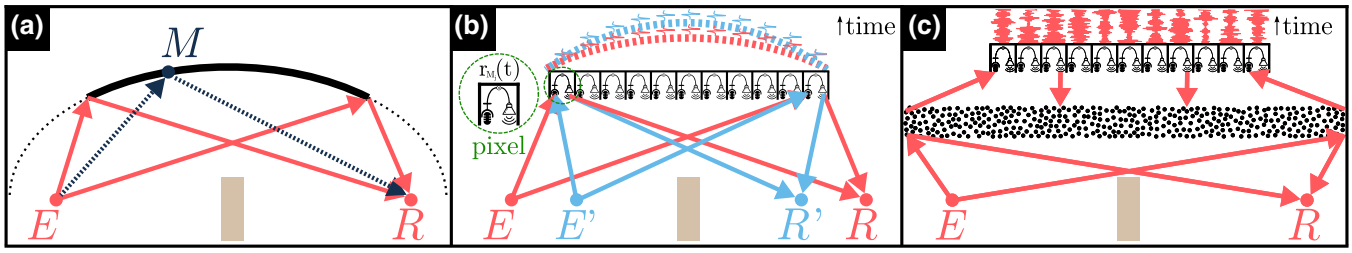

FIG. 1. (a) An ellipsoid is the ideal configuration to focus one emitter toward one receiver (and vice versa). (b) The ellipsoidal mirror is replaced by a transient RIS where each pixel is a programmable acoustic unit capable of hearing, computing, and emitting acoustic signals. By mimicking multiple ellipsoids on the RIS, one can create multiple independent audio channels. (c) Applied in a more-complex environment such as a multiple-scattering medium, the transient RIS should benefit from the additional spatiotemporal degrees of freedom offered by the medium.

To address these two issues, one needs to replace the ellipsoid by a more-versatile solution. This is where the concept of an RIS comes into play [10,12]. The curved reflecting wall of an ellipsoid is replaced by a planar surface [Fig. 1(b)]. At one frequency, to mimic the reflection on the walls of the ellipsoid, a dephasing mirror is required. To satisfy the same path length as in Eq. (1), the reflection coefficient on the pixel M_j of the surface must satisfy the condition

$$r_{M_j}(\omega) = e^{i\omega t_j}, \quad \text{with } t_j = \underbrace{\frac{2a}{c}}_{T} - \underbrace{\frac{EM_j}{c}}_{t_{E\to M_j}} - \underbrace{\frac{M_jR}{c}}_{t_{M_j\to R}}. \tag{2}$$

Overall, the signal received by R when the emitter E emitted a monochromatic wave is the sum of all the secondary emitted waves on the pixels of the surface. Mathematically, if we introduce the monochromatic free-space Green's functions $G^0_{A\to B}(\omega)$ linking a source at A to a receiver at B, it corresponds to

$$s_R(\omega) = \sum_{M_j}\left[G^0_{E\to M_j}(\omega) r_{M_j}(\omega) G^0_{M_j\to R}(\omega)\right]. \tag{3}$$

Note that we have ignored from the beginning the line-of-sight contribution from E to R.

Since we are working in the acoustic regime, where the emissions are extremely wide band, we need to fulfill the phase-matching conditions for all the frequencies. Or, said differently, we need to create a transient response in reflection for the design of our RIS. The reflection coefficient (a complex number) $r_{M_j}(\omega)$ becomes a temporal signal $r_{M_j}(t)$: the phase law becomes a delay law. Also, the multiplications in Eq. (3) are replaced by convolution products:

$$s_R(t) = \sum_{M_j}\left[G^0_{E\to M_j}(t) \underset{t}{\circledast} r_{M_j}(t) \underset{t}{\circledast} G^0_{M_j\to R}(t)\right], \tag{4}$$

with the temporal reflection coefficients being

$$r_{M_j}(t) = \delta(t-T) \underset{t}{\circledast} \delta(t + t_{E\to M_j}) \underset{t}{\circledast} \delta(t + t_{M_j\to R}). \tag{5}$$

As a result, this delay law is supposed to mimic the initial ellipsoid as shown in Fig. 1(b). By simply changing the delay law on each pixel of the RIS, one can change the ellipsoid configuration, thus changing the emitter and the receiver—let us say that we change E to E' and R to R' [Fig. 1(b)]. And, the icing on the cake is that one can also add the two delay laws and create simultaneously two telecommunication channels. By adopting the language of the telecommunication community, we switch from a single-input, single-output (SISO) telecommunication scheme to a MU-MIMO scheme [Fig. 1(b)].

Interestingly, these mirror coefficients in Eq. (5) are proportional to the product of two time-reversed mirrors [5], with each one associated, respectively, with antenna E and antenna R:

$$r_{M_j}(t) \propto G^0_{E\to M_j}(-t) \underset{t}{\circledast} G^0_{M_j\to R}(-t). \tag{6}$$

Previous research [22,23] have proven that a disordered medium improves the performances of time-reversal focusing (both in time and in space). The more degrees of freedom offered by the medium, the more efficient the time-reversal process is [23,24]. We therefore propose incorporating between the emitter and the mirror elements, and similarly between the mirror and the receivers, a multiple-scattering medium [Fig. 1(c)]. This time it becomes impossible to find the analytical solution for the reflection coefficients $r_{M_j}(t)$, but one can find them experimentally. Indeed, this requires one to measure all the Green's functions between the emitters and the mirror elements $G_{E_i\to M_j}(t)$, as well as the Green's functions between the mirror elements and the receivers $G_{M_j\to R_k}(t)$. At the end, the transient reflection coefficients will be given by Eq. (6) where the free-space Green's functions are replaced by the more-complex ones. In this configuration, we not only ignore the line of sight between E and R, but we also

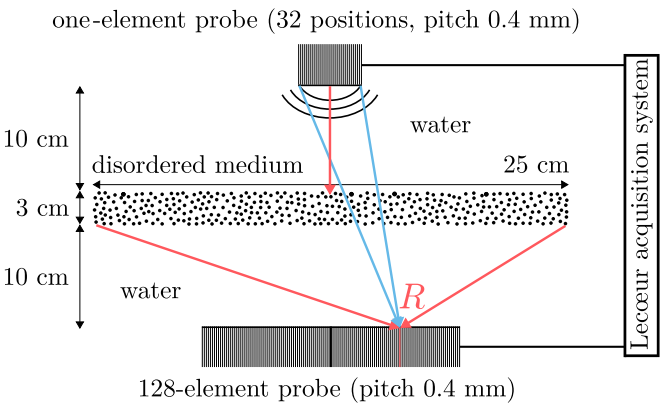

FIG. 2. Experimental setup for the acquisition of the Green's function. The blue arrows depict the RIS's focusing aperture in free space, whereas the red arrows illustrate the RIS's focusing aperture within the disordered medium.

ignore all the paths linking these two points without reaching the RIS. To tackle this issue, we make the assumption that the controllable gain of the RIS allows us to overpass the uncontrollable paths. The RIS can emit sound with an arbitrary amplitude. Thus, the amplitude of the controlled path can be an order of magnitude higher than the amplitude of the uncontrolled path.

To try this concept experimentally, we have designed an underwater ultrasonic experiment. To measure the Green's functions, we use a 128-element ultrasonic probe with a central frequency of 3.5 MHz and a pitch equal to the wavelength, i.e., 0.4 mm, facing a single emitting ultrasonic transducer that corresponds to one element of a linear transducer array (0.4 mm $\times$ 1.2 cm, central frequency of 3.5 MHz). The single-element transducer is moved parallel to the 128-element probe so as to emulate a 32-element probe (see Fig. 2). These two ultrasonic devices are directed one to the other from a distance of 23 cm. They are immersed in water in the middle of a nonreverberating water tank of size 135 $\times$ 90 $\times$ 80 cm^3 (length $\times$ width $\times$ height). The disordered medium is added without modification of any other parameter of the acquisition, only by our putting a forest of steel rods between both devices. The forest of steel rods is composed of randomly distributed, vertical parallel steel rods of diameter 0.8 mm with a density of 18.75 rods per square centimeter. The multiple-scattering medium used in this experiment is the same as the medium reported in Refs. [22,25]. We obtain 128 $\times$ 32 Green's functions by emitting a one-cycle pulse from the single element and receiving the signals for all the 128 elements of the linear array. For the rest of the experiments, we decided to divide the 128-element part in two to obtain 64 emitters and 64 receivers. The 32-element part corresponds to the RIS side. In conclusion, we have a bank of temporal Green's functions from the 64 emitters to the 32 mirrors and from the 32 mirrors to the 64 receivers for each medium (free space and disordered medium).

With the knowledge of all these Green's functions, it is possible to calculate what would be the temporal signal $s_R(t)$ received at any position R by simply evaluating with a computer the convolution product represented in Eqs. (4) and (6). Note that for the case of the multiple-scattering medium, one has to replace the free-space Green's function by the more-complex measured Green's function. For computing purposes, the calculations are performed in the Fourier domain, where the convolution product is replaced by a simple multiplication. When the RIS is OFF, the reflection coefficients are simply 1 at each frequency, or equivalently in time, their impulse response is a Dirac function. In all the results presented herein, the data are normalized by the maximum received signal. As we assume that everything is linear, it is easier and without consequences to compare the normalized results.

The results for the SISO configuration (emitter E emits a pulse that is received by receiver R) are shown in Fig. 3 for the RIS ON and OFF and in free space or a disordered medium. The SISO configuration is equivalent to the aforementioned ellipsoidal room with E and R situated at the foci. Having recorded all the Green's functions, we create the reflection coefficients of the metasurface with a computer by referring to Eq. (6).

We start the analysis by considering the RIS as OFF; that is, the reflection coefficient at each frequency is 1, or equivalently in time, their impulse response is a Dirac function. In this case, the signal received in free space at receiver R corresponds to a pulse [Fig. 3(a)], similar to the one initially emitted at E. Indeed, a spherical wave exits from point E, reflects on a planar mirror, and then propagates towards R. When looking at the full field received around the point R as shown in the color map, one can recognize the wave front associated with this reflection on a planar mirror.

In the disordered medium, things are more complicated [Fig. 3(b)]. During its propagation from E to R by passing through the mimicked planar mirror, the wave generates multiple secondary waves each time a scatterer is encountered. As a result, the signal received by R exhibits a first pulse followed by multiple echoes. The signal resembles a coda [22] that extends over more than 100 times the initial pulse duration.

Looking at the spatiotemporal distribution of the wave field [color map in Fig. 3(b)], one can recognize a spatiotemporal speckle [24]: the ballistic wave front is barely visible and one can see only bright and dark spots both in time and in space as a result of multipath interferences.

Next, the RIS is turned on: each element of the RIS is configured to perform a double time-reversal operation as in Eq. (6). As shown in Figs. 3(c) and 3(d), the RIS has drastically modified the wave field received at and around point R. In free space, the double time-reversal operation

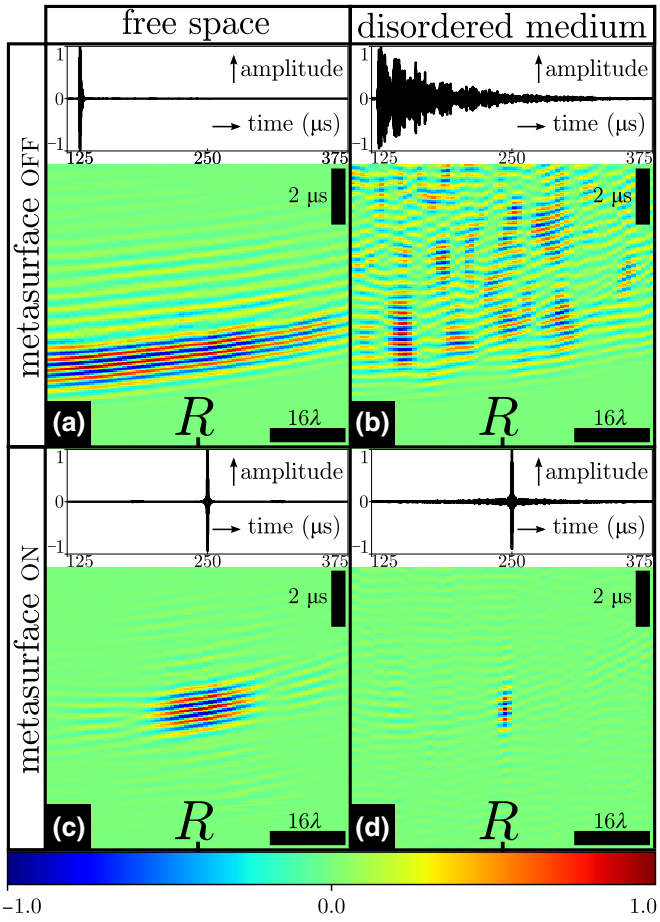

FIG. 3. SISO configuration. When the RIS is OFF, the color-coded spatiotemporal pressure map shows (a) a wave front in free space and (b) a spatiotemporal speckle in the disordered medium . When the RIS is turned on, spatiotemporal focusing is obtained in the two media [(c),(d)]. In the disordered medium, a thinner focal spot is obtained thanks to the increase of the focusing aperture.

corresponds to a delay law as in Fig. 1(b), which mimics an ellipsoidal mirror. As a result, the signal received at point R still resembles a pulse because all the secondary waves arrive simultaneously since the delay law compensate for the differences in paths' length, but it shows a delay compared with the initial case. Spatially, the effect of the metasurface is indeed to focus the wave, and a diffraction-limited spot is visible on the color map. Its width corresponds to $\lambda F/D$, where λ is the wavelength associated with the central frequency of the pulse, in free space $F \approx 500\lambda$ is the distance between the RIS and point R, and $D = 32\lambda$ is the overall dimension of the RIS, so the focal width is approximately $\lambda F/D \approx 16\lambda$ [Fig. 3(c)].

In the disordered medium, things are even more impressive. The double time-reversal operation transforms the long coda signal received at R into a short pulse: all path durations are compensated by the RIS and constructive interferences are built at a specific time at R. Spatially, the wave field is again focused at point R, but the focal width is much smaller, with a width comparable to λ. This focal spot is still limited by the diffraction limit, but we benefit from the increase of the angular aperture, as sketched in Fig. 1(c). This effect was observed earlier in work concerning focusing through a multiple-scattering medium [23,26], but in our experiment this effect is observed in a double time-reversal operation. This reducing of the focal width (around λ) can be interpreted geometrically from Fig. 2, where an increase in the focusing angle is evidenced in the presence of the steel rods (F is reduced and D is increased).

Let us now imagine a more-complex scenario: instead of having just one emitter and one receiver, let us imagine that we want to create four independent channels between four couples of emitter and receiver. In terms of telecommunications, this corresponds to a MU-MIMO scenario [27]; or in terms of audible acoustics, it corresponds to the cocktail party effect where four distinct people want to listen to four different speakers [28,29]. In that configuration, the transient reflection coefficient of the RIS simply corresponds to the summation of the four double time-reversal signals.

As an illustration, we imagine that each emitter emits a series of two pulses separated by different time intervals. This choice is only a visual presentation of the performances, and more-complex signals are required to assess the performance, as we will see later with the emission of white noise. The metasurface has absolutely no prior knowledge of the signal types emitted. In practical scenarios, the metasurface can focus any type of signal sent by emitters, even signals overlapping in the temporal domain. As shown in Fig. 4(a), it is completely impossible to distinguish the different messages in free space. Indeed, the separation between the receivers is smaller than the focal width observed in Fig. 3(c) and all the messages are mixed. Note that, by reciprocity, we can say that the emitters are also within a focal width on the other side and increasing the distance between the receivers could not solve this problem.

However, in the multiple-scattering medium, the separation distance between the receivers becomes greater than the focal width and one recognizes the four different messages [Fig. 4(b)]. Still, the messages received on receivers R_1 and R_4 do not clearly evidence two well-separated pulses. This comes from the fact that time reversal does not optimize the level of spatiotemporal side lobes. There are two problems that degrade the time-reversal focusing and do not permit one to achieve a perfect inverse-filter focusing. First, dissipation breaks the time-reversal invariance of the wave equation. A wave will suffer twice the dissipation effects during its forward and backward propagation, and this irreversibility degrades the time-reversal focusing quality. Second, a closed time-reversal cavity is difficult to build, and in practice the time-reversal operation is achieved on a limited aperture. This angular aperture

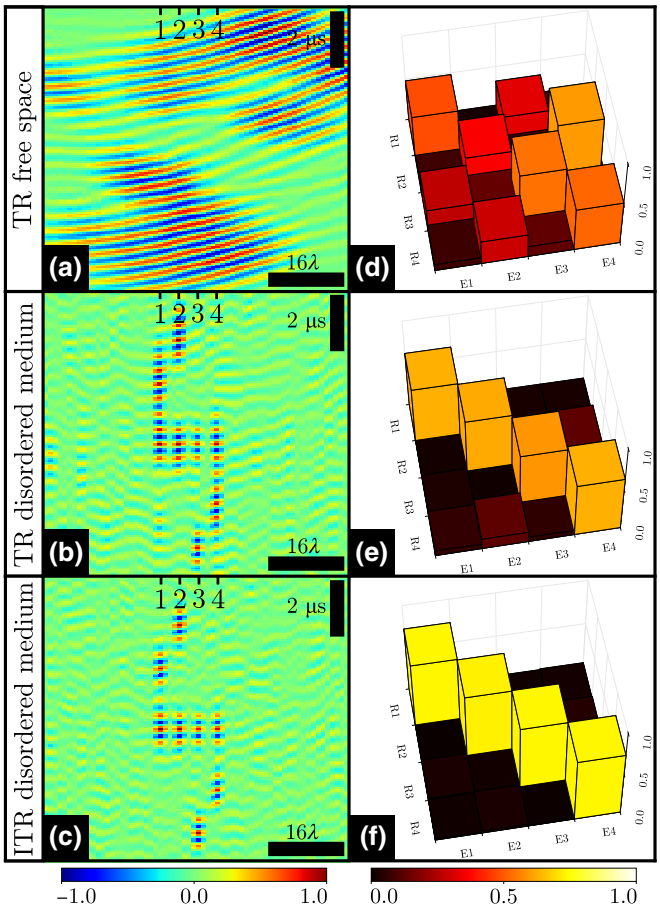

FIG. 4. MU-MIMO configuration: four emitters emit different signals (a succession of two pulses or noises). (a) RIS ON performing time reversal (TR) in free space; the received signals are all mixed. (b) RIS ON performing time reversal in a disordered medium; the received signals on four distinct receivers are distinct, but the signals have been slightly deformed. (c) RIS ON performing ITR in a disordered medium; the received signals are focused both in space and in time. (d)–(f) When the four emitters emit independent white noise, the correlation matrices between the received signals and the emitted signals are evaluated in the same three scenarios. Only in the multiple-scattering medium with ITR are four independent audio channels created.

limitation gives rise to new information losses that may cause the focusing quality to decrease, particularly in strongly diffracting media.

As already proven in previous research [30,31], it is possible to improve the time-reversal operation with an iterative algorithm. We call it the "iterative-time-reversal (ITR) operation." The ITR method iterates the time-reversal process in order to cancel the focusing degradation due to absorption and other information losses. To that end, since we have all the Green's functions at hand, we prepared a set of ITR signals in order to reduce the number of spatiotemporal side lobes on the four receivers. This time, the spatiotemporal wave field around the four receivers better evidences the performances of the RIS for MU-MIMO purposes.

In a real cocktail party configuration, the room is filled with multiple emitters emitting simultaneously audio signals to multiple receivers. The signals emitted are more complex than pulses. In that sense, to mimic such phenomena, we decided to use multiple emitters to emit broadband independent noises.

To quantify the quality of the transmission channels in this telecommunication scenario, we keep the same four transmission channels with the help of the RIS but now the emitted signals $s_{E_i}(t)$ correspond to four different noises that have the same bandwidth as our transducers. The signal received on each receiver $s_{R_k}(t)$ is then compared with each emitted signal, with the normalized correlation coefficient defined as

$$\Gamma_{E_i,R_k} = \frac{\int s_{E_i}(t) s_{R_k}(-t)\mathrm{d}t}{\sqrt{\int s_{E_i}^2(t)\mathrm{d}t}\sqrt{\int s_{R_k}^2(t)\mathrm{d}t}}. \tag{7}$$

When the RIS is OFF, in all cases, the correlation coefficients Γ_{E_i,R_k} have an average value of around 0.02. When the RIS is turned on, the normalized correlation coefficients are as shown in Figs. 4(d)–4(f). In free space, we cannot distinguish a strong correlation relation between the emitters and the receivers: receiver R_1 receives a strong signal from E_1 (target emitter) and E_3 (nontarget emitter), etc. In a disordered medium, there is a strong correlation between emitter signals and receiver signals. The off-diagonal coefficients are low compared with the diagonal ones, meaning that independent audio channels have been created. Even more impressive, in the case of the ITR, the off-diagonal coefficients seem to have vanished to the benefit of the diagonal ones: we proved that the ITR is more efficient than the time reversal in a disordered medium in a MU-MIMO configuration.

In this work we have developed a temporal RIS capable of establishing acoustic communication in a MU-MIMO configuration in a noisy environment thanks to a double time-reversal method. We also enhance such a method with an ITR algorithm, outperforming the results of the time-reversal method.

This proof-of-concept experiment has successfully demonstrated the efficacy of using transient RISs in the ultrasound range. Our laboratory's prior experience with single time-reversal experiments in this medium yielded consistent and promising results, affirming the potential of integrating transient RISs with the concept of time reversal. While there may be direct applications in the ultrasound regime, particularly for underwater telecommunications, our primary objective is the transposition of this technology to airborne audible acoustics. It is true that, in our scenario of ultrasound focusing, we use rigid

scatterers. In a real environment such as a room full of absorbing objects and people, the time-reversal operation will focus sound waves as intended [32–34] but the performance of the focusing will be degraded. To underscore the feasibility of this transposition, we can assess certain order-of-magnitude considerations. The typical diffusion time of the measured coda through the multiple scattering medium used corresponds to a T60 [35] of 200 μs (time at which the amplitude has decreased to −60 dB compared with its maximum value). Operating at 3.5 MHz in the ultrasound range, we can extrapolate to the audible acoustic realm at 1 kHz, resulting in a T60 of 700 ms. This duration implies a relatively high reverberation environment, akin to a reverberating room, yet remains within reasonable bounds [32]. Furthermore, the relative bandwidth used in our experiments is approximately 100%, a parameter that could be significantly expanded in audible acoustics, ranging from 20 Hz to 20 kHz. This could greatly improve the performance of the transient RIS by increasing the number of temporal degrees of freedom [23]. Considering implementation aspects, the necessary hardware for this endeavor is already in place (microphones, speakers and multiple input-outputs devices are everywhere), laying a solid foundation for the next crucial step: conducting experiments specifically tailored to audible acoustics. This progression is essential to unlock the full potential of transient RISs in addressing the realm of cocktail party listening.

ACKNOWLEDGMENTS

This work has received support under the program "Investissements d'Avenir" launched by the French Government and by the Simons Collaboration on Symmetry-Driven Extreme Wave Phenomena.

[1] Constantine A. Balanis, *Antenna Theory: Analysis and Design* (Wiley, Hoboken, New Jersey, 2016), 4th ed.

[2] Hidetsugu Yagi, Beam transmission of ultra short waves, Proc. Inst. Radio Eng. **16**, 715 (1928).

[3] Aris L. Moustakas, Harlod U. Baranger, Leon Balents, Anirvan M. Sengupta, and Steven H. Simon, Communication through a diffusive medium: Coherence and capacity, Science **287**, 287 (2000).

[4] Didier Cassereau and Mathias Fink, Focusing with plane time-reversal mirrors: An efficient alternative to closed cavities, J. Acoust. Soc. Am. **94**, 2373 (1993).

[5] Mathias Fink, Time-reversed acoustics, Sci. Am. **281**, 91 (1999).

[6] Geoffroy Lerosey, Julien de Rosny, Arnaud Tourin, Arnaud Derode, Gabriel Montaldo, and Mathias Fink, Time reversal of electromagnetic waves, Phys. Rev. Lett. **92**, 193904 (2004).

[7] Geoffroy Lerosey, Julien de Rosny, Arnaud Tourin, and Mathias Fink, Focusing beyond the diffraction limit with far-field time reversal, Science **315**, 1120 (2007).

[8] Ivo Micha Vellekoop and Allard P. Mosk, Universal optimal transmission of light through disordered materials, Phys. Rev. Lett. **101**, 120601 (2008).

[9] Sébastien M. Popoff, Geoffroy Lerosey, Rémi Carminati, Mathias Fink, Albert Claude Boccara, and Sylvain Gigan, Measuring the transmission matrix in optics: An approach to the study and control of light propagation in disordered media, Phys. Rev. Lett. **104**, 100601 (2010).

[10] Nadège Kaina, Matthieu Dupré, Geoffroy Lerosey, and Mathias Fink, Shaping complex microwave fields in reverberating media with binary tunable metasurfaces, Sci. Rep. **4**, 6693 (2014).

[11] Philipp Del Hougne, Mathias Fink, and Geoffroy Lerosey, Optimally diverse communication channels in disordered environments with tuned randomness, Nat. Electron. **2**, 36 (2019).

[12] Yuanwei Liu, Xiao Liu, Xidong Mu, Tianwei Hou, Jiaqi Xu, Marco Di Renzo, and Naofal Al-Dhahir, Reconfigurable intelligent surfaces: Principles and opportunities, IEEE Commun. Surv. Tutor. **23**, 1546 (2021).

[13] Marco Di Renzo, Konstantinos Ntontin, Jian Song, Fadil H. Danufane, Xuewen Qian, Fotis Lazarakis, Julien De Rosny, Dinh-Thuy Phan-Huy, Osvaldo Simeone, Rui Zhang, Meroaune Debbah, Geoffroy Lerosey, Mathias Fink, Sergei Tretyakov, and Shlomo Shamai, Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison, IEEE Open J. Commun. Soc. **1**, 798 (2020).

[14] Odysseas Tsilipakos, Anna C. Tasolamprou, Alexandros Pitilakis, Fu Liu, Xuchen Wang, Mohammad Sajjad Mirmoosa, Dimitrios C. Tzarouchis, Sergi Abadal, Hamidreza Taghvaee, and Christos Liaskos, *et al.*, Toward intelligent metasurfaces: the progress from globally tunable metasurfaces to software-defined metasurfaces with an embedded network of controllers, Adv. Opt. Mater. **8**, 2000783 (2020J).

[15] George C. Alexandropoulos, Nir Shlezinger, and Philipp Del Hougne, Reconfigurable intelligent surfaces for rich scattering wireless communications: Recent experiments, challenges, and opportunities, IEEE Commun. Mag. **59**, 28 (2021).

[16] Chongwen Huang, Alessio Zappone, George C. Alexandropoulos, Mérouane Debbah, and Chau Yuen, Reconfigurable intelligent surfaces for energy efficiency in wireless communication, IEEE Trans. Wireless Commun. **18**, 4157 (2019).

[17] Ertugrul Basar, Marco Di Renzo, Julien De Rosny, Merouane Debbah, Mohamed-Slim Alouini, and Rui Zhang, Wireless communications through reconfigurable intelligent surfaces, IEEE Access **7**, 116753 (2019).

[18] Mohamed A. ElMossallamy, Hongliang Zhang, Lingyang Song, Karim G. Seddik, Zhu Han, and Geoffrey Ye Li, Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities, IEEE Trans. Cogn. Commun. Netw. **6**, 990 (2020).

[19] Hongkuan Zhang, Qiyuan Wang, Mathias Fink, and Guancong Ma, Optimizing multi-user indoor sound communications with acoustic reconfigurable metasurfaces, Nat. Commun. **15**, 1270 (2024).

[20] Athanasius Kircher and Francesco Eschinardi, Phonurgia nova sive conjugium mechanico-physicum artis naturae paranympha phonosophia concinnatum (1673), https://gallica.bnf.fr/ark:/12148/bpt6k111869n/?lang=EN.

[21] Athanasius Kircher and Tobias Nißlen, Neue Hall- und Thon-Kunst, oder Mechanische Gehaim-Verbindung der Kunst und Natur, durch Stimme und Hall-Wissenschafft Gestifftet (1684), https://www.digitale-sammlungen.de/en/details/bsb10497374.

[22] Arnaud Derode, Arnaud Tourin, and Mathias Fink, Random multiple scattering of ultrasound. II. Is time reversal a self-averaging process?, Phys. Rev. E **64**, 036606 (2001).

[23] Fabrice Lemoult, Geoffroy Lerosey, Julien de Rosny, and Mathias Fink, Manipulating spatiotemporal degrees of freedom of waves in random media, Phys. Rev. Lett. **103**, 173902 (2009).

[24] Allard P. Mosk, Ad Lagendijk, Geoffroy Lerosey, and Mathias Fink, Controlling waves in space and time for imaging and focusing in complex media, Nat. Photonics **6**, 283 (2012).

[25] Arnaud Derode, Arnaud Tourin, and Mathias Fink, Random multiple scattering of ultrasound. I. Coherent and ballistic waves, Phys. Rev. E **64**, 036605 (2001).

[26] Arnaud Derode, Arnaud Tourin, Julien de Rosny, Mickaël Tanter, Sylvain Yon, and Mathias Fink, Taking advantage of multiple scattering to communicate with time-reversal antennas, Phys. Rev. Lett. **90**, 014301 (2003).

[27] Jonathan Duplicy, Biljana Badic, Rajarajan Balraj, Rizwan Ghaffar, Péter Horváth, Florian Kaltenberger, Raymond Knopp, István Z. Kovács, Hung T. Nguyen, Deepaknath Tandur, and Guillaume Vivier, MU-MIMO in LTE systems, EURASIP J. Wirel. Commun. Netw. **2011**, 1 (2011).

[28] Barry Arons, A review of the cocktail party effect, J. Am. Voice I/O Soc. **12**, 35 (1992).

[29] Simon Haykin and Zhe Chen, The cocktail party problem, Neural Comput. **17**, 1875 (2005).

[30] Gabriel Montaldo, Geoffroy Lerosey, Arnaud Derode, Arnaud Tourin, Julien de Rosny, and Mathias Fink, Telecommunication in a disordered environment with iterative time reversal, Waves Random Medias **14**, 287 (2004).

[31] Gabriel Montaldo, Mickaël Tanter, and Mathias Fink, Real time inverse filter focusing through iterative time reversal, J. Acoust. Soc. Am. **115**, 768 (2004).

[32] Sylvain Yon, Mickael Tanter, and Mathias Fink, Sound focusing in rooms: The time-reversal approach, J. Acoust. Soc. Am. **113**, 1533 (2003).

[33] Michael H. Denison and Brian E. Anderson, Time reversal acoustics applied to rooms of various reverberation times, J. Acoust. Soc. Am. **144**, 3055 (2018).

[34] Guillemette Ribay, Julien de Rosny, and Mathias Fink, Time reversal of noise sources in a reverberation room, J. Acoust. Soc. Am. **117**, 2866 (2005).

[35] Yoichi Ando and Daniel R. Raichel, *Architectural Acoustics: Blending Sound Sources, Sound Fields, and Listeners*. (Springer, New York, NY, 1998).