nature portfolio

Peer Review File

Dual-domain metamaterials co-integrated with a compact ultrasonic transducer for highly directional audio generation

Corresponding Author: Professor Junsuk Rho

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The manuscript presents a "Metasurface-integrated Parametric Array Loudspeaker" (MiPAL), which combines a single-body piezoelectric transducer with a co-designed acoustic metasurface and elastic meta-units. The authors demonstrate that this system can generate highly directional audio over a broad bandwidth (500 Hz to 10 kHz) by leveraging the nonlinear interaction of ultrasound in air. By utilizing a single Langevin transducer enhanced by dual-domain metamaterials (fluid-borne acoustic and structure-borne elastic), the proposed architecture aims to reduce the hardware complexity typically associated with array-based PALs.

The design strategy presented is technically sound, and the combination of numerical simulations with experimental validations appears reliable. The integration of elastic meta-units to suppress parasitic Combination Resonance (CR) modes is a great engineering solution.

However, I have concerns regarding the broad interest and conceptual novelty required for publication. In its current form, the manuscript reads primarily as a specialized technical optimization of the authors' previous work on single-transducer PALs (e.g., Refs. 21-24). To justify publication in a high-impact multidisciplinary journal, the authors should elevate the narrative from "device performance improvement" to "fundamental physical insights". Other comments are appended below.

1. In Fig. 3b, the legend labels the curves using audio frequencies (e.g., 0.5 kHz, 1 kHz). This is confusing because the plot depicts the propagation of the ultrasonic waves. It is suggested to label these curves with the actual ultrasonic frequencies to avoid confusion.

2. In Fig. 3d (without acoustic metasurface), the directivity is shown to be very broad. However, for a source with f = 60 kHz and radius a = 50 mm, the ka value is approximately 55, which typically suggests a highly directional beam for a piston source. Why is the directivity so broad in the absence of the metasurface? Is it solely due to the high-order flexural mode shape?

3. The suppression of the CR modes relies on the precise tuning of the elastic meta-units (mass-spring systems) attached to the plate. How sensitive are these meta-units to manufacturing tolerances (given the resolution of 3D printing). If the local resonance shifts slightly away from the plate's eigenmode due to environmental factors, will the "omnidirectional leakage" return?

4. Figure 5a proposes "in-cabin personal audio delivery" as a key application. However, the results in Fig. 5 appear to be measured in an anechoic chamber. The authors should explicitly clarify the measurement environment in the caption and main text. Actually, practical cabin environments present more challenges such as reflections, scattering, and reverberation.

5. The manuscript acknowledges that the prototype produces "relatively limited SPL" compared to conventional loudspeakers. The authors are suggested to provide a quantitative analysis of the system's efficiency. How much acoustic energy is lost due to the insertion/blocking effect of the acoustic metasurface? Is the low SPL a fundamental limitation of the metasurface's low transmission ratio, or is it merely a power handling issue of the PZT driver? A comparative analysis of the radiation efficiency between the proposed MiPAL and the previous SPPAL is suggested to evaluate.

6. The manuscript lacks specific dimensional details of the critical structures. Please add dimension annotations (e.g., plate radius) directly in Figs. 1 or 2.

7. In Fig. 4f, what are circles and triangles for?

Reviewer #2

(Remarks to the Author)
The manuscript proposes a metasurface-integrated parametric array loudspeaker (MiPAL), which achieves ultra-broadband directional sound emission (spanning 500 Hz to 10 kHz) by combining a compact single-body ultrasonic transducer with an acoustic metasurface. The study is content-rich, featuring rigorous theoretical analysis, simulations, and experimental validation. However, key aspects including novelty, physical insight, and practical significance are insufficiently addressed, and the manuscript thus does not meet the publication standards of Nature Communications.

1.The acoustic metasurface of the MiPAL adopts a unit-cell structure consisting of "radially arranged Helmholtz resonators (HRs) interconnected by space-coiling (SC) channels". Such structures have been utilized in numerous existing studies (e.g., Appl. Phys. Lett. 122, 183506 (2023)). The authors fail to clearly illustrate the innovative differences. Furthermore, the manuscript does not adequately elaborate on strategies to address the dispersion effects of the metamaterial structure over a broad frequency range.

2.The design concept of "integrating a single-body transducer with metamaterial structures" has been proposed in prior research (e.g., Appl. Phys. Lett. 127, 151701 (2025)), and no new breakthroughs in this integration scheme are demonstrated in the current work. Additionally, the manuscript does not compare performance under different transducer geometries scenarios, lacking further validation of the scheme's applicability and flexibility.

3.The core difference-frequency wave (DFW) generation technology described in the manuscript does not show essential distinctions in physical mechanisms compared with existing studies (e.g., Proc. Natl. Acad. Sci. U.S.A. 122(12), e2408975122 (2025); Phys. Rev. Lett. 131, 234001 (2023)). No corresponding physical analysis is provided to confirm performance improvements in the current work.

4.The authors assert that the MiPAL possesses the advantage of "ultra-broadband directional sound emission". However, its broadband range of 500 Hz to 10 kHz does not demonstrate obvious superiority compared with state-of-the-art technologies (e.g., the ultra-broadband acoustic metamaterial lens reported in Appl. Phys. Lett. 127, 151701 (2025), which can omnidirectionally emit sound between 1000 Hz and 10000 Hz). A systematic comparison of key performance indicators (such as broadband range and directivity precision) with current state-of-the-art directional sound technologies is lacking.

Reviewer #3

(Remarks to the Author)
The work reports the design of a novel ultrasonic transducer that leverages artificial structures—including an acoustic metasurface and elastic meta-units—to achieve directional sound. The topic is of clear interest to the acoustic wave engineering community, with promising potential for practical applications. The work presents a comprehensive study combining numerical simulations and experimental validation, including audio tests that demonstrate enhanced directivity. Overall, the paper is well organized and clearly written. However, several points require clarification or further discussion before the work can be fully assessed. My specific comments are listed follow.

• The abstract does not mention the elastic meta-unit design, despite this component being discussed extensively in the main text and playing a key role in the system's performance. Including this element in the abstract would improve consistency and better reflect the core contributions of the work.

• Figure 2a shows multiple geometric parameters associated with both space-coiling channels and Helmholtz resonators in the metasurface design. Since several physical mechanisms may contribute to wave collimation, it is not clear which mechanism dominates the observed directional propagation. A more detailed discussion or quantitative analysis clarifying the relative contributions of these mechanisms would strengthen the interpretation.

• In Fig. 2c, the transmittance appears to remain strong for $t1 \gtrsim 1.5$mm, while the influence of t2 is relatively weak. Given that t1 may also be constrained by other geometric parameters (e.g., total thickness t), it is unclear whether this trend persists over a broader parameter space. Please clarify the explored parameter ranges and discuss the robustness of this observation.

• In Fig. 2g, it is not immediately clear which curve corresponds to the case with elastic meta-units. Adding explicit labels or annotations directly in the figure would improve readability and help readers quickly identify the key results.

• In Fig. S6a, the reduced wavenumber appears to be incorrectly associated with the real-space axis. This labeling issue should be corrected.

• The distinction between modes 1a and 1b, as well as 2a and 2b, is not clearly explained. In addition, the color bar in Fig. 2h does not seem to correspond consistently to the modes involving elastic meta-units.

• In Fig. 3b, a pronounced valley is observed around 0.23m for different audio modulation offsets. Please provide a physical explanation or discussion of the origin of this feature.

• Figure 4d indicates that elastic meta-units enhance performance mainly within specific low-frequency ranges (approximately 0.5–0.8 kHz, ~2 kHz, and ~6.2 kHz). Outside these ranges, it is less clear that the configuration with elastic meta-units consistently outperforms the one without them. If the primary role of the elastic meta-units is to suppress combination-resonance modes, as suggested in Fig. 1g, then the acoustic metasurface appears to be the dominant contributor to high directivity. This point deserves further discussion. Additionally, it would be helpful to include supplementary audio test comparisons—similar to Fig. 5d—directly contrasting configurations with and without the acoustic metasurface.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have fully addressed my concerns, and I am happy to recommend the manuscript for publication.

Reviewer #2

(Remarks to the Author)
The authors have responded point-by-point to initial comments, yet the core flaws—insufficient substantial novelty and inadequate elucidation of fundamental physical mechanisms—remain unresolved. This manuscript does not meet the rigorous publication standards of Nature Communications and it may be more suitable for a specialized journal such as Communications Engineering, where its practical implementation and platform-specific contributions would be better appreciated.
Specific comments are as follows:
1. The authors claim to have expanded discussions on the MiPAL's novelty and physical mechanisms, but only conducted superficial textual rephrasing (e.g., rhetorical emphasis on "Dual-domain metamaterials interwoven") with no core physical mechanism analysis added. The so-called "dual-domain metamaterial integration" is merely a conceptual description, which fails to reveal the multi-physical field coupling laws between the acoustic metasurface and elastic meta-units, nor explain how this coupling achieves performance breakthroughs unavailable with single-domain technologies. Their argument of "dual-domain interweaving rather than simple integration" is only conceptual. No specific theoretical framework or engineering methodology for the proposed "co-design" paradigm is provided, leaving it without practical support.
2. For broadband dispersion control, the supplementary upper-sideband modulation selection and phase-matching RMSE analysis are merely simple numerical simulations. They fail to uncover the physical nature of the coupling between the transducer's structural dispersion and the acoustic metasurface's dispersion, and no quantitative theoretical model for dispersion control is established. Thus, generalizability and effectiveness of the strategy cannot be verified.
3. The difference-frequency wave (DFW) generation mechanism has no essential difference from existing studies. The authors' claim of "systematic innovation via dual-domain integration" (as opposed to previous "local nonlinear effects") completely evades the core issue: the manuscript still lacks substantive physical analysis and experimental validation to prove that this dual-domain integration brings fundamental improvements to DFW generation compared with state-of-the-art technologies.

Reviewer #3

(Remarks to the Author)
I believe the authors have addressed my concerns and I would like to recommend the manuscript for publication.
My only minor suggestion concerns the revised title. I am not sure if the term "dual-domain metamaterials" accurately captures the integration of the elastic and acoustic design aspects presented in the work. Besides, the metamaterial design itself appears to be the key element enabling directional audio. It may therefore be worth reconsidering whether the phrase "dual-domain metamaterials interwoven with a compact ultrasonic transducer" is suitable, or if a more concise title could better highlight the central contribution.

Response to Reviewers

Dual-domain metamaterials interwoven with a compact ultrasonic transducer for highly directional audio generation (NCOMMS-25-100737-T)

Woongji Kim[1†], Beomseok Oh[1†], Wonkyu Moon[1*], Junsuk Rho[1,2,3,4*]

[1]Department of Mechanical Engineering, Pohang University of Science and Technology (POSTECH), Cheongam-ro 77, Pohang-si, 37683, Gyeongsangbuk-do, Republic of Korea. [2]Department of Chemical Engineering, Pohang University of Science and Technology (POSTECH), Cheongam-ro 77, Pohang-si, 37683, Gyeongsangbuk-do, Republic of Korea. [3]Department of Electrical Engineering, Pohang University of Science and Technology (POSTECH), Cheongam-ro 77, Pohang-si, 37683, Gyeongsangbuk-do, Republic of Korea. [4]POSCO-POSTECH-RIST Convergence Research Center for Flat Optics and Metaphotonics, Pohang University of Science and Technology, Cheongam-ro 77, Pohang-si, 37683, Gyeongsangbuk-do, Republic of Korea.

*Corresponding author(s). E-mail(s): wkmoon@postech.ac.kr; jsrho@postech.ac.kr; Contributing authors: wj.kim@postech.ac.kr; bs.oh@postech.ac.kr; †These authors contributed equally to this work.

General Response

We would like to express our sincere gratitude to the editor and the reviewers for their insightful evaluations and constructive feedback for our manuscript retitled **"Dual-domain metamaterials interwoven with a compact ultrasonic transducer for highly directional audio generation"** submitted to *Nature Communications* (NCOMMS-25-100737-T). Their suggestions have been instrumental in refining the conceptual clarity and elevating the academic rigor of our manuscript.

In this response, we have thoroughly addressed all comments from the reviewers through comprehensive additional experiments, rigorous analysis, simulations, and discussions. We have carefully examined the referees' primary concerns, particularly regarding (i) the conceptual novelty and (ii) the physical insights and impact of our proposed architecture. This revision process has further solidified our belief that our work is not merely a technical refinement of existing methods, but a transformative advancement in metamaterial-integrated platforms. We would like to highlight that our work centers on dual-domain wave manipulation, enabled by the monolithic integration and sophisticated structural-acoustic coupling between the electromechanical transduction, acoustic metasurfaces, and elastic meta-units. By orchestrating this synergistic interplay where the transducer's vibration and its structure-borne acoustic radiation is simultaneously governed by both elastic meta-units and the acoustic metasurface, we demonstrate a unique capability to overcome the inherent key limitations of existing parametric array loudspeakers (PALs).

Here, we wish to articulate the broader scientific impact and significance of this work through the following key foundational contributions that define why this approach represents a paradigm shift in metamaterials-augmented platforms:

1. While metasurfaces have previously been applied for wavefront manipulation, our work pioneers the interweaving of dual-domain metamaterials (acoustic and elastic) with a compact, modular, and cost-effective single-driver ultrasonic transducer. Unlike previous studies that treat the transducer as a fixed "black-box" source, our work introduces a monolithic co-design framework. By integrating elastic meta-units directly into the transducer's radiating structure and coupling them with an acoustic metasurface, we effectively resolve the fundamental bottlenecks of existing studies. Furthermore, our work offers new physical insights into the interplay between structure-borne elastic waves and airborne acoustics, bridging the gap between structural mechanics and wave engineering for a wider scientific community. We believe such unprecedented "device-level integration of metamaterials" with the real-device highlight the far-reaching impact of metamaterials in practical applications, promising to influence a wide array of fields.
2. The proposed architecture is not a single-application device but a highly adaptable platform capable of meeting diverse application requirements. We demonstrate through extended analyses that the

MiPAL can be tailored for various scenarios representing a transformative demonstration of how integrated metamaterials can be leveraged to create a versatile and scalable platform for advanced sound manipulation. This inherent design flexibility underscores that our work provides not just a standalone device, but a robust engineering roadmap for addressing multi-physics wave-control challenges at the practical device level.

In the detailed point-by-point response that follows, we have revised the manuscript and Supplementary Information accordingly and added further explanations and supporting content to address each concern. We once again thank the Reviewers for their invaluable time and constructive evaluation, and we respectfully hope that our revised manuscript will now be considered favorably in the next stage of the review process.

Reviewer 1

Comment 1. The manuscript presents a "Metasurface-integrated Parametric Array Loudspeaker (MiPAL)", which combines a single-body piezoelectric transducer with a co-designed acoustic metasurface and elastic meta-units. The authors demonstrate that this system can generate highly directional audio over a broad bandwidth (500 Hz to 10 kHz) by leveraging the nonlinear interaction of ultrasound in air. By utilizing a single Langevin transducer enhanced by dual-domain metamaterials (fluid-borne acoustic and structure-borne elastic), the proposed architecture aims to reduce the hardware complexity typically associated with array-based PALs.

Response 1. We would like to thank the referee for the thorough evaluation and the accurate summary of our work. We are particularly encouraged that the referee recognized the core intention of our architecture: to overcome the long-standing hardware complexity of phased-array PALs by interweaving dual-domain metamaterials within a compact ultrasonic transducer architecture. We have significantly revised the manuscript for greater clarity, incorporating additional simulations, experiments, and discussions to fully address your concerns as well as those of the other referees. Below, we provide detailed point-by-point responses to each of your comments.

Comment 2. The design strategy presented is technically sound, and the combination of numerical simulations with experimental validations appears reliable. The integration of elastic meta-units to suppress parasitic Combination Resonance (CR) modes is a great engineering solution.

Response 2. We sincerely thank the reviewer for the positive assessment of our design strategy and the recognition of our experimental validation. We appreciate your positive feedback on the technical soundness of our novelty of design strategy and the effectiveness of integrating elastic meta-units to suppress parasitic CR modes. To further underscore the scientific impact and novelty of this approach, we would like to highlight the following points:

- **From theoretical abstraction to functional integration**
  While locally resonant elastic metamaterials have been extensively studied in elastics, most research has remained confined to idealized, infinite media—such as infinite elastic plates—focused primarily on idealized wave attenuation[R1–R3]. Our work represents a transformative transition from these conceptual demonstrations to a high-potential, real-device implementation.

- **Enhancing fidelity through CR mode suppression**
  Beyond the demonstration of the simple frequency-domain suppression, the integration of elastic meta-units provides critical improvements in the real-time domain audio experience. As discussed in Supplementary Notes 12 and 13, the CR modes cause "ringing" effects. This leads to temporal smearing and perceptual degradation in the received sound quality. Our elastic meta-units effectively suppress these phenomena, enhancing transient clarity and ensuring a crisper audio output. These improvements are qualitatively demonstrated in the provided audio samples (Supplementary Videos 1 to 5), which offer a direct comparison of the acoustic fidelity. This also represents a vital contribution from a psychoacoustic perspective.

- **Providing insight for advanced audio engineering**
  The engineering challenge we address parallels the principles of commercial balanced mode radiator (BMR) technology[R4,R5]. Conventional diaphragms suffer from "breakup" and "power beaming" when mechanical wavelengths become similar to the diaphragm's dimensions. BMR technology manages these structural modes by "re-balancing" the diaphragm with precisely placed masses at average nodal positions to ensure wide sound power dispersion and a flat response. Similarly, our elastic meta-units offer a light-weight, deep sub-wavelength for structural mode control. These structures allow for precise suppression of parasitic resonances without affecting the device's footprint, providing significant insights for the next generation of compact, high-fidelity audio engineering.

To clearly demonstrate the novelty and impact of the elastic meta-units, we have revised the manuscript as follows:

- **Supplementary Note 7**: In response to Comment 6 by Reviewer 1, we have expanded the discussion regarding the practical implementation of elastic meta-units, emphasizing their transition from theoretical models to functional, real-world device applications.

Response 2. (*continued*) Furthermore, to substantiate the physical validity of the elastic meta-units and provide a comprehensive technical foundation for a broad readership, we have elaborated on our experimental protocols and validation procedures as follows:

- **Supplementary Note 10**: We have documented the calibration of the acoustic measurement system required to isolate true acoustic fields. This rigorous suppression of measurement artifacts was fundamental to identifying the CR effects and validating the effectiveness of the elastic meta-units.
- **Supplementary Note 11**: We have detailed the rigorous experimental methodology used to characterize the plate's vibration characteristics, specifically focusing on the CR modes. This multi-stage validation of the measurement system's linearity and intermodulation distortion floor confirms that the observed CR phenomena are authentic physical effects rather than instrumentation artifacts.

Comment 3. However, I have concerns regarding the broad interest and conceptual novelty required for publication. In its current form, the manuscript reads primarily as a specialized technical optimization of the authors' previous work on single-transducer PALs (e.g., Refs. 21–24). To justify publication in a high-impact multidisciplinary journal, the authors should elevate the narrative from "device performance improvement" to "fundamental physical insights". Other comments are appended below.

Response 3. We appreciate the reviewer's constructive critique regarding the broad interest and conceptual novelty of our work. We agree that for a high-impact multidisciplinary journal such as *Nature Communications*, the narrative must transcend technical optimization and offer fundamental insights that resonate across diverse scientific fields. To address this, we have fundamentally restructured the manuscript's narrative—starting from the title and abstract to the discussion—to emphasize that our device is not merely a performance-enhanced version of our previous studies[21–24]. Instead, it represents a paradigm shift in wave-manipulation platforms through the following reasons:

- **Dual-domain metamaterial integration and engineering**
  We introduce a pioneering framework that synergistically combines acoustic metasurfaces and elastic meta-units within a single, modular, and compact transducer architecture. This approach represents a fundamental departure from conventional single-domain metamaterial research, which typcially focuses on manipulating waves in a single physical domain. We firmly believe that this dual-domain integration demonstrates the immense potential of metamaterials to govern complex multi-physics phenomena simultaneously. Such a paradigm shift not only enhances the performance of personal audio systems but also provides a significant multidisciplinary impact, offering a versatile design methodology for diverse fields such as medical ultrasound, structural health monitoring, and advanced audio devices.

- **Overcoming the limitations of previous studies**
  Our work demonstrates a paradigm shift by effectively resolving the long-standing challenges inherent in traditional phased-array PALs—such as prohibitive hardware complexity, limited design flexibility, and high production costs—and the critical limitations of earlier single-transducer PALs (stepped-plate PALs, SPPALs[R21–R24]), which were limited by the design complexity of intricate step designs, low transducer efficiency stemming from inherent material damping (as also detailed in Supplementary Note 16 (newly added) and our response for Comment 8 by Reviewer 1), and a fundamental loss of beam control caused by combination resonance (CR) mode. By replacing these complex mechanical designs with the interweaving of dual-domain metamaterials, our approach achieves more efficient and adaptable wave-control solutions with a significantly more compact footprint. To our knowledge, this approach is unprecedented, presenting an innovative method to overcome conventional design constraints. We believe this advancement represents more than a "technical refinement" it constitutes a foundational reimagining of how metamaterials can be integrated to overcome deep-seated engineering constraints in functional devices.

- **Versatile and scalable platform for advanced wave engineering**
  We demonstrate that our architecture can be tailored for various applications, showcasing its potential as a highly adaptable platform for metamaterials-enabled advanced wave control. This versatility opens new opportunities for research in metamaterials and wave engineering, extending its impact well beyond conventional audio device innovation.

We understand the importance of framing our research within a context that highlights its fundamental insights rather than solely as a technical optimization of previous designs. In response to your valuable feedback, we have substantially revised the manuscript to elevate the narrative and emphasize the foundational contributions of our work.

Response 3. (*continued*) The following outlines the revisions that were made:

- Revisit the **title** and **abstract**
  The manuscript title has been revised to better reflect the core novelty and broader scientific impact of our work. As you acknowledged in Comment 2 by Reviewer 2 (and as similarly noted in Comment 2 by Reviewer 3), the effectiveness of elastic meta-unit-based engineering represents a key contribution of our study. The abstract has also been rewritten to emphasize the impact and transformative potential of our architecture.

- **Introduction** and **discussion** enhancement
  We have revised the introduction and discussion sections to clearly establish broader impact of our work and strengthen its narrative coherence.

The corresponding changes in the title, abstract, and manuscript are as follows:

Title	
Original: Metasurface-integrated compact ultrasonic transducer for highly directional audio generation	**Revised:** Dual-domain metamaterials interwoven with a compact ultrasonic transducer for highly directional audio generation

Abstract	
Original: ... Here, we report an alternative form factor: a compact, single-body ultrasonic transducer augmented with an acoustic metasurface, termed metasurface-integrated parametric array loudspeakers (MiPALs). Unlike conventional PALs, the MiPAL utilizes a single piezoelectric driver combined with a co-designed metasurface, yielding a simple and cost-effective architecture that achieves compact directional sound systems. To realize this, we exploit a unified design strategy that integrates theoretical modeling, transducer-metasurface co-design, and comprehensive quantitative analyses of both linear and non-linear acoustic fields. ... These findings underscore the potential of metasurface-transducer integrated platforms for advanced audio systems, laying the foundation for developments in immersive audio, spatial sound delivery, and extended reality.	**Revised:** ... Here, we report an alternative form factor: a compact, single-body ultrasonic transducer integrated with dual-domain metamaterials, termed metamaterials-integrated parametric array loudspeakers (MiPALs). Unlike conventional PALs, the MiPAL utilizes a single piezoelectric driver interwoven with dual-domain metamaterials—consisting of a co-designed acoustic metasurface and elastic meta-units—to yield a simple, cost-effective, and highly integrated architecture. To realize this, we exploit a unified design strategy that integrates theoretical modeling, transducer-metamaterials co-design, and comprehensive quantitative analyses of both linear and non-linear acoustic fields. ... These findings underscore the potential of metamaterial-transducer integrated platforms for advanced audio systems, laying the foundation for developments in immersive audio, spatial sound delivery, and extended reality.

Response 3. (*continued*)

In the Introduction (Lines 26–28 and Lines 39–46)

Added:

... This reliance fundamentally limits design flexibility, as the transducer is conventionally treated as a given black-box source with predefined operating characteristics, rather than as a designable variable.

... Furthermore, the inherent material damping often associated with integrated step structures—particularly those utilizing polymers—leads to significant efficiency degradation, representing a bottleneck for achieving sufficient sound volume (see Supplementary Note 16 for detailed discussion). Crucially, earlier SPPALs inevitably suffer from the compromise of their core functionality—directional sound reproduction—due to combination resonance (CR)[24]. This parasitic phenomenon, an inherent side-effect of nonlinear intermodulation, generates nearly omnidirectional acoustic leakage that represents a primary drawback of existing single-body PAL architectures.

In the Introduction (Lines 47–58)

Original:

... In this paper, we propose a new form factor for PALs—the metasurface-integrated PAL (MiPAL)—which combines a modular, single-body piezoelectric ultrasonic transducer with a co-designed acoustic metasurface. The integrated metasurface features subwavelength spatial resolution and high design flexibility, both considered well-known advantages of metamaterials[25–32,41], and effectively redirects the non-directional ultrasound generated by the transducer's flexural vibration into highly collimated and directional wavefront. As a result, directional audible sound is produced through DFW generation via nonlinear interactions of high-intensity ultrasonic waves. Furthermore, embedding locally resonant elastic meta-units on the radiating plate effectively suppresses the omnidirectional structure-borne sound associated with the combination-resonance (CR) mode—an undesired by-product nonlinearly induced through intermodulation[24] and one of the key drawbacks of SPPALs.

Revised:

... In this paper, we propose a new form factor for PALs—the metamaterials-integrated PAL (MiPAL)—which consists of a modular, single-body piezoelectric ultrasonic transducer interwoven with dual-domain metamaterial architectures. Within this unified system, the acoustic metasurface serves as the primary functional layer; by leveraging subwavelength spatial resolution and high design flexibility, both considered well-known advantages of metamaterials[25–32], it effectively redirects the non-directional ultrasound generated by the transducer's flexural vibration into highly collimated and directional wavefront. As a result, directional audible sound is produced through DFW generation via nonlinear interactions of high-intensity ultrasonic waves. Simultaneously, the elastic meta-units act as a critical compensatory layer. By embedding deep-subwavelength tiny locally resonant structures[33] onto the transducer's radiating plate, we effectively suppress the omnidirectional sound leakage associated with the CR mode that undermines the core functionality of the PAL by compromising its defining high-directivity.

Response 3. (*continued*)

In the Introduction (Lines 61–64)

Original:	Revised:
... This framework comprehensively encompasses the design of transducer components, geometries of metamaterial structures, and rigorous full-field analysis of both linear and nonlinear acoustic behavior.	... This framework encompasses the comprehensive design of the MiPALs, integrating electro-mechanical transduction with dual-domain wave manipulation through theoretical investigation and rigorous full-field analyses of both linear and nonlinear behaviors in acoustic and elastic regimes.

In the Discussion (Lines 297–308)

Original:	Revised:
... In this work, we introduce the MiPAL, which fundamentally redefines the architecture for personal audio systems. By integrating an acoustic metasurface with a compact single-body ultrasonic transducer, the MiPAL achieves broadband directional sound generation without relying on costly multi-emitter arrays constructed from acoustically oversized off-the-shelf ultrasonic emitters or the complex phase-control electronics[1,19]. The metasurface redirects the incoherent flexural-mode radiation into collimated wavefront, while embedded elastic meta-units effectively suppress parasitic vibration modes that otherwise cause undesired acoustic leakage. As a result, the MiPAL demonstrates directional audio across an ultra-broad frequency range from 500 Hz to 10 kHz—covering over four octaves. This compact platform overcomes the cost, complexity, and scalability limitations of conventional PALs.	... In this work, we introduce the MiPAL, which fundamentally redefines the architecture for advanced personal audio systems through the dual-domain integration of metamaterials. By synergistically combining an acoustic metasurface and elastic meta-units with a compact single-body ultrasonic transducer, the MiPAL achieves broadband directional sound generation without relying on costly multi-emitter arrays or the complex phase-control electronics typically required in conventional systems[R1,R19]. The metasurface redirects the incoherent flexural-mode radiation into collimated wavefronts, while embedded elastic meta-units effectively suppress parasitic vibration modes that otherwise cause undesired omnidirectional leakage. As a result, the MiPAL demonstrates directional audio performance across an ultra-broad frequency range from 500 Hz to 10 kHz—covering over four octaves within a physically compact platform. Our approach overcomes the cost, complexity, and scalability limitations of conventional PALs, paving the way for widespread adoption in diverse applications.

Response 3. (*continued*)

In the Discussion (Lines 319–323)

Added:

... Furthermore, we foresee that the interweaving of architected metamaterials with real-world functional platforms will catalyze a paradigm shift across diverse fields, including biomedicine, non-destructive testing, and energy harvesting, by enabling unprecedented control over wave-matter interactions.

Comment 4. In Fig. 3b, the legend labels the curves using audio frequencies (e.g., 0.5 kHz, 1 kHz). This is confusing because the plot depicts the propagation of the ultrasonic waves. It is suggested to label these curves with the actual ultrasonic frequencies to avoid confusion.

Response 4. We agree with the reviewer that labeling the propagation curves with audio frequencies in an ultrasonic context may lead to confusion. To address this, we have updated Fig. 3b and its corresponding caption to reflect the actual primary ultrasonic frequencies. We have implemented the following changes:

- Legend update: The legend in Fig. 3b now labels the curves using the ultrasonic frequencies to produce the intended audio frequencies.
- Caption revision: The caption for Fig. 3b has been updated to clarify that these curves represent the propagation of the ultrasonic primary tones:

The caption of Fig. 3b	
Original: Simulated and measured propagation curves at multiple modulating audio frequencies (0.5, 1, 2, 4, and 8 kHz) with 60 kHz primary ultrasound. The SPL peaks near 0.45 m correspond to the designed target hearing distance (dotted line).	**Revised:** Simulated and measured propagation curves at multiple primary ultrasonic frequencies (60.5, 61, 62, 64, and 68 kHz) corresponding to audio modulation. The SPL peaks near 0.45 m (dotted line), aligned with the target hearing distance of 0.5 m.

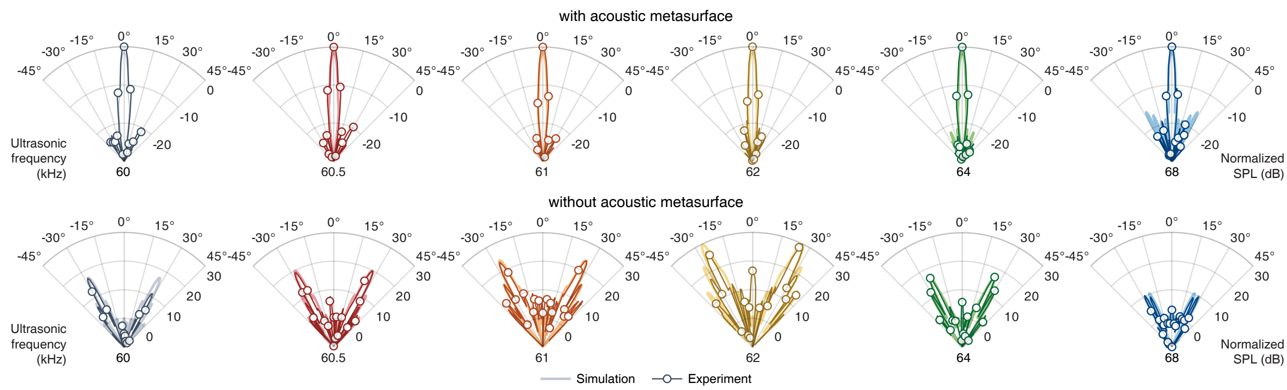
Comment 5. In Fig. 3d (without acoustic metasurface), the directivity is shown to be very broad. However, for a source with $f = 60$ kHz and radius $a = 50$ mm, the ka value is approximately 55, which typically suggests a highly directional beam for a piston source. Why is the directivity so broad in the absence of the metasurface? Is it solely due to the high-order flexural mode shape?

Response 5. Thank you for your comment. Yes, the broad directivity observed in Fig. 3d without the acoustic metasurface is primarily attributed to the excitation of high-order flexural modes in the radiating plate (please also refer to Fig. 1f). While a high ka value generally indicates a directional beam for an ideal piston source, the actual radiation pattern is significantly influenced by the mode shapes of the vibrating structure. In our case, the transducer's circular plate does not behave as a perfect rigid piston owing to the elasticity and resulting flexural modes, which lead to complex vibration patterns and consequently "non-directivity". The high-order flexural modes introduce different radiation patterns.

For clarity, we add the representative 1D beam patterns with and without the acoustic metasurface in below. We can clearly see how the metasurface effectively shapes the beam pattern, enhancing directivity by compensating for the incoherent radiation arising from the complex mode shapes of the plate.

In Supplementary Note 9

Added:



Supplementary Fig. 19 | Simulated and measured beam patterns of ultrasound at representative ultrasonic frequencies (60, 60.5, 61, 62, 64, and 68 kHz) with acoustic metasurface (top row) and without acoustic metasurface (bottom row).

To address this point in the manuscript, we have revised the relevant part as follows:

In the subsection 'Experimental characterization of the MiPAL' (Lines 202–205)

Original:

... In contrast, the introduction of the metasurface substantially enhances beam directivity, producing a sharply confined main lobe along the propagation axis.

Revised:

... In contrast, the introduction of the metasurface substantially enhances beam directivity, producing a sharply confined main lobe along the propagation axis (additional beam patterns at representative frequencies are provided in Supplementary Figs. 8 and 19).

Comment 6. The suppression of the CR modes relies on the precise tuning of the elastic meta-units (mass-spring systems) attached to the plate. How sensitive are these meta-units to manufacturing tolerances (given the resolution of 3D printing). If the local resonance shifts slightly away from the plate's eigenmode due to environmental factors, will the "omnidirectional leakage" return?

In Supplementary Note 7

Original:

The designed elastic meta-units were fabricated using SLA 3D printing following the procedures in Methods, and their performance was experimentally validated in two stages before integration into the MiPAL transducer. In the first stage (Supplementary Fig. 11b), the velocity frequency response of the mass element in each meta-unit was measured using an LDV. Although SLA printing provides high geometric precision, the material properties inevitably exhibit some variation; therefore, three groups of meta-units with spring lengths $l_s = \{4.3, 4.5, 4.7\}$ mm were fabricated and characterized. Their measured responses show distinct resonance peaks, from which the group-averaged resonance frequencies were determined.

In the second stage (Supplementary Fig. 11c), the circular plate was instrumented with 18 selected meta-units, and the resulting mechanical impedance was measured. Comparison with the impedance of the bare plate allows identification of the combination resonance frequency, confirming that the attached meta-units introduce a strong localized dynamic interaction near the designed bandgap.

Revised:

The designed elastic meta-units were fabricated using SLA 3D printing following the procedures in Methods. Given that the suppression of combination resonance modes relies on precise tuning, an iterative calibration and characterization procedure was implemented to account for variations in material properties and manufacturing tolerances.

In the first stage, the mechanical impedance of the bare circular plate was measured to experimentally verify the exact target resonance frequencies of the plate modes (Supplementary Fig. 11c). To ensure the meta-units were accurately tuned to these frequencies, an iterative design-and-test loop was performed. Initially, trial meta-units were designed via FEM using the manufacturer's nominal material properties[S10]. Three sets of samples were fabricated with varying spring lengths (l_s±0.2 mm) to sample the design space. The density of the printed resin was estimated by measuring the mass of the samples. Subsequently, the velocity transmissibility of each sample was measured using an laser Doppler vibrometer (LDV) (Supplementary Fig. 11b), and the actual Young's modulus and quality factor was back-calculated by matching the experimental resonance peaks with FEM results.

Based on these calibrated material properties, the elastic meta-units were redesigned to precisely target the plate's eigenfrequencies. This process was repeated iteratively—fabricating, measuring, and refining—until the measured resonance frequency of the meta-units matched the target mode frequency. Once validated, the selected meta-units were integrated into the plate for final characterization.

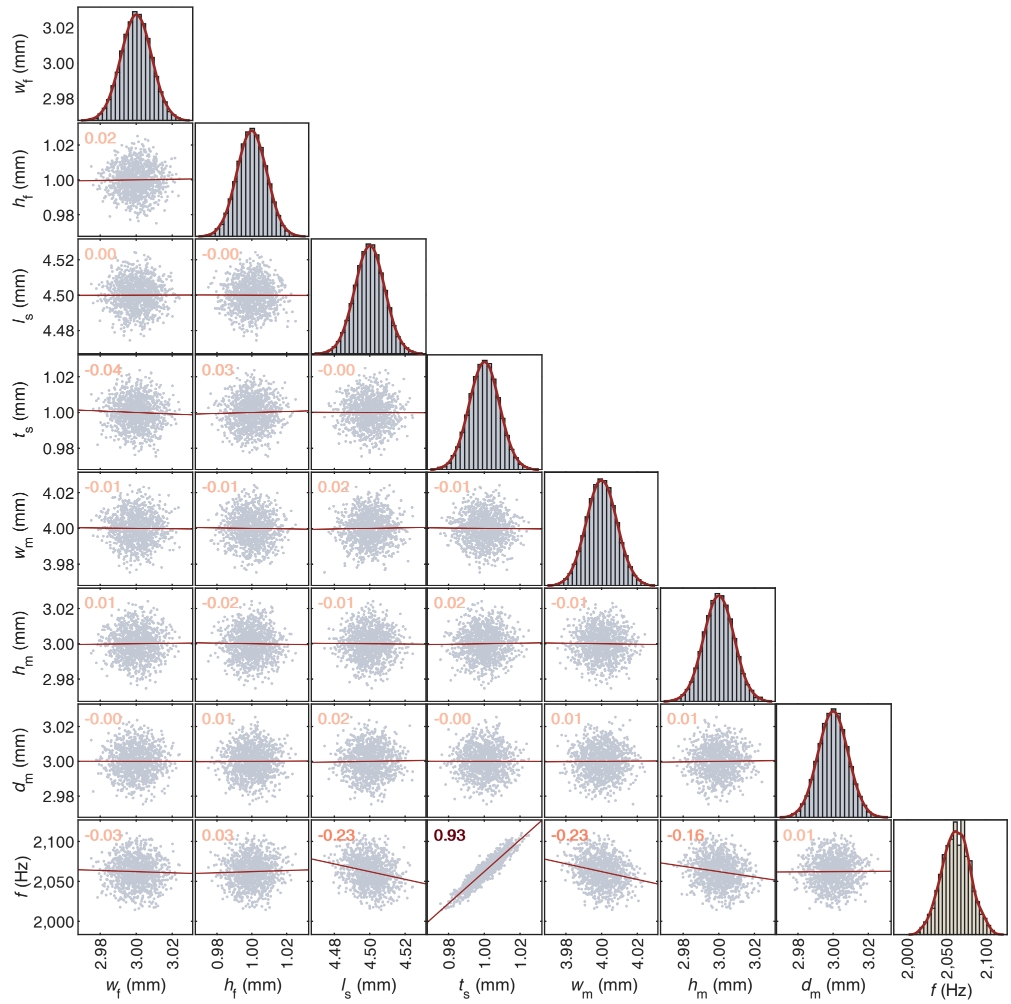
During the integration, the meta-units were attached to the plate using a cyanoacrylate adhesive, which was also used for the jig during the individual unit characterization. To minimize assembly-induced variance, a standardized attachment protocol was strictly followed by a single operator, ensuring consistent glue thickness and placement. In the second stage (Supplementary Fig. 11c), the mechanical impedance of the instrumented plate was measured to confirm that the attached meta-units introduced the required localized dynamic interaction near the designed bandgap.

Response 6. (*continued*) We thank the reviewer for this insightful comment regarding the practical implementation of the meta-units. To address this, we have first provided a detailed description of the iterative fabrication and tuning procedure to clarify how precise spectral alignment is achieved as the above (Supplementary Note 7).

To address the concern regarding manufacturing tolerances, we performed a statistical sensitivity analysis based on the specific resolution of our SLA 3D printer. By mapping the printer's accuracy to a Gaussian distribution, we quantitatively evaluated the resonance frequency variability, confirming that the manufacturing process provides sufficient precision for effective suppression as following (Supplementary Fig. 12):

In Supplementary Note 7

Added:



Supplementary Fig. 12 | Correlation analysis between geometric tolerances and local resonance frequency for the elastic meta-unit. The analysis assumes a Gaussian distribution for manufacturing errors, where the standard deviation ($\sigma \approx$ 8.33 μm) is derived from the 3D printer's stated accuracy (± 25 μm) using the 3σ rule. Based on the nominal dimensions of the second-order mode (Supplementary Table 3), the resulting resonance frequency exhibits a mean of 2062.2 Hz with a standard deviation of 17.6 Hz. The diagonal subplots display the distribution of each parameter via histograms, overlaid with kernel density estimations.

Response 6. (*continued*)

- **Manufacturing tolerance**
  The sensitivity analysis revealed that the resonance frequency of the meta-units has a standard deviation of 17.6 Hz, which is in good agreement with the experimental deviations observed across multiple samples (approximately 22 to 52 Hz; Supplementary Fig. 11b). This indicates that the current manufacturing method provides sufficient precision to effectively suppress the targeted combination resonance modes.

In Supplementary Note 7

Added:

Manufacturing tolerance and sensitivity analysis

To quantitatively assess the impact of manufacturing tolerances on the performance of the elastic meta-units, a statistical sensitivity analysis was performed using a correlation matrix approach (Supplementary Fig. 12). The analysis utilized the official specifications of the SLA 3D printer (ProJet 7000 HD; 3D Systems), which cites a base accuracy of ±25 μm[S11]. Following the 3σ rule[S12], this tolerance interval was mapped to a 3σ Gaussian distribution ($\sigma \approx 8.33$ μm). For this study, 1,000 samples were generated using Latin hypercube sampling to ensure an efficient and uniform coverage of the parameter space. While surface roughness was neglected to focus on macro-geometric variance, the geometric parameters were assumed to follow a Gaussian distribution.
The resulting resonance frequency for the second-order meta-unit mode exhibits a mean of 2062.2 Hz with a standard deviation of 17.6 Hz. As shown in Supplementary Fig. 12, the histogram (bottom-right) for the resonance frequency slightly deviates from a Gaussian profile; however, the kernel density estimation (line) confirms that the distribution remains tightly clustered around the target frequency. Notably, this simulated standard deviation (17.6 Hz) is in good agreement with the experimental deviations observed in Supplementary Fig. 11b (approximately 22 to 52 Hz across 10 samples per group). Ultimately, these results demonstrate that the current manufacturing method provides sufficient precision to effectively suppress the targeted combination resonance modes. While this prototyping approach is suitable for validating academic feasibility through precise iterative tuning, further improvements in frequency consistency and suppression performance are anticipated with the transition to well-controlled, high-volume industrial manufacturing processes.

Response 6. (*continued*) Furthermore, we conducted a sensitivity analysis to evaluate the impact of environmental factors—specifically thermal softening—on the suppression performance and omnidirectional leakage. Our results demonstrate that even under a significant ±20% variation in Young's modulus, the mechanical impedance of the meta-units remains high enough to prevent the restoration of the plate's original high-amplitude vibration modes as following (Supplementary Fig. 13):

- **Resonant frequency shifts**

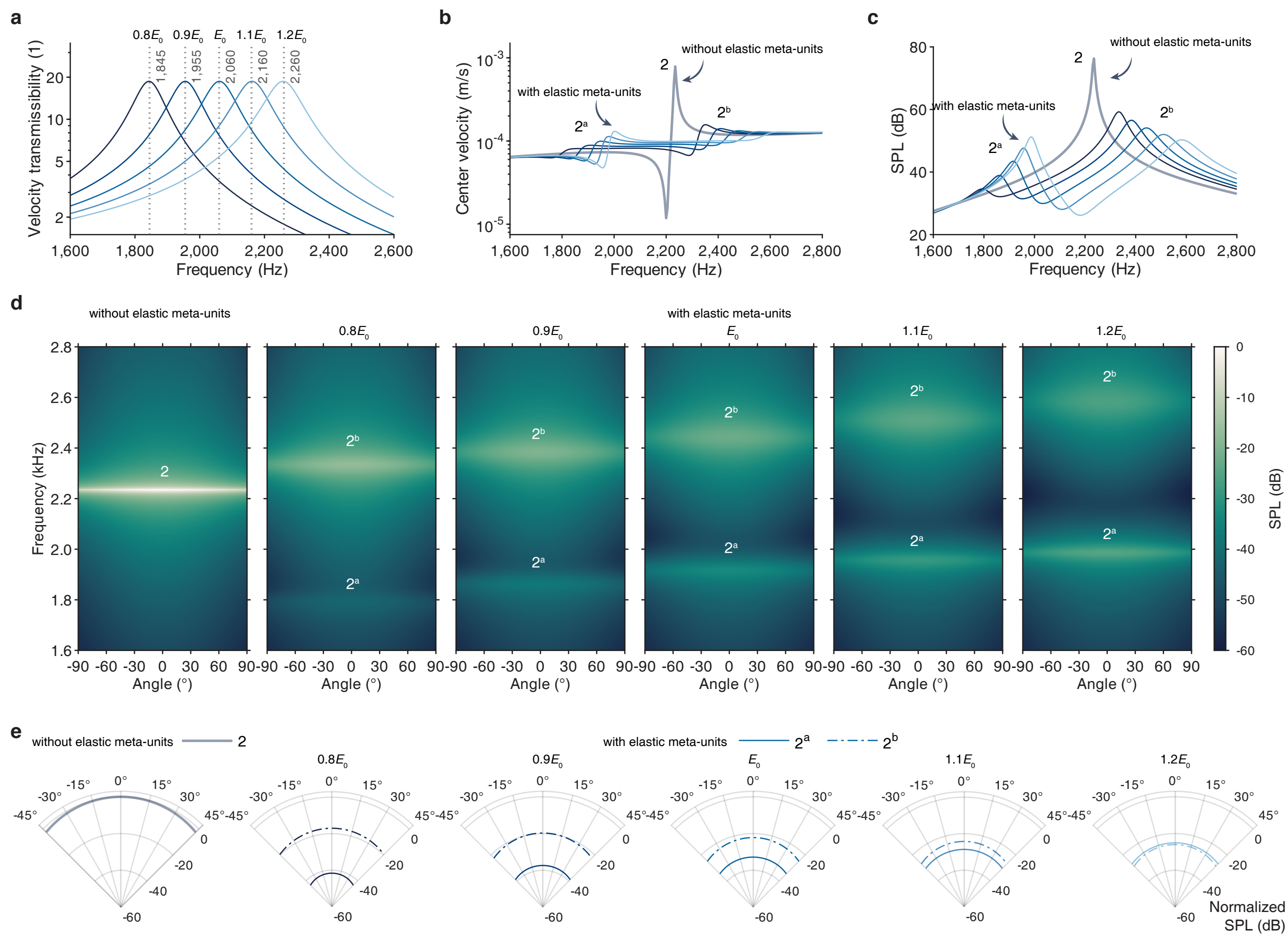
  As expected, variations of ±10% and ±20% in Young's modulus lead to predictable shifts in the local resonance frequency of the meta-units. This interaction with the plate modes results in the emergence of split modes (2^a and 2^b) from the original 2nd-order mode.

- **Omnidirectional leakage**

  Despite these detunings, the plate's center velocity frequency response indicates that the resonance peaks of the split modes remain substantially lower than that of the bare plate.

In Supplementary Note 7

Added:



Supplementary Fig. 13 | Sensitivity analysis of resonant suppression and omnidirectional leakage under stiffness variations. **a,** Local resonance frequency shifts of the meta-unit for ±10% and ±20% variations in Young's modulus in the velocity transmissibility. **b,** Frequency response of plate center velocity showing the split modes (2^a, 2^b) and sustained attenuation despite detuning. **c,** SPL at 0.5 m highlighting suppression depths for all stiffness cases compared to the bare plate. **d, e,** Normalized beam patterns at 0.5 m for the split modes, normalized to the on-axis peak of the original 2nd-order mode.

Response 6. (*continued*)

In Supplementary Note 7

Added:

Sensitivity analysis of combination resonance suppression and associated omnidirectional leakage

The efficacy of the elastic meta-units in suppressing combination resonance depends on the precise spectral alignment between the local resonance of the units and the structural eigenmodes of the plate. Given that the Young's modulus (E) of the 3D-printed material (Accura ClearVue[S10]; 3D Systems) forming the meta-unit is sensitive to environmental factors (e.g., temperature), this section evaluates the robustness of the suppression performance under a ±20% variation in stiffness.

First, the impact of thermal expansion on the meta-unit's geometry was evaluated to isolate it from stiffness-induced shifts. Based on the coefficient of thermal expansion (122 µm/m℃; while the ProJet 7000 HD was utilized for fabrication, the values for SLA printers were adopted for the simulation due to the unavailability of manufacturer-specific data.), a 10 mm spring element expands by only 1.22 µm/℃. Even with a 20 ℃ temperature rise, the resulting 24.4 µm expansion remains within the printer's native manufacturing tolerance (±25 µm), exerting a negligible effect on the resonance frequency (Supplementary Fig. 12). Consequently, geometric expansion is not a primary driver of detuning; instead, the dominant factor is the thermal softening of the material.

All acoustic and structural simulations in this analysis were performed under a linear driving condition of 50 V. While the actual combination resonance is an inherently nonlinear phenomenon arising from the amplitude modulation scheme of the parametric array, the transducer's specific nonlinear electromechanical conversion characteristics are primarily determined through experimental characterization rather than closed-form analytical models. Therefore, linear operation is utilized as a reliable proxy to infer the behavior of resonant shifts and the resulting magnitude of omnidirectional leakage, noting that the absolute amplitudes in this study appear higher than those observed during actual parametric modulation.

The resonance frequency of the elastic meta-unit is fundamentally governed by the stiffness of the spring element. A parametric sweep reveals that variations of ±10% and ±20% in E result in predictable frequency shifts, as shown in Supplementary Fig. 13a. This interaction between the plate and the 18 meta-units leads to the emergence of split modes, denoted as 2^a and 2^b, which originate from the original 2nd-order plate mode (Mode 2). Comparisons of the plate's center velocity frequency response indicate that even with a ±20% shift in E, the magnitude of the resonance peaks remains substantially lower than the bare plate response (Supplementary Fig. 13b). While ideal tuning provides maximum attenuation, these detuned states still introduce sufficient mechanical impedance to prevent the restoration of the original mode's high-amplitude vibration (Fig. 2h).

To evaluate the omnidirectional leakage, the sound pressure level was calculated at a distance of 0.5 m (Supplementary Fig. 13c). In all cases, the meta-units provide clear suppression of modes 2^a and 2^b compared to the bare plate, although the specific suppression depth of each split mode varies with E.

Response 6. (*continued*) Finally, we have included a discussion on advanced design strategies, such as asymmetric frequency distribution[S14], to further enhance the system's robustness against real-world environmental fluctuations in practical applications.

We would like to express our sincere gratitude once again for this highly constructive comment. The reviewer's insight prompted us to perform a much more rigorous sensitivity analysis and to think more deeply about the long-term reliability of these elastic meta-units. As a result, we have significantly expanded Supplementary Note 7 to include formal documentation of manufacturing tolerances and environmental sensitivities. We believe these additions have substantially strengthened the technical depth and practical relevance of our work.

In Supplementary Note 7

Added:

(*continued*) This implies that while omnidirectional leakage may slightly increase due to environmental detuning, its amplitude remains within acceptable limits, often falling near or below the human hearing threshold or ambient noise levels. The spatial radiation characteristics were further analyzed via beam patterns at 0.5 m (Supplementary Figs. 13d and 13e), focusing on the frequency range surrounding the split modes. These results are normalized to the on-axis peak of the original 2nd-order mode. The patterns confirm that the meta-units effectively disrupt the radiation efficiency of the 2nd-order mode. Even under the worst-case stiffness variation (-20%), the radiation remains sufficiently suppressed in magnitude (≈ 20 dB for mode 2^{b}).

Furthermore, the beam pattern at 2 kHz (Fig. 4f) shows that while the meta-units effectively suppress omnidirectional leakage, a minor residual leakage persists due to imperfect suppression. This residual leakage could be further reduced by increasing the number of elastic meta-units or the mass of each unit, thereby enhancing the suppression effect[S13]. To improve robustness against environmental shifts, an asymmetric frequency distribution strategy could be adopted[S14]. By introducing slight geometric or stiffness variations across the array, the narrow suppression notch is transformed into a broader suppression band. This frequency-spreading approach ensures that a subset of meta-units remains active even if the plate's eigenmodes or material properties drift thermally. Transitioning to materials with lower thermal sensitivity and utilizing such heterogeneous arrays would provide consistent effectiveness in practical applications.

Comment 7. Figure 5a proposes "in-cabin personal audio delivery" as a key application. However, the results in Fig. 5 appear to be measured in an anechoic chamber. The authors should explicitly clarify the measurement environment in the caption and main text. Actually, practical cabin environments present more challenges such as reflections, scattering, and reverberation.

Response 7. We appreciate the reviewer for raising this point. We agree that the measurement environment is crucial for interpreting the experimental results. We would like to clarify that all measurements in Fig. 5 were conducted in an anechoic chamber for rigorous quantitative validation of the MiPAL's performance under controlled conditions—specifically to clearly verify the effects of the elastic meta-units on the resulting real-time audio performance without interference e.g., external boundary reflections or noise. As you suggested, we have updated the caption of Fig. 5 and the main text to explicitly state the measurement environment:

The caption of Fig. 5b

Original:	**Revised:**
Geometry for the speech demonstration at 0.5 m: microphones at 0°, 15°, 30°, 45°.	Geometry for the speech demonstration at 0.5 m conducted in an anechoic chamber: microphones at 0°, 15°, 30°, 45°.

In the subsection 'Potential applications of the MiPALs' (Lines 264–268)

Original:	**Revised:**
... To demonstrate this capability, we performed a speech-based experiment in the same measurement environment as Fig. 4a, recording at a distance of 0.5 m and at the angles (0°, 15°, 30°, 45°) defined in Fig. 5b.	... To demonstrate this capability, we performed a speech-based experiment in a controlled anechoic environment—consistent with the measurement setup in Fig. 4a—for a rigorous characterization of the intrinsic properties of the MiPAL, free from any external environmental factors. Recording was conducted at a distance of 0.5 m and at the angles (0°, 15°, 30°, 45°) defined in Fig. 5b.

Response 7. (*continued*) We also agree that practical cabin environments introduce additional complexities such as reflections, scattering, and reverberation, which can affect audio delivery performance. Regarding the cabin environment depicted in Fig. 5a, we would like to clarify that this is a conceptual illustration highlighting the MiPAL's primary advantage: **compactness**. Unlike array-based PALs (APALs) that typically require bulky arrays and complex electronic driving systems, the MiPAL offers a streamlined architecture suitable for integration into confined spaces like aircraft or train cabins. Beyond the conceptual illustration, we would like to further address the reviewer's concern by considering two primary scenarios in a cabin environment:

- **When passengers are seated (scattering)**
  The presence of passengers introduces scattering, primarily from the human head, which can be modeled simply as a rigid sphere[R6]. In this scenario, the SPL of PA sound in the region directly behind the head decreases significantly because the primary ultrasound waves are largely reflected by the sphere, as their wavelength is much smaller than the sphere's diameter[R7]. We further simulate the scattering of the PA sound field by a rigid sphere (diameter: 0.2 m; standard human head[R8]), as shown in Fig. R1. The rigid sphere is positioned 0.5 m away from the MiPAL along the propagation axis. The results indicate that PA sound reflections occur predominantly toward the front of the sphere. Regarding the lateral influence of this scattering (perpendicular to the propagation radiation), we consider a -20 dB difference as the threshold for a perceptually distinct change in sound pressure level. The lateral impact of the passenger's head on the audio field is confined to a radial range of approximately 0.4 m. For example, in economy class, the seat pitch is typically around 31–32 inches (approximately 0.8 m), meaning that within half this distance (0.4 m radius), the scattering effect from adjacent passengers' heads is relatively minimal. In business/first class, where seat pitches are generally larger, the application of MiPALs for private sound delivery is expected to be even more feasible.

- **When seats are empty (reflection, absorption)**
  In this scenario, the reflection and absorption are dominated by the structural materials of the seats[R9]. For example, experimental data in Ref. R10 show that various seat foams and fabrics exhibit absorption coefficients typically reaching 0.8 to 0.9 at high audio frequencies up to 6.3 kHz. Since the effectiveness of a porous absorber increases when the material thickness is large relative to the wavelength, the short wavelengths of ultrasonic carrier waves ($>$ 20 kHz) make the seat structure "acoustically thick". This ensures that incident ultrasonic energy can be effectively dissipated rather than being reflected back as reverberation. Consequently, reflection effects are expected to be relatively minimal in such an environment.

Response 7. (*continued*)

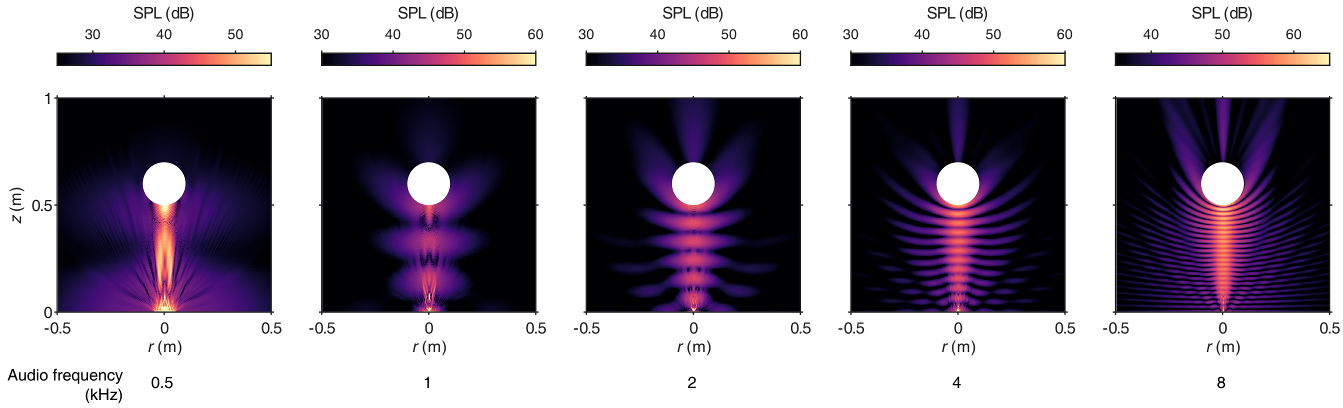


Fig. R1 | Scattering effects from a rigid sphere on resulting secondary audio sound fields.

While the physical acoustic effects in cabin environments have been simply addressed above, we acknowledge that additional factors in both the frequency domain (e.g., head-related transfer function, HRTF) and time domain (e.g., head-related impulse response, HRIR) remain largely unexplored in PAL research. As the reviewer noted, incorporating such elements is essential for a more comprehensive evaluation of real-world applicability. We plan to address these aspects through detailed analyses and experiments under various realistic acoustic conditions in our future work.

In the subsection 'Potential applications of the MiPALs' (Lines 260–263)

Original:	**Revised:**
Owing to its cost-effectiveness and scalability, the MiPAL architecture can be well suited for in-cabin entertainment systems where per-unit cost is critical (Fig. 5a).	Due to its compact form factor and scalability, the MiPAL architecture can be well-suited for deployment in confined environments like aircraft cabin (Fig. 5a). Unlike traditional APALs, its space-efficient design allows for seamless integration into seating areas where per-unit cost and mounting volume are critical constraints.

Comment 8. The manuscript acknowledges that the prototype produces "relatively limited SPL" compared to conventional loudspeakers. The authors are suggested to provide a quantitative analysis of the system's efficiency. How much acoustic energy is lost due to the insertion/blocking effect of the acoustic metasurface? Is the low SPL a fundamental limitation of the metasurface's low transmission ratio, or is it merely a power handling issue of the PZT driver? A comparative analysis of the radiation efficiency between the proposed MiPAL and the previous SPPAL is suggested to evaluate.

Response 8. We thank the reviewer for the suggestion to provide a quantitative analysis of the system's efficiency. We provide our response to move from the underlying physical principles of the components to a full system-level power flow anatomy, followed by a comparative analysis.

To elucidate the origins of the system's SPL, we first evaluate the physical behavior of the two primary stages:

1. **Insertion/blocking effect of the acoustic metasurface**

   In the acoustic metasurface, sound energy attenuation is dominated by the boundary layer effect arising from thermoviscous losses in the narrow subwavelength channels. Since the meta-atoms are typically designed with the frequency-dependent unit-cell periodicities, the resulting losses are inherently frequency-dependent, reflecting the narrowness of the channel geometries relative to the viscous/thermal boundary layer thicknesses. The effects of viscous and thermal boundary layers at the interface result in sound absorption. The thickness of viscous and thermal boundary layers is defined by $\delta_v = \sqrt{2\mu/\omega\rho_0}$, $\delta_t = \delta_v/\sqrt{\mathrm{Pr}}$ where μ is the dynamic viscosity, ω is the angular frequency, ρ_0 denotes the mass density of the medium, and Pr is Prandtl number[R11,R12]. At the carrier frequency of 60 kHz in air (with $\mu = 1.814 \times 10^{-5}\,\mathrm{Pa\cdot s}$ at 20°C), the viscous boundary layer thickness is approximately 8.9 µm. Given that the characteristic dimensions of the acoustic metasurface unit cells are on the order of millimeters, these boundary layers occupy a small fraction of the channel cross-sectional area. To provide a more intuitive understanding of the influence of the acoustic metasurface, we characterize the acoustic transparency of the constituent unit cells. The average transmittance (or power transmission coefficient) of the unit cells across the acoustic metasurface is approximately 0.55 (with a corresponding average amplitude transmission of around 0.74). This level of acoustic transparency indicates that the metasurface does not act as a significant physical barrier that limits the SPL of the generated audio sound.

2. **Conversion efficiency of PA sound in air**

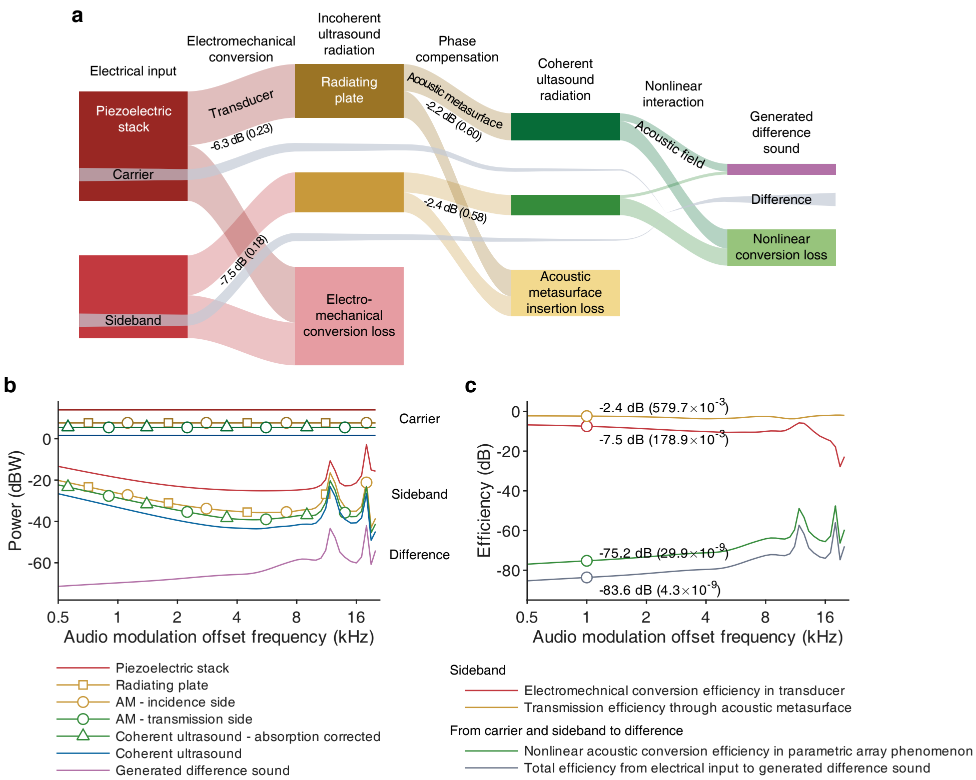
   It is well known that PA sound generation in air suffers from low conversion efficiency. This is primarily dictated by the nonlinear parameter (B/A)[R13], which originates from the Taylor expansion of the acoustic pressure in terms of density variations. This parameter characterizes the dominant finite-amplitude contribution to the sound speed, and its effect on acoustic wave propagation is manifest through the coefficient of nonlinearity $\beta = 1 + B/2A$.

   More specifically, in air, $\beta \approx 1.2$ (where $B/A \approx 0.4$), which is significantly lower than that of liquid media (e.g., $\beta \approx 3.5$ for water). Consequently, the overall conversion efficiency from ultrasound to audible sound is typically reported to be below 0.05 % in air under the idealized case[R14]. Therefore, the limited SPL resulting from this low efficiency is a fundamental physical limitation. This intrinsic constraint underscores the necessity for high-performance transducers as well as optimized integrated architectures. We anticipate that these challenges also will be addressed more comprehensively in our future studies.

Response 8. (*continued*) Building on these component-level principles, we provide a comprehensive power flow anatomy (detailed in Supplementary Note 15 and Supplementary Fig. 31). Briefly, the transmission efficiency through the acoustic metasurface is quantified as approximately 0.60 and 0.58 for the carrier and sideband frequencies, respectively; these results are consistent with the aforementioned component-level analysis. The nonlinear parametric conversion efficiency is calculated to be on the order of 10^{-8} at 1 kHz, which is lower than the 10^{-4} ratio reported for idealized cases at 10 kHz. This discrepancy is discussed in Supplementary Note 15.

In Supplementary Note 15

Added:



Supplementary Fig. 31 | **Power flow and efficiency analysis of the MiPAL. a,** Power flow and efficiency diagram illustrating the multi-stage process from initial electrical input to the final generated difference sound. The diagram provides a quantitative analysis at both the component level—from piezoelectric stack to acoustic field—and the process level. Two primary branches track the power streaming paths for the ultrasonic carrier and sideband components, which undergo nonlinear interaction in the acoustic field to generate the difference sound. Key efficiency benchmarks, including electromechanical conversion, metasurface insertion, and nonlinear parametric conversion, are indicated at their respective stages. Efficiencies for the transducer and acoustic metasurface are denoted in dB (and ratio) scale, specifically at the carrier and 1 kHz sideband frequencies. **b,** Power of the ultrasonic primary waves and generated difference sound. The spectral power of the carrier (darker tones) and sidebands (brighter tones) is compared with the resulting difference sound across the audio frequency range (0.5–20 kHz). The overlapping lines (the radiating plate and metasurface incident side coincide, and the metasurface transmission side and absorption-corrected coherent ultrasound coincide) provide a cross-verification of the power across different calculation methods. **c,** Efficiency of each conversion stage as a function of audio frequency. The electromechanical transduction efficiency and acoustic metasurface transmission efficiency are plotted for the sideband frequency only, as the carrier efficiency remains constant. Nonlinear conversion efficiency and total efficiency are calculated as the ratio of the difference-frequency power to the total input power (sum of carrier and sideband powers). The markerd points indicate the values at 1 kHz audio frequency.

Response 8. (*continued*)

In Supplementary Note 15

Added:

This note provides a comprehensive quantitative analysis of the power flow anatomy within the MiPAL system to elucidate the energy distribution across its various physical domains and verify the fundamental limits of its acoustic performance.

As described in Concept and overall principle of the MiPALs in Results, the generation of the target difference sound via parametric interaction requires the emission of two primary ultrasonic waves—the carrier and the sideband—from the transducer. Accordingly, we analyze the power flow anatomy for each ultrasonic frequency component separately (carrier and sideband branches) before evaluating their combined contribution to the final difference branch. As illustrated in Supplementary Fig. 31a, the MiPAL system comprises seven key processes categorized into power nodes and conversion stages. The power nodes—comprising the electrical input at the piezoelectric stack, incoherent radiation at the radiating plate, coherent radiation at the metasurface transmission interface, and the generated difference sound—represent the physical landmarks where active power is quantified. Conversely, the conversion stages—electromechanical transformation in the transducer, phase-compensated transmission through the metasurface, and nonlinear parametric interaction in the acoustic field—characterize the fundamental energy transitions between these nodes.

The total efficiency (η_{total}) from the initial electrical input to the final difference sound is approximately 4.3×10^{-9} (-83.6 dB at 1 kHz audio frequency), as calculated using Eq. (S15). The detailed process of power flow are as follows:

1. **Electrical input at piezoelectric stack**

   The power flow anatomy of the MiPAL system initiates with the electrical input at the piezoelectric stack. This nodal stage represents the primary power injection that drives the subsequent electromechanical transduction process within the transducer. To characterize the system's performance, the stack is driven with a peak amplitude of 50 V at both the carrier and sideband frequencies. This results in an electrical input power of approximately 24.7 W (13.9 dBW at f_1) and 0.013 W (-18.8 dBW at $f_2 = f_1 + 1$ kHz), respectively, as calculated using Eq. (S5).

2. **Electromechanical transduction in the transducer**

   This conversion stage characterizes the transformation of electrical power injected into the piezoelectric stack into mechanical vibrations of the radiating plate. The electromechanical conversion efficiency ($\eta_{\text{electromechanical}}$) at the carrier frequency (f_1) is approximately 0.23 (-6.3 dB), while the sideband frequency exhibits a slightly lower efficiency of 0.18 (-7.5 dB at $f_2 = f_1 + 1$ kHz), as determined by Eq. (S12).

3. **Incoherent radiation at the radiating plate**

   The flexural vibrations of the radiating plate generate an ultrasonic acoustic field. However, due to the high-order mode operation of the plate, the resulting acoustic radiation is inherently incoherent, as illustrated in Fig. 1f. The power radiated from the plate is quantified using Eqs. (S6) and (S7), yielding approximately 5.8 W (7.6 dBW at f_1) for the carrier and 0.0023 W (-26.3 dBW at $f_2 = f_1 + 1$ kHz).

4. **Phase compensation in the acoustic metasurface**

Response 8. (*continued*)

As illustrated in Fig. 1f, the acoustic metasurface plays a crucial role in transforming the incoherent ultrasonic field emitted by the radiating plate into a coherent wavefront, which involves an inevitable power loss as described in Fig. 2c. The metasurface transmission efficiency ($\eta_{\text{metasurface}}$) is approximately 0.60 (−2.2 dB at f_1) for the carrier and 0.58 (−2.4 dB at $f_2 = f_1 + 1$ kHz), as calculated using Eq. (S13).

5. **Coherent ultrasound radiation at the metasurface transmission interface**
   The phase-compensated ultrasonic waves exiting the metasurface form a coherent acoustic field. The power of the coherent ultrasound is quantified using Eqs. (S8) and (S9), yielding approximately 3.5 W (5.4 dBW at f_1) for the carrier and 0.0014 W (−28.6 dBW at $f_2 = f_1 + 1$ kHz).

6. **Nonlinear interaction in the acoustic field**
   The coherent ultrasonic waves at the carrier and sideband frequencies interact nonlinearly in the acoustic field to generate the desired directional difference sound. The nonlinear conversion efficiency ($\eta_{\text{nonlinear conversion}}$) is approximately 29.9×10^{-9} (−75.2 dB at 1 kHz audio frequency), as determined by Eq. (S14). This calculation based on the absorption-corrected coherent ultrasound power accounts for not atmospheric attenuation at ultrasonic frequencies but also parametric interaction losses. Omitting the atmospheric absorption correction[R14] leads to a slight overestimation of conversion efficiency. Beyond this, a significant source of the efficiency discrepancy between the MiPAL and idealized models[R14] stems from the disparity in reference audio frequencies (1 and 10 kHz) and the physical limitations of the transducer response. While idealized benchmarks assume identical, flat velocity amplitudes for both primary waves, practical transducers inevitably exhibit a non-flat frequency response. As described in Eq. (3), the generated difference-frequency pressure is proportional to both the product of the carrier and sideband pressures ($p_a \propto p_c \cdot p_s^*$) and the square of the audio frequency ($p_a \propto f_a^2$). Consequently, any spectral roll-off or deviation from resonance in either primary component, especially when combined with a lower target audio frequency, significantly compounds the efficiency loss relative to idealized models. Furthermore, the high-order modal operation of the MiPAL results in a complex surface velocity distribution (Fig. 1f), which potentially reduces the effective nonlinear interaction compared to the uniform, piston-like velocity profiles used in idealized numerical studies.

7. **Generated difference sound**
   The final nodal stage characterizes the secondary field resulting from the nonlinear interaction of the ultrasonic primary waves. The active power of the difference sound is quantified via Eq. (S11), yielding approximately 1.07×10^{-7} W (−69.7 dBW) at a 1 kHz audio frequency. This power is derived by integrating the acoustic intensity over a finite control surface $S_{\text{control surface}}$; consequently, the value represents the power contained within the defined computational volume. While this spatial truncation implies that the result is a conservative estimate of the total generated audio power, the domain is sufficiently large to capture the significant portion of the virtual source region. Notably, atmospheric absorption is omitted for this stage, as atmospheric attenuation at audio frequencies is negligible compared to the primary ultrasonic waves.

Response 8. *(continued)*

The power flow and efficiency are analyzed through Full FEM approach in Supplementary Fig. 31. The power at each stage is calculated using the following equations:

$$\Pi_{\text{piezoelectric stack}} = \frac{1}{2}\mathfrak{Re}\{\mathbf{V}_{\text{e}} \cdot \mathbf{I}_{\text{e}}^{*}\} \tag{S5}$$

$$\Pi_{\text{radiating plate}} = \int_{S_{\text{plate}}} \frac{1}{2}\mathfrak{Re}\{\mathbf{p}_{\text{a}} \cdot \mathbf{u}_{z,\text{m}}^{*}\}\, dS_{\text{plate}} \tag{S6}$$

$$\Pi_{\text{metasurface, incident}} = \int_{S_{\text{metasurface, incident}}} \frac{1}{2}\mathfrak{Re}\{\mathbf{p}_{\text{a}} \cdot \mathbf{u}_{z,\text{a}}^{*}\}\, dS_{\text{metasurface, incident}} \tag{S7}$$

$$\Pi_{\text{metasurface, transmission}} = \int_{S_{\text{metasurface, transmission}}} \frac{1}{2}\mathfrak{Re}\{\mathbf{p}_{\text{a}} \cdot \mathbf{u}_{z,\text{a}}^{*}\}\, dS_{\text{metasurface, transmission}} \tag{S8}$$

$$\Pi_{\text{absorption}-\text{corrected coherent ultrasound}} = \int_{S_{\text{control surface}}} \mathbf{I}_{\text{a}} \cdot e^{2\alpha R} \cdot \mathbf{n}\, dS_{\text{control surface}} \tag{S9}$$

$$\Pi_{\text{coherent ultrasound}} = \int_{S_{\text{control surface}}} \mathbf{I}_{\text{a}} \cdot \mathbf{n}\, dS_{\text{control surface}} \tag{S10}$$

$$\Pi_{\text{difference sound}} = \int_{S_{\text{control surface}}} \mathbf{I}_{\text{a}} \cdot \mathbf{n}\, dS_{\text{control surface}} \tag{S11}$$

where $\mathbf{V_e}$ and $\mathbf{I}_{\text{e}}$ are the complex voltage and current of the piezoelectric stack, respectively; $\mathbf{p}_{\text{a}}$ is the complex acoustic pressure; $\mathbf{u}_{z,\text{m}}$ and $\mathbf{u}_{z,\text{a}}$ are the complex velocities of the radiating plate and metasurface surface on the acoustical-axis, respectively; $\mathbf{I}_{\text{a}}$ is the acoustic intensity vector in the control surface; α is the absorption coefficient of air at the ultrasonic frequency; R is the distance from the transducer to the control surface; and $\mathbf{n}$ is the outward normal vector of the respective surface. The control surface $S_{\text{control surface}}$ is defined in a cylindrical coordinate system with a radius of $0.5\mathscr{D}_{\text{u,R}}$ and a height of $3\mathscr{D}_{\text{u,R}}$, centered on the transducer, where $\mathscr{D}_{\text{u,R}}$ is the Rayleigh distance of the ultrasonic carrier.

The efficiency at each stage is calculated as follows:

$$\eta_{\text{electromechanical}} = \frac{\Pi_{\text{radiating plate}}}{\Pi_{\text{piezoelectric stack}}} \tag{S12}$$

$$\eta_{\text{metasurface}} = \frac{\Pi_{\text{metasurface, transmission}}}{\Pi_{\text{metasurface, incident}}} \tag{S13}$$

$$\eta_{\text{nonlinear conversion}} = \frac{\Pi_{\text{difference sound}}}{\sum \Pi_{\text{absorption}-\text{corrected coherent ultrasound}}} \tag{S14}$$

$$\eta_{\text{total}} = \frac{\Pi_{\text{difference sound}}}{\sum \Pi_{\text{piezoelectric stack}}} \tag{S15}$$

where $\eta_{\text{electromechanical}}$ is the electromechanical conversion efficiency of the transducer, $\eta_{\text{metasurface}}$ is the acoustic metasurface transmission efficiency, $\eta_{\text{nonlinear conversion}}$ is the nonlinear parametric conversion efficiency, and η_{total} is the overall efficiency from electrical input to generated difference sound. The summation symbol in Eqs. (S14) and (S15) accounts for both the carrier and sideband ultrasonic powers contributing to the nonlinear interaction.

Response 8. (*continued*) To reflect and address the reviewer's suggestion, we have modified the main text to include a brief summary of the above analysis, as follows:

In the Discussion (Lines 309–314)

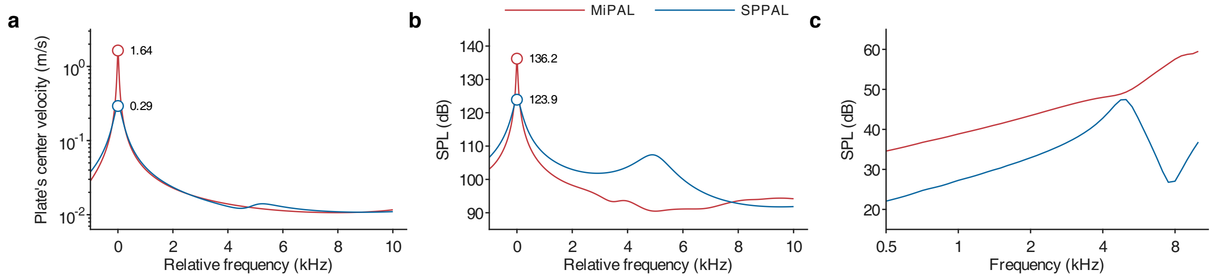
Original:	**Revised:**
... The current prototype still produces a relatively limited SPL compared to conventional loudspeakers, and therefore further optimization is required to achieve sufficient SPL for personal listening applications, where preferred levels typically reach approximately 70 dBA[R15].	... The current prototype still produces a relatively limited SPL compared to conventional loudspeakers, primarily constrained by the inherently low nonlinear conversion efficiency of parametric arrays; a quantitative power-flow analysis is provided in Supplementary Note 15. Further optimization is therefore required to achieve sufficient SPL for personal listening applications, where preferred levels typically reach approximately 70 dBA[R15].

To contextualize the engineering decisions that led to the MiPAL architecture, we have included a detailed in Concept rationale and narrative: the evolutionary transition from SPPAL to MiPAL in following. This note is intended for those interested in the specific evolutionary transition, highlighting the realization of structural-acoustic decoupling. While not strictly necessary for interpreting the primary findings, we invite the reviewer to consult this section for further insight into the system's radiation efficiency.

As shown in the comparative analysis (Supplementary Fig. 32), the MiPAL allows the radiating plate to operate at its undamped resonance, achieving a center velocity gain factor of 5.7 over the SPPAL. Even after accounting for the metasurface's transmission coefficient, the MiPAL achieves a net ultrasonic pressure gain of 4.2 (12.3 dB). This demonstrates that the MiPAL architecture successfully enhances the source power available for nonlinear conversion, proving that the metasurface is an enabler of high surface velocity rather than a hindrance to the system's power handling.

In Supplementary Note 16

Added:



Supplementary Fig. 32 | **Comparative performance analysis of MiPAL and SPPAL. a, b,** Simulated center velocity of the radiating plate and the resulting ultrasonic frequency response for MiPAL and SPPAL configurations using an identical transducer. The frequency axes are plotted as an offset relative to the respective resonant frequency of each configuration ($f - f_r$) to facilitate a direct comparison of bandwidth and peak magnitude. **c,** Comparison of the demodulated audible SPL as a function of audio frequency.

Response 8. *(continued)*

In Supplementary Note 16

Added:

To evaluate the relative radiation efficiency of our proposed architecture, we performed a comparative FEM analysis (through Full FEM approach in Supplementary Fig. 31) between the MiPAL and the previous SPPAL, utilizing an identical transducer and radiating plate geometry. Compared to the MiPAL, the SPPAL relies on integrated polymer steps that introduce significant material damping[R16–R18], which severely restricts the surface velocity of the radiating plate. In contrast, the MiPAL achieves structural-acoustic decoupling, allowing the plate to vibrate at its undamped maximum.
At the carrier resonance frequency, the MiPAL achieves a center surface velocity of 1.64 m/s, representing a gain factor of approximately 5.7 compared to the 0.29 m/s recorded for the SPPAL as shown in Supplementary Fig. 32a. When evaluating the ultrasonic radiation efficiency and accounting for the metasurface amplitude transmission coefficient, approximately 0.74, the theoretical ultrasonic pressure gain is approximately $5.7 \times 0.74 = 4.2$. This theoretical expectation is strongly corroborated by our simulation results at a distance of 0.45 m, where the MiPAL exhibits an ultrasonic SPL of 136.2 dB compared to 123.9 dB for the SPPAL (Supplementary Fig. 32b). The resulting 12.3 dB difference corresponds to a pressure ratio of 4.12, manifesting an excellent agreement with our predicted performance gain. This significant improvement in carrier radiation directly facilitates the enhanced audible SPL observed in the MiPAL system (Supplementary Fig. 32c), proving that the metasurface is an enabler of high surface velocity rather than a hindrance to the system's power handling.

To reflect and address the reviewer's suggestion, we have added the main text to bridge the readers interseted in the comparative performance analysis, as follows:

In the Introduction (Lines 39–42)

Added:

... Furthermore, the inherent material damping often associated with integrated step structures—particularly those utilizing polymers—leads to significant efficiency degradation, representing a bottleneck for achieving sufficient sound volume (see Supplementary Note 16 for detailed discussion).
...

Response 8. (*continued*)

Concept rationale and narrative: the evolutionary transition from SPPAL to MiPAL

The development of the MiPAL represents a strategic departure from the SPPAL framework, necessitated by the fundamental physical trade-offs between structural damping, radiator mass, and effective surface velocity.

1. **The initial phase**: monolithic stepped-plate architecture[21]

   Early iterations of the SPPAL utilized half-wavelength height annular steps attached directly to a radiating plate to compensate for phase disparities and generate a coherent ultrasonic beam[R19,R20]. Theoretically, reducing the plate thickness is advantageous, as a lower mass facilitates higher surface velocities for a given transducer excitation. However, when the plate and its integrated steps were fabricated from the same monolithic material (e.g., aluminum), significant engineering obstacles emerged:

   - **Modal interference**: In thin-plate configurations, the integrated steps act as structural discontinuities that disrupt the intended vibration modes, making it nearly impossible to maintain a predictable resonance[22].
   - **Thermal detuning**: Conversely, utilizing a thicker plate to preserve modal integrity increased the mechanical load on the PZT transducer. This not only limited the maximum achievable surface velocity but also induced significant heat generation within the PZT, leading to thermal degradation and undesirable shifts in the resonance frequency during operation[22].

2. **The intermediate solution**: integrated polymer steps[22–24]

   To mitigate the structural loading and modal interference of metal steps, an intermediate design was explored using polymer materials with lower Young's modulus and density. While the reduced stiffness of the polymer allowed for the use of thinner plates without compromising the vibrational mode, this approach introduced a new performance bottleneck:

   - **Intrinsic damping**: The high viscoelastic losses inherent in polymers resulted in severe parasitic damping.
   - **Velocity attenuation**: Experimental observations indicated that this damping caused the plate's surface velocity to decrease to approximately one-tenth (1/10) of its potential undamped amplitude, representing a significant loss in mechanical-to-acoustic conversion efficiency.

3. **The evolutionary transition**: structural-acoustic decoupling

   The MiPAL was conceptualized as a comprehensive solution to these stability and performance issues by fundamentally decoupling the phase-compensation mechanism from the vibrating structure. By shifting the phase control to an external acoustic metasurface, the system achieves several critical advancements:

   - **Maximum radiator intensity**: The radiating surface can now consist of a thin, undamped aluminum plate, allowing it to reach its maximum theoretical resonance velocity without the mass-loading of steps or the parasitic damping of polymers.
   - **Superior radiated velocity**: Although the metasurface itself introduces a minor transmission loss (amplitude transmission ≈ 0.74), this is negligible compared to the 0.1 velocity factor of the polymer SPPAL. Consequently, for an identical transducer-plate setup, the MiPAL configuration yields an effective radiated velocity approximately seven times higher than its predecessor.

Response 8. (*continued*)

- **Modular fabrication and operational robustness**: Beyond the gain in acoustic intensity, the MiPAL introduces a *modular design workflow* that enhances manufacturing process. In the SPPAL, the transducer, plate, and steps were structurally interdependent, requiring exhaustive iterative tuning to match the target resonance. In the MiPAL, the transducer-plate assembly is optimized and characterized as a standalone unit first. The acoustic metasurface is then independently designed based on the observed vibration profile and integrated as a separate layer.

Comment 9. The manuscript lacks specific dimensional details of the critical structures. Please add dimension annotations (e.g., plate radius) directly in Figs. 1 and 2.

Response 9. Thank you for your detailed assessment. We agree that providing explicit dimensional details is crucial for the reproducibility and technical clarity of the proposed MiPAL system. We have updated the figures as follows:

- Update to Fig. 1: We have added direct dimension annotations for the primary structural components, including the radiating plate radius, the acoustic metasurface radius, and the overall length of the elastic meta-units.
- Update to Fig. 2: To provide specific details of the acoustic and elastic metamaterial geometries without compromising the aesthetic clarity of the schematics, we have adopted a variable-labeling approach. We have added annotations for the critical geometric parameters as variables within the figure and provided the corresponding numerical values in the revised figure caption:

<table>
<tr><th colspan="2">The caption of Fig. 2a</th></tr>
<tr><td>Original:
Acoustic metasurface and its unit cells. Schematic of the metasurface colored according to the target phase distribution (left), consistent with (b) and (c). Unit-cell configuration comprising a central waveguide and Helmholtz resonators connected through a space-coiling channel (right), where the geometric parameters (t, w, $t_0 - t_2$, $w_1 - w_3$) define the acoustic response.</td><td>Revised:
Acoustic metasurface and its unit cells. Schematic of the metasurface colored according to the target phase distribution (left), consistent with (b) and (c). Unit-cell configuration comprising a central waveguide and Helmholtz resonators connected through a space-coiling channel (right), where the geometric parameters ($t = 2.43$ mm, $w = 8.58$ mm, $t_0 = 2.86$ mm, $w_1 = 0.3$ mm, $w_2 = 0.9$ mm, t_1, t_2) define the acoustic response.</td></tr>
</table>

<table>
<tr><th colspan="2">The caption of Fig. 2e</th></tr>
<tr><td>Original:
High-order flexural-mode plate and its elastic meta-units. Comparison of the radiating plate without and with attached elastic meta-units (left). Configuration of an elastic meta-unit (right), defined by the geometrical parameters (w_f, h_f, t_s, l_s, w_m, d_m, h_m).</td><td>Revised:
High-order flexural-mode plate and its elastic meta-units. Comparison of the radiating plate without and with attached elastic meta-units (left). Configuration of an elastic meta-unit (right), defined by the geometrical parameters ($w_f = 3$ mm, $h_f = 1$ mm, $d_m = 3$ mm, t_s, l_s, w_m, h_m, h_m). The foot and spring segments are integrated with a 1 mm longitudinal overlap.</td></tr>
</table>

We believe these additions provide the necessary technical transparency while maintaining the visual quality required for the manuscript.

Comment 10. In Fig. 4f, what are circles and triangles for?

Response 10. We apologize for the omission of these labels in the original figure. We have updated Fig. 4f to clearly define these markers and have implemented the following changes:

- Legend addition: A clear legend has been added to Fig. 4f to identify the data sets.
- Marker definitions: In line with the convention established throughout the manuscript, the circular markers (○) represent the audio beam pattern of the MiPAL with elastic meta-units, while the triangular markers (△) represent the baseline configuration without elastic meta-units.
- Cross-domain synchronization: As detailed in our response to Comment 5 by Reviewer 3, these marker styles have been synchronized with the structural vibration data in Fig. 2g. This consistency ensures that the reader can intuitively correlate the suppression of structural resonances with the resulting improvement in audio directivity.

Reviewer 2

Comment 1. The manuscript proposes a metasurface-integrated parametric array loudspeaker (MiPAL), which achieves ultra-broadband directional sound emission (spanning 500 Hz to 10 kHz) by combining a compact single-body ultrasonic transducer with an acoustic metasurface. The study is content-rich, featuring rigorous theoretical analysis, simulations, and experimental validation. However, key aspects including novelty, physical insight, and practical significance are insufficiently addressed, and the manuscript thus does not meet the publication standards of Nature Communications.

Response 1. We sincerely thank the reviewer for careful evaluation of our manuscript and for raising concerns regarding the novelty, insight, and practical significance of our work. In this revision, we have thoroughly addressed all your comments point-by-point and significantly revised and improved the manuscript resulting in increased technical rigor. In particular, we have substantially expanded the discussion on the novelty and physical mechanisms underlying the proposed metamaterials-integrated parametric array loudspeaker (MiPAL). We clarify how interweaving dual-domain metamaterials with the transducer's architecture enables wave manipulation capabilities—such as the simultaneous suppression of parasitic resonances and precise wavefront shaping—that are not accessible in existing directional sound technologies.

Moreover, we have added additional analyses, simulations, experimental results, and discussions to further strengthen the quality of the manuscript. Collectively, these revisions substantially improve the technical rigor, clarity, and overall impact of the work. We sincerely hope that, with these comprehensive revisions and clarifications, the referee will recognize the distinct novelty and broader significance of our study and will join us in envisioning MiPAL as a promising and extensible practical platform for metamaterial-integrated advanced wave engineering.

Comment 2. The acoustic metasurface of the MiPAL adopts a unit-cell structure consisting of "radially arranged Helmholtz resonators (HRs) interconnected by space-coiling (SC) channels". Such structures have been utilized in numerous existing studies (e.g., Appl. Phys. Lett. 122, 183506 (2023))[R21]. The authors fail to clearly illustrate the innovative differences. Furthermore, the manuscript does not adequately elaborate on strategies to address the dispersion effects of the metamaterial structure over a broad frequency range.

Response 2. Thank you for your comments and the suggested reference. As you correctly noted, various geometrical configurations, such as HR and SC channels, have been extensively studied as **building blocks** for acoustic metasurface design over the past decade. We wish to clarify that we do not claim the geometry of the HR+SC hybrid unit cell configuration itself to be the primary innovation of this work. This configuration to be just one of several viable geometry for metasurface configuration, alongside other established solutions such as pure HR arrays or space-coiling channels[R22–R24]. Instead, the conceptual novelty of the MiPAL lies in the **interweaving of acoustic and elastic metamaterial structures into a monolithic, multi-physics transduction platform**. In order to avoid any misunderstanding regarding our original description, the manuscript has been revised as follows:

In the subsection 'Design, characterization, and fabrication of the MiPALs' (Lines 128–133)

Original:	**Revised:**
... Based on the preliminary established design process, we proceed to design the acoustic metasurface with a 60 kHz carrier frequency. The key design component of the MiPAL is the acoustic metasurface (Fig. 2a). The metasurface comprises radially arranged unit-cells featuring Helmholtz resonators (HRs) interconnected by a space-coiling (SC) channel.	... Based on the preliminary established design process, we proceed to design the acoustic metasurface with a 60 kHz carrier frequency, a 100 mm aperture diameter, and a 0.5 m target hearing distance. The radiating plate of the transducer is identical in size, with a diameter of 100 mm, aligning with the metasurface aperture. For its implementation, we adopt and tailor a hybrid unit-cell geometry[40,41] comprising radially arranged Helmholtz resonators (HRs) interconnected by a space-coiling (SC) channel (Fig. 2a).

Response 2. (*continued*) Regarding the **strategies to address dispersion effects over a broad frequency range**, we agree that the original manuscript did not sufficiently detail the underlying physical mechanisms and our design strategies. In the original manuscript, we briefly mentioned that the broadband characteristics in Supplementary Note 6. To provide a more comprehensive understanding, we have expanded this in the revised Supplementary Information and added discussions to the main text. Indeed, ultra-broadband performance of the MiPAL in the audio frequency range is achieved through a two-fold interplay:

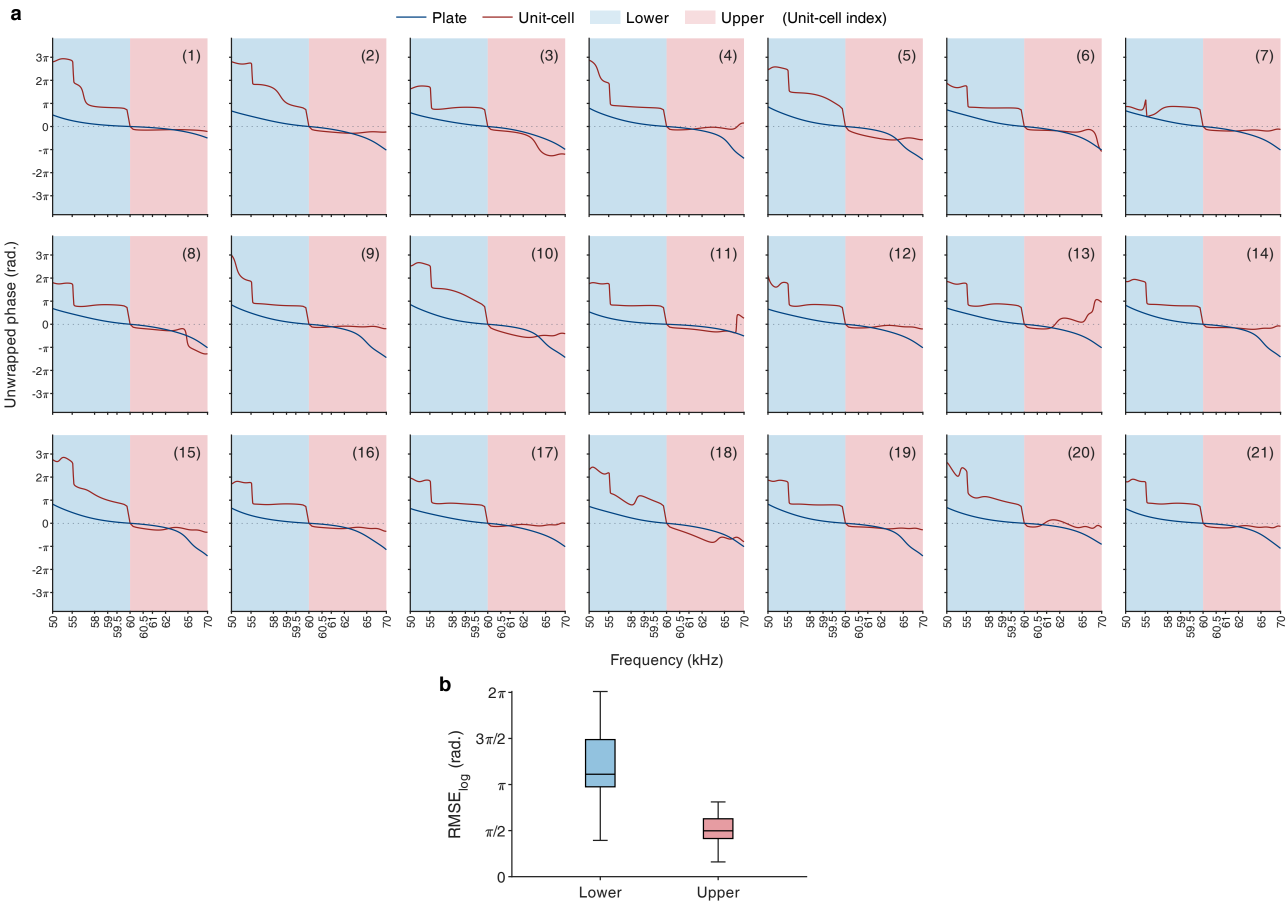
1. **Bandwidth translation from ultrasound to audio enabled by nonlinear interaction**
   As described in the manuscript, our system operates within a relatively narrow bandwidth in the ultrasonic range (60–70 kHz), which corresponds to a modest fractional bandwidth of ∼15% (around 0.2 octaves). However, through the nonlinear process, this range is effectively translated into a broadband response in the audio range (0.5–10 kHz) spanning over **four octaves**. This effectively bypasses the inherent narrowband limitations typically encountered in resonance-based acoustic metamaterials operating in the linear regime.

2. **Mitigation of dispersive effects via adopting upper-sideband modulation**
   Beyond the simple bandwidth translation, the MiPAL can reduce beam distortion by synchronizing the transducer's structural dispersion with the acoustic metasurface's dispersion characteristics. As illustrated in Supplementary Fig. 10a, we found that the dispersive slope of the transducer in the upper sideband (60–70 kHz) is more suitable for directional sound beam generation than that of the lower sideband (50–60 kHz). More specifically, in the lower sideband, another resonant peak (around 55 kHz) in the transducer emerges, leading to abrupt phase shifts and significant phase discrepancies relative to the metasurface's profile.
   To quantitatively validate this, we analyze the averaged (log-weighted) RMSE by Eq. (R1) between the radiated phase profile of the plate and that of the unit-cells mounted at each plate's radial position, as shown in Supplementary Fig. 10b. The results confirm that the upper sideband operation consistently yields smaller phase errors compared to the lower sideband. This mismatch directly translates to distorted wavefronts.
   $$\mathrm{RMSE}_{\log} = \sqrt{\frac{1}{\ln(f_{\max}/f_{\min})}\int_{f_{\min}}^{f_{\max}} \left|\phi_{\mathrm{unit-cell}}(f) - \phi_{\mathrm{plate}}(f)\right|^2 \frac{1}{f}\, df} \tag{R1}$$
   where $f_{\min} = 500$ Hz and $f_{\max} = 10$ kHz.
   Consequently, by strategically selecting the upper sideband modulation, we achieved broadband characteristics for both primary ultrasonic and secondary audio range (Figs. 3e and 4d to 4f and Supplementary Fig. 19). While the upper sideband does not possess a perfectly matched phase slope—which may still introduce minor wavefront distortion—the measured directivity patterns clearly demonstrate that the MiPAL maintains robust performance across the broad audio frequency range. We believe that further refining this dispersion matching represents a promising avenue for future performance enhancement; however, for the scope of this work, we have adopted this straightforward yet reasonable strategy.

Response 2. (*continued*)

In Supplementary Note 6

Added:



Supplementary Fig. 10 | **Dispersion anaylsis of the designed MiPAL. a,** Frequency-dependent unwrapped phase profiles of individual unit-cells and the plate-induced acoustic radiation. The phase of the metasurface unit-cells (red lines) are compared with the radiation phase from the transducer plate (blue) at 21 representative positions. All phase values are normalized to zero at the center frequency of 60 kHz. The frequency axis is depicted on a symmetrical logarithmic scale centered at 60 kHz. **b,** Quantitative comparison of phase matching performance. The averaged log-weighted root mean square error (RMSE) of the phase difference between the plate and the unit-cells is calculated for the lower band (50–60 kHz) and the upper band (60–70 kHz). The low RMSE in the upper band quantitatively confirms the selection of this region for robust, broadband directional audio generation with minimal phase-induced distortion.

In Supplementary Note 6

Added:

Design rationale for adopting upper-sideband modulation

The operational principle of the MiPAL system relies on a two-step mechanism that overcomes the fundamental limitations of conventional acoustic metamaterials. First, the system leverages a bandwidth translation process facilitated by nonlinear acoustic interaction. Although the physical device operates within a relatively narrow ultrasonic bandwidth (60–70 kHz), representing a fractional bandwidth of approximately 15% (~ 0.2 octaves), this range is effectively down-converted into a

Response 2. (*continued*)

wide audio spectrum (0.5–10 kHz) spanning over four octaves. This nonlinear translation allows the MiPAL to bypass the narrowband constraints inherent in linear, resonance-based metamaterials, achieving broadband audio performance while maintaining the design advantages of high-frequency operation.

Beyond this bandwidth translation, the MiPAL can reduce beam distortion by synchronizing the transducer's structural dispersion with the acoustic metasurface's dispersion. As demonstrated in Supplementary Fig. 10a, we analyze the phase difference characteristics across both the lower (50–60 kHz) and upper (60–70 kHz) sidebands. One can see that the dispersive slope of the transducer in the upper sideband (60–70 kHz) is more suitable for directional sound beam generation than that of the lower sideband (50–60 kHz). More specifically, in the lower sideband, another resonant peak (around 55 kHz) in the transducer emerges, leading to abrupt phase shifts and significant phase discrepancies relative to the metasurface's profile.

We further calculate the averaged (log-weighted) RMSE by Eq. (S4) between the phase profile radiated by the plate and the response of the unit-cells across all radial positions Supplementary Fig. 10b. The analysis confirms that the upper sideband consistently yields lower phase errors compared to the lower sideband. By strategically operating in this regime where the dispersive effects are synchronized and minimized, the MiPAL achieves robust, broadband characteristics for both the primary ultrasonic waves and the secondary audio signals.

$$\mathrm{RMSE}_{\log} = \sqrt{\frac{1}{\ln(f_{\max}/f_{\min})} \int_{f_{\min}}^{f_{\max}} \left|\phi_{\text{unit-cell}}(f) - \phi_{\text{plate}}(f)\right|^2 \frac{1}{f}\, df} \tag{S4}$$

where $f_{\min} = 500$ Hz and $f_{\max} = 10$ kHz.

Response 2. (*continued*) Furthermore, as the reviewer correctly pointed out, broadband chromatic (frequency-dependent) dispersion can affect the overall performance and cause beam distortion if not properly managed. While it is possible to apply various dispersion compensation techniques that have been extensively studied in the fields of optics and acoustics[R25–R27], we found that a straightforward yet effective approach is to compare the relative dispersive effects between the transducer and the metasurface, and specifically selecting the optimal operation band. As such, we adopted a metasurface design methodology that utilizes a phase modulation scheme based on the carrier frequency, which proved to be robust within our directivity characteristics. We anticipate that dispersion engineering techniques could be applied to further enhance the performance—for instance, in designing focused PALs to maximize local nonlinear effects to achieve achromatic focusing response. However, this requires a distinct approach from conventional dispersion-engineered metasurfaces, as the dispersive characteristics of the transducer must be simultaneously accounted for within the system design. This would be an interesting direction for future research, and we have clarified this point in the revised manuscript. We once again thank the referee for the valuable feedback, and we list the relevant revisions below:

In the subsection 'Design, characterization, and fabrication of the MiPALs' (Lines 138–141)

Original:	**Revised:**
... To achieve desired wavefront shaping, each unit-cell is individually encoded to counteract the incoherent radiation modes produced by the radiating plate (Fig. 2b).	... To achieve desired wavefront shaping, each unit-cell is individually encoded to compensate for the incoherent radiation modes produced by the radiating plate (Fig. 2b), ensuring accurate phase modulation and high transmission at the center frequency (see Supplementary Note 5 for details).

In the subsection 'Experimental characterization of the MiPAL' (Lines 221–229)

Added:

... Nevertheless, we also note that owing to the inherent dispersive nature of both the acoustic metasurface and the transducer, chromatic (frequency-dependent) aberrations may arise, potentially distorting the wavefront across the operational bandwidth. These effects are discussed in detail in Supplementary Note 6, and can be mitigated using sophiscated dispersion engineering techniques[46–48]. However, we emphasize that in integrated systems like the MiPAL, it is essential to simultaneously account for not only the metasurface's intrinsic dispersion but also the transducer's structural dispersion—a complex interplay that represents a significant leap from conventional metasurface design and will be a key subject of our future study.

Comment 3. The design concept of "integrating a single-body transducer with metamaterial structures" has been proposed in prior research (e.g., Appl. Phys. Lett. 127, 151701 (2025))[R28], and no new breakthroughs in this integration scheme are demonstrated in the current work. Additionally, the manuscript does not compare performance under different transducer geometries scenarios, lacking further validation of the scheme's applicability and flexibility.

Response 3. We thank the referee for raising this point. However, we respectfully disagree with your assertion that **no new breakthroughs** are demonstrated in our work. We believe this assessment may stem from a narrow interpretation of "integration" as a simple combination of existing components.

Foremost, we wish to clarify that our claim of novelty does not rest on the mere integration of a transducer with metamaterial structures *per se*. Instead, we would like to emphasize that the **functionally interweaving** of dual-domain metamaterials (both acoustic and elastic) within a compact and modular ultrasonic transducer platform, thereby successfully demonstrating its potential as a compelling candidate for next-generation PALs. Also, our approach enables unprecedented, simultaneous control over both acoustic and elastic waves across a wide frequency spectrum from audible to ultrasonic. We believe that our work is **categorically distinct** from prior studies that primarily focus on metamaterial structures as dependent elements. A critical distinction lies in the intrinsic coupling of the transducer's internal mechanisms within the comprehensive design framework:

- **Prior works**: Existing studies include the Ref. R28 and Refs. 3, 34 (as pointed out also in Comment 4 by Reviewer 2) treat the transducer as a "given" or "black-box" source, where metamaterials are added as external components to manipulate the radiated acoustic field. As a result, these approaches are inherently limited because they treat the transducer as an immutable source rather than a designable variable. This prevents them from addressing the fundamental constraints imposed by the transducer's internal characteristics—such as electromechanical coupling, structural vibrations, and design flexibility.

- **This work**: Departing from simple integration approach, the MiPAL represents a new design paradigm that co-designs the electromechanical transduction, structural vibrations and elastic wave, and acoustic radiation within a single monolithic system. To the best of our knowledge, this work provides the first demonstration of a framework where the transducer's internal mechanics and dual-domain metamaterials are so intricately linked that their interplay resolves the bottlenecks of previous approaches—a feat unattainable through a simple combination of discrete components.

We believe that our work provides a robust engineering solution for complex wave manipulation in dual-domain environments: a challenge that has remained largely unaddressed in this field. Our work provides profound insights that resonate with the multidisciplinary readership of *Nature Communications*, spanning from mechanical engineering to wave physics. We hope that the reviewer appreciates the significance of our contributions in this context. Reflecting your valuable feedback and concerns, we have strengthened the narrative of our manuscript to better highlight the contributions of our work as follows:

Response 3. (*continued*) Firstly, to better represent the dual-domain integration strategy and the broadband impact of our research, we first revised the title to more accurately reflect our core contribution:

Title

Original:	**Revised:**
Metasurface-integrated compact ultrasonic transducer for highly directional audio generation	Dual-domain metamaterials interwoven with a compact ultrasonic transducer for highly directional audio generation

Next, the abstract has been revisited to better articulate the novelty and significance of our approach:

Abstract

Original:	**Revised:**
... Here, we report an alternative form factor: a compact, single-body ultrasonic transducer augmented with an acoustic metasurface, termed metasurface-integrated parametric array loudspeakers (MiPALs). Unlike conventional PALs, the MiPAL utilizes a single piezoelectric driver combined with a co-designed metasurface, yielding a simple and cost-effective architecture that achieves compact directional sound systems. To realize this, we exploit a unified design strategy that integrates theoretical modeling, transducer-metasurface co-design, and comprehensive quantitative analyses of both linear and non-linear acoustic fields. ... These findings underscore the potential of metasurface-transducer integrated platforms for advanced audio systems, laying the foundation for developments in immersive audio, spatial sound delivery, and extended reality.	... Here, we report an alternative form factor: a compact, single-body ultrasonic transducer integrated with dual-domain metamaterials, termed metamaterials-integrated parametric array loudspeakers (MiPALs). Unlike conventional PALs, the MiPAL utilizes a single piezoelectric driver interwoven with dual-domain metamaterials—consisting of a co-designed acoustic metasurface and elastic meta-units—to yield a simple, cost-effective, and highly integrated architecture. To realize this, we exploit a unified design strategy that integrates theoretical modeling, transducer-metamaterials co-design, and comprehensive quantitative analyses of both linear and non-linear acoustic fields. ... These findings underscore the potential of metamaterial-transducer integrated platforms for advanced audio systems, laying the foundation for developments in immersive audio, spatial sound delivery, and extended reality.

We also have revised the Introduction to explicitly distinguish our approach from prior works:

In the Introduction (Lines 26–28 and Lines 39–46)

Added:

... This reliance fundamentally limits design flexibility, as the transducer is conventionally treated as a given black-box source with predefined operating characteristics, rather than as a designable variable.

... Furthermore, the inherent material damping often associated with integrated step structures—particularly those utilizing polymers—leads to significant efficiency degradation, representing a

Response 3. (*continued*)

bottleneck for achieving sufficient sound volume (see Supplementary Note 16 for detailed discussion). Crucially, earlier SPPALs inevitably suffer from the compromise of their core functionality—directional sound reproduction—due to combination resonance (CR)[24]. This parasitic phenomenon, an inherent side-effect of nonlinear intermodulation, generates nearly omnidirectional acoustic leakage that represents a primary drawback of existing single-body PAL architectures.

In the Introduction (Lines 47–58)

Original:	**Revised:**
... In this paper, we propose a new form factor for PALs—the metasurface-integrated PAL (MiPAL)—which combines a modular, single-body piezoelectric ultrasonic transducer with a co-designed acoustic metasurface. The integrated metasurface features subwavelength spatial resolution and high design flexibility, both considered well-known advantages of metamaterials[25–32,41], and effectively redirects the non-directional ultrasound generated by the transducer's flexural vibration into highly collimated and directional wavefront. As a result, directional audible sound is produced through DFW generation via nonlinear interactions of high-intensity ultrasonic waves. Furthermore, embedding locally resonant elastic meta-units on the radiating plate effectively suppresses the omnidirectional structure-borne sound associated with the combination-resonance (CR) mode—an undesired by-product nonlinearly induced through intermodulation[24] and one of the key drawbacks of SPPALs.	... In this paper, we propose a new form factor for PALs—the metamaterials-integrated PAL (MiPAL)—which consists of a modular, single-body piezoelectric ultrasonic transducer interwoven with dual-domain metamaterial architectures. Within this unified system, the acoustic metasurface serves as the primary functional layer; by leveraging subwavelength spatial resolution and high design flexibility, both considered well-known advantages of metamaterials[25–32], it effectively redirects the non-directional ultrasound generated by the transducer's flexural vibration into highly collimated and directional wavefront. As a result, directional audible sound is produced through DFW generation via nonlinear interactions of high-intensity ultrasonic waves. Simultaneously, the elastic meta-units act as a critical compensatory layer. By embedding deep-subwavelength tiny locally resonant structures[33] onto the transducer's radiating plate, we effectively suppress the omnidirectional sound leakage associated with the CR mode that undermines the core functionality of the PAL by compromising its defining high-directivity.

Response 3. (*continued*)

In the Introduction (Lines 61–64)

Original:	Revised:
... This framework comprehensively encompasses the design of transducer components, geometries of metamaterial structures, and rigorous full-field analysis of both linear and nonlinear acoustic behavior.	... This framework encompasses the comprehensive design of the MiPALs, integrating electro-mechanical transduction with dual-domain wave manipulation through theoretical investigation and rigorous full-field analyses of both linear and nonlinear behaviors in acoustic and elastic regimes.

We have revised the Discussion section to better highlight the contributions of our work:

In the Discussion (Lines 297–308)

Original:	Revised:
... In this work, we introduce the MiPAL, which fundamentally redefines the architecture for personal audio systems. By integrating an acoustic metasurface with a compact single-body ultrasonic transducer, the MiPAL achieves broadband directional sound generation without relying on costly multi-emitter arrays constructed from acoustically oversized off-the-shelf ultrasonic emitters or the complex phase-control electronics[1,19]. The metasurface redirects the incoherent flexural-mode radiation into collimated wavefront, while embedded elastic meta-units effectively suppress parasitic vibration modes that otherwise cause undesired acoustic leakage. As a result, the MiPAL demonstrates directional audio across an ultra-broad frequency range from 500 Hz to 10 kHz—covering over four octaves. This compact platform overcomes the cost, complexity, and scalability limitations of conventional PALs.	... In this work, we introduce the MiPAL, which fundamentally redefines the architecture for advanced personal audio systems through the dual-domain integration of metamaterials. By synergistically combining an acoustic metasurface and elastic meta-units with a compact single-body ultrasonic transducer, the MiPAL achieves broadband directional sound generation without relying on costly multi-emitter arrays or the complex phase-control electronics typically required in conventional systems[R1,R19]. The metasurface redirects the incoherent flexural-mode radiation into collimated wavefronts, while embedded elastic meta-units effectively suppress parasitic vibration modes that otherwise cause undesired omnidirectional leakage. As a result, the MiPAL demonstrates directional audio performance across an ultra-broad frequency range from 500 Hz to 10 kHz—covering over four octaves within a physically compact platform. Our approach overcomes the cost, complexity, and scalability limitations of conventional PALs, paving the way for widespread adoption in diverse applications.

In the Discussion (Lines 319–323)

Added:

... Furthermore, we foresee that the interweaving of architected metamaterials with real-world

Response 3. *(continued)*

functional platforms will catalyze a paradigm shift across diverse fields, including biomedicine, non-destructive testing, and energy harvesting, by enabling unprecedented control over wave-matter interactions.

Response 3. (*continued*) Furthermore, regarding the reviewer's concern on the **lack of further validation of the scheme's applicability and flexibility**, we acknowledge that our original submission did not include a sufficiently broad comparison of various transducer configurations. To address this, we have conducted additional simulations across a range of radiator dimensions to verify our design methodology, which is central to the MiPAL transduction platform.

Specifically, we established several "design splits" by varying the ultrasound critical distance ($\mathscr{D}_{\mathrm{u,c}}$) from 0.4 to 0.5 meters (at 0.05-meter intervals) and the plate's flexural vibration mode orders (m; 7th, 8th, and 9th). For reference, the primary design parameters presented in the main manuscript are $\mathscr{D}_{\mathrm{u,c}} = 0.45$ m and the 8th-order plate mode. As shown in Fig. R2, the behaviors of the transducer, plate, acoustic metasurface, and resulting acoustic fields are accurately predicted based on our design methodology across all configurations.

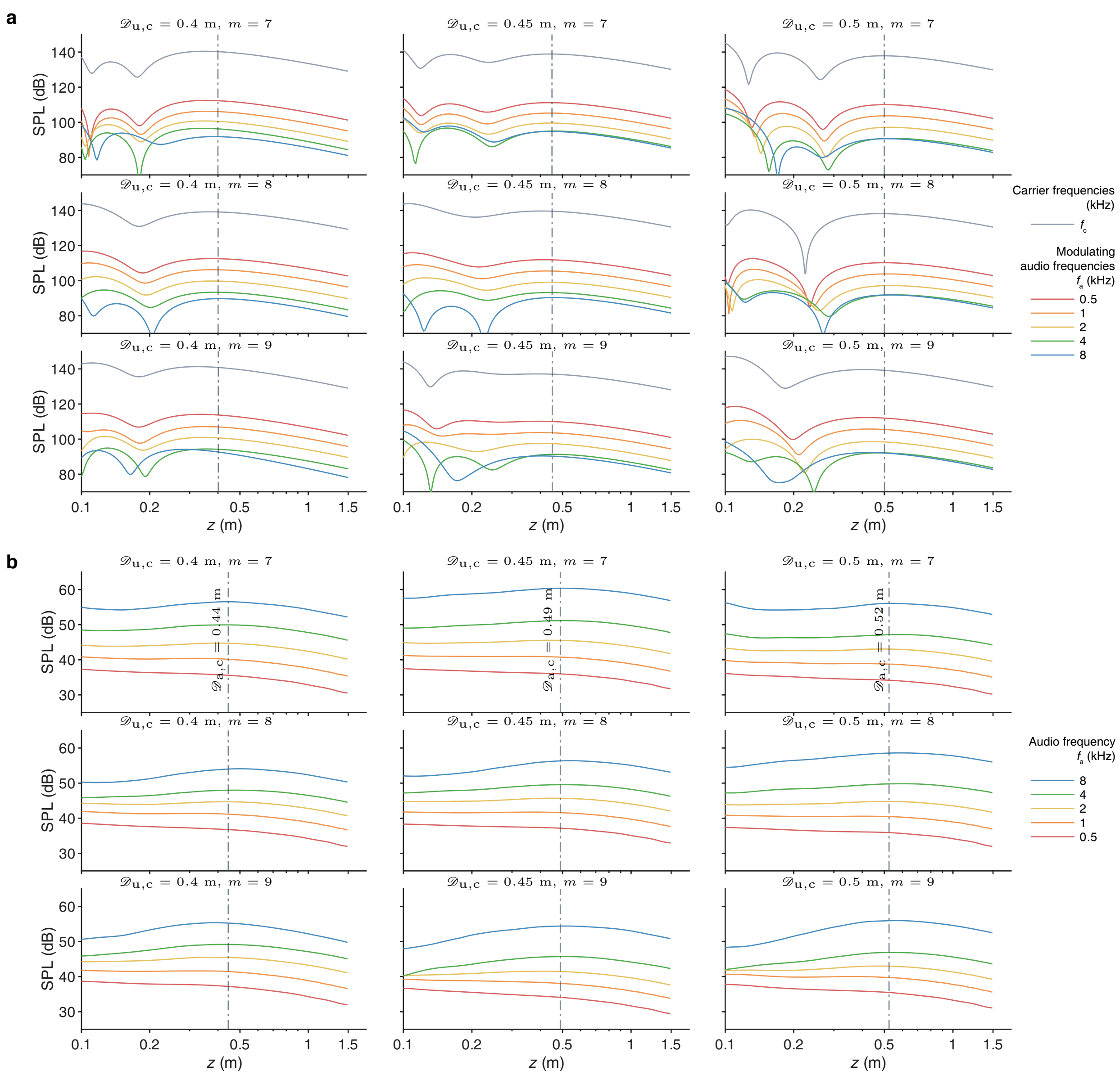


Fig. R2 | Parametric validation of the MiPAL design framework across diverse radiator dimensions. **a,** Ultrasound propagation curves for varying design parameters (ultrasound critical distance $\mathscr{D}_{\mathrm{u,c}}$ and plate flexural mode order). Vertical dash-dotted lines indicate the specific critical distance for each design split. **b,** Audio sound propagation curves, where vertical dash-dotted lines represent the corresponding critical distance ($\mathscr{D}_{\mathrm{a,c}}$) for the demodulated audio signal.

Response 3. (*continued*) Notably, the ultrasound propagation curves in Fig. R2a and the critical distance for the demodulated audio sound in Fig. R2b are both well-matched with the theoretical rigid piston model described in Supplementary Note 3. The ultrasound beam patterns illustrated in Fig. R2e align with our characterization of the acoustic metasurface and plate behavior, and all design variants demonstrate robust, high-fidelity audio beam patterns as shown in Fig. R2f. These results definitively confirm that the MiPAL design scheme is a versatile framework adaptable to diverse engineering requirements and radiator scales. We believe these additions substantiate the versatility and robustness of our proposed MiPAL architecture across varying transducer geometries and operating conditions.

As a concluding remark, we sincerely invite the Reviewer to evaluate the impact of our work through the interwoven dual-domain architecture and the practical demonstrations, which we believe hold significant potential for a wide range of engineering and acoustic applications.

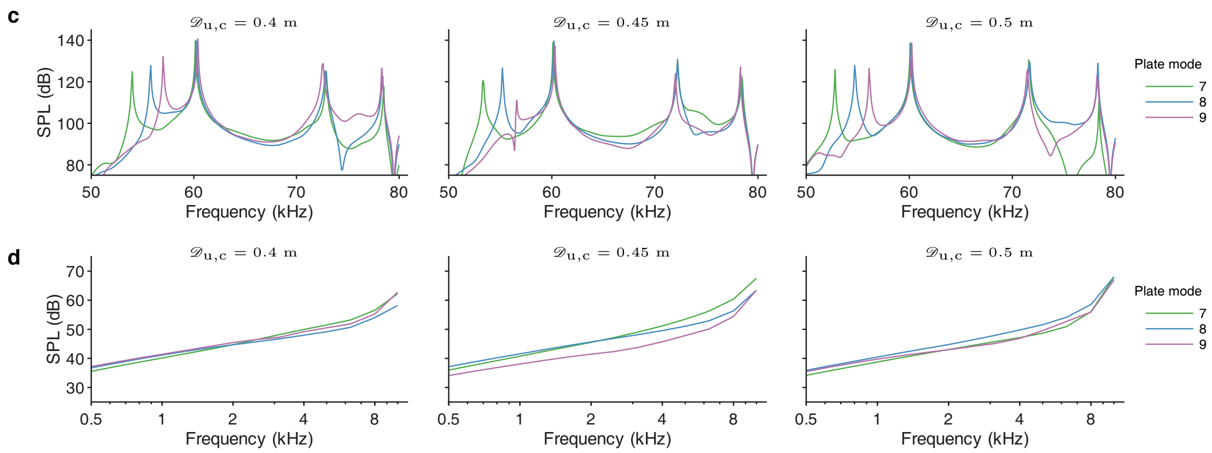
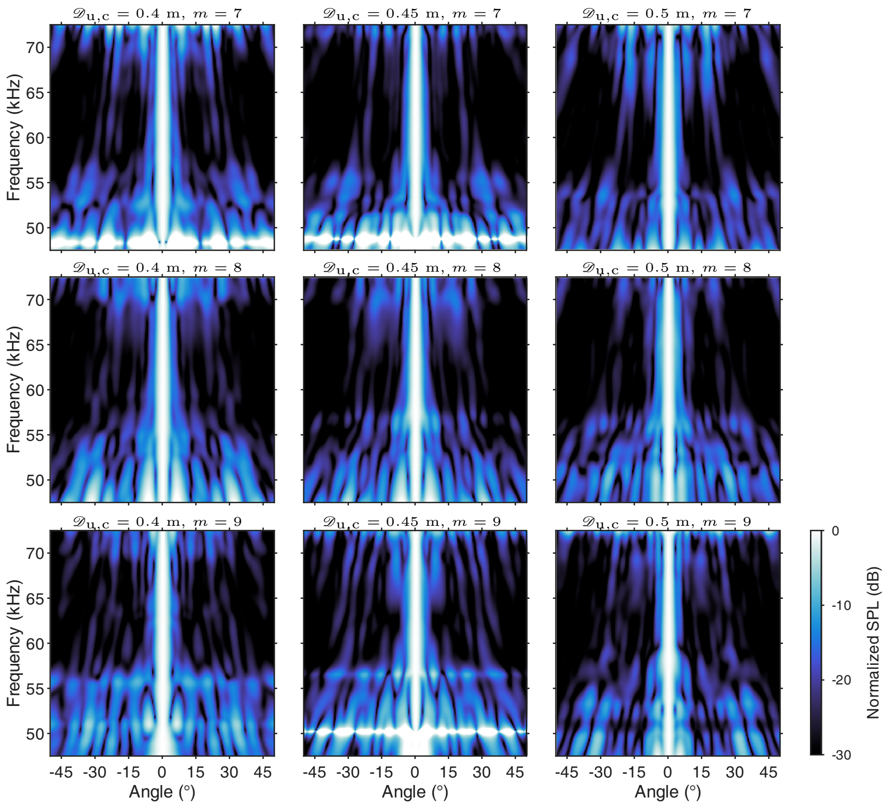


Fig. R2 | (*continued*) **c,** Ultrasound frequency responses calculated at the ultrasound critical distance, $\mathscr{D}_{\mathrm{u,c}}$. **d,** Audio sound frequency responses calculated at the audio critical distance, $\mathscr{D}_{\mathrm{a,c}}$.

Response 3. (*continued*)

e



f

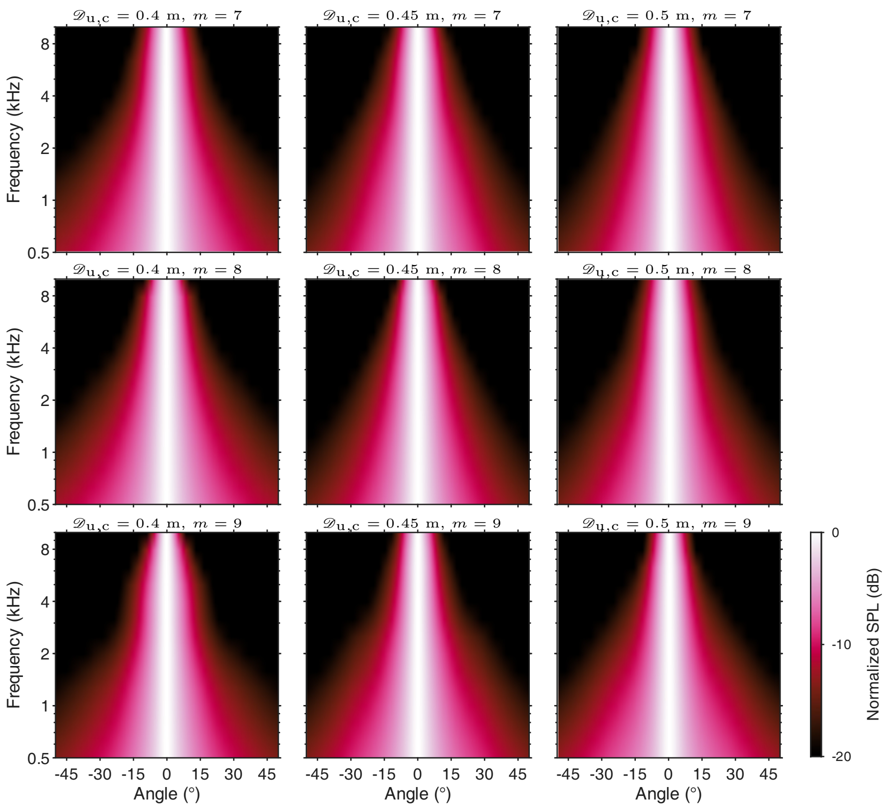


Fig. R2 | (*continued*) **e,** Ultrasound directivity patterns at the ultrasound critical distance, $\mathscr{D}_{u,c}$, showing consistent wavefront control across different geometries. **f,** Audio sound directivity patterns at the audio critical distance, $\mathscr{D}_{a,c}$, demonstrating the maintenance of high directivity across the parametric design space.

Comment 4. The core difference-frequency wave (DFW) generation technology described in the manuscript does not show essential distinctions in physical mechanisms compared with existing studies (e.g., Proc. Natl. Acad. Sci. U.S.A. 122(12), e2408975122 (2025)[3]; Phys. Rev. Lett. 131, 234001 (2023)[34]). No corresponding physical analysis is provided to confirm performance improvements in the current work.

Response 4. We thank the reviewer for this comment. As noted, the fundamental principle of DFW generation through self-demodulation is a well-established physical phenomenon, originating from the seminal work by Westervelt in 1963[6]. Over the decades, this principle has been successfully applied in highly specialized domains, such as parametric array SONAR and low-frequency sub-bottom profiling for geophysical exploration. However, while these underwater applications have proven the conceptual viability of the technology, its transition to high-performance airborne systems has been historically limited by structural and acoustic trade-offs. Despite numerous explorations in both array-based and single-transducer configurations, the technology remains largely confined to these industrial niches due to the difficulty of achieving high-fidelity performance.

We have already included both references mentioned by the referee in our original manuscript as Refs. 3, 34. As detailed below, these previous works focus primarily on the extremely localized generation of sounds through "local" acoustic nonlinear effects. More specifically, Ref. 34 utilized 1D metasurface masks for ultrasound focusing to create limited nonlinear interaction regions specifically within focal spots. Their mechanism realizes deep-subwavelength DFW focal spots, which are potentially applicable to high-precision tasks such as super-resolution imaging. Ref. 3 exploits two separate metasurfaces designed for different operating frequencies to generate two self-bending beams. The resulting ultrasonic field is specifically engineered to ensure localized nonlinear interaction beyond an obstacle (scatterer).

We acknowledge that while acoustic metasurfaces have been previously explored for nonlinear wave manipulation[3,34], the MiPAL represents a significant departure from existing approaches. Metamaterials have indeed matured into a transformative technology across various physical domains (electromagnetics, thermodynamics, elastics, and acoustics), and replacing conventional acoustic components is a natural evolution. As metamaterials demonstrate increasingly clear advantages over conventional methods, such as phased-array based wavefront control or diffractive acoustic elements, it is a natural progression to replace conventional acoustic components with metamaterials. However, it would be premature to conclude that all upcoming metamaterial-based innovations in these areas are merely incremental or similar rather than transformative. True potential of the MiPAL is realized not through simple replacement, but through the **interweaving of dual-domain metamaterials** that transcends traditional boundaries. As we noted in our responses (Response for Comment 3 by Reviewer 1 and Response for Comment 3 by Reviewer 2), we believe that the impact and novelty of a study are not determined by the simple integration of metamaterials and existing methods, but by how they are orchestrated within specific applications to reveal their versatility and potentials across diverse fields, thereby providing meaningful insights into broader engineering challenges.

Comment 5. The authors assert that the MiPAL possesses the advantage of "ultra-broadband directional sound emission". However, its broadband range of 500 Hz to 10 kHz does not demonstrate obvious superiority compared with state-of-the-art technologies (e.g., the ultra-broadband acoustic metamaterial lens reported in Appl. Phys. Lett. 127, 151701 (2025)[R28], which can omnidirectionally emit sound between 1,000 Hz and 10,000 Hz). A systematic comparison of key performance indicators (such as broadband range and directivity precision) with current state-of-the-art directional sound technologies is lacking.

Response 5. We appreciate the reviewer's constructive feedback and the opportunity to clarify the functional distinctions between our work and the cited literature. To ensure a technically accurate and contextually appropriate comparison, we would like to provide a point of clarification regarding the performance specifications of the referenced metamaterial lens[R28]. While the reviewer notes the lens can "**omnidirectionally emit sound** between 1,000 Hz and 10,000 Hz," our study focuses specifically on **directional sound emission**. As the cited paper includes both omnidirectional and directional modes, we have focused our comparison on its **directional regime**, which maintains "a consistent directivity pattern between 2,000 Hz and 20,000 Hz." We believe this alignment is necessary for a technically accurate evaluation.

To demonstrate the superiority of our system's frequency range, we provide a systematic comparison of key performance indicators grounded in the principles of human auditory perception and audio engineering standards.

1. **Broadband comparison**

   Human hearing is characterized by a logarithmic frequency resolution. A standard engineering measure of broadband utility is the octave span.

   - **Prior work**: Demonstrates a consistent directivity pattern from 2,000 Hz to 20,000 Hz, covering more than 3 octaves ($3.32 \approx \log_2 \frac{20{,}000}{2{,}000}$).
   - **This work**: Experimentally validated to achieve directional sound emission over a frequency range from 500 Hz to 10,000 Hz, covering more than 4 octaves ($4.32 \approx \log_2 \frac{10{,}000}{500}$).

   While the linear bandwidth in Hertz may appear comparable, the MiPAL system offers a **30% wider perceptual span**, capturing a significantly larger portion of the audible spectrum as processed by the human ear.

2. **Perceptual coverage**

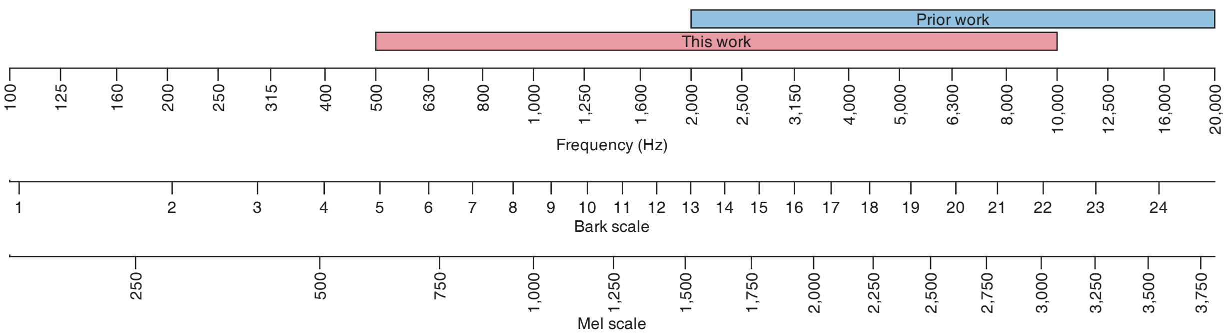


Fig. R3 | Systematic comparison of perceptual frequency coverage between the proposed MiPAL and prior works. The figure illustrates the mapping of ISO preferred frequencies (sampled at one-third octave intervals) to the critical-band rate (Bark scale) and ratio pitch (mel scale).

Response 5. (*continued*) The most robust measure of an acoustic system's broadband significance is its coverage of critical bands, which are the biological frequency-selective filters of the human hearing system. The audible frequency range up approximately to 16 kHz is subdivided into 24 abutting critical bands, measured on the "Bark" scale[R29].

- **Prior work**: Spans from Bark 13 to Bark 24. This covers 11 bands, representing less than 46% of the perceptual range.
- **This work**: Spans from Bark 5 to approximately Bark 22.5. This range accounts for 17.5 of the 24 critical bands, or approximately 73% of the functionally significant auditory spectrum.

By extending the directional emission down to 500 Hz, the MiPAL captures the lower 8 full critical bands (Bark 5 to 13) that are entirely absent in the cited directional device.

3. **Operational mechanism and performance potential**

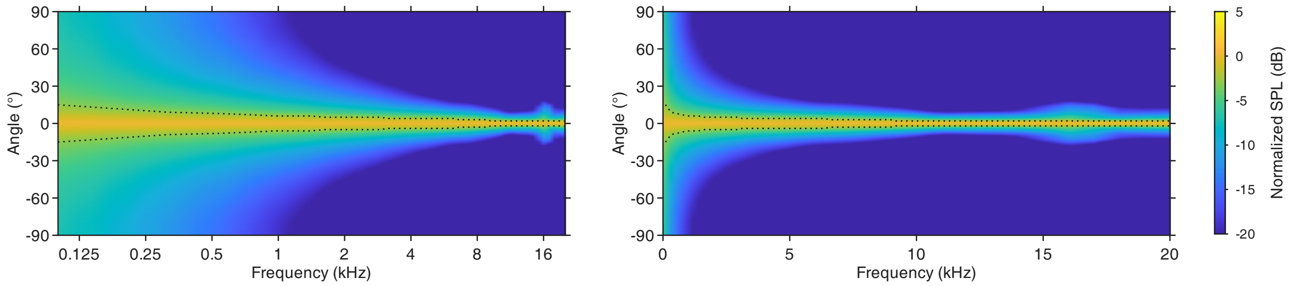


Fig. R4 | Numerical simulation of the MiPAL acoustic beam patterns presented on logarithmic (left) and linear (right) frequency axes. The dotted lines denote the half-power beamwidth (-3 dB) at each frequency.

A fundamental distinction exists between the operating mechanisms of the cited metamaterial lens and the proposed MiPAL. While the referenced lens relies on passive wavefront manipulation, it significantly loses directivity at lower frequencies. As illustrated in Fig. 3f[R28], the system's minor lobes begin to exceed the major lobe below 2 kHz, indicating a breakdown in directional control. In contrast, the MiPAL utilizes the nonlinear self-demodulation (parametric acoustic array phenomenon[R30]), allowing it to bypass the ka dependent diffraction limits that constrain passive linear-regime based devices (see Bandwidth translation from ultrasound to audio enabled by nonlinear interaction in response to Comment 2 by Reviewer 2 for details).

- **Broadband directional potential**: We have experimentally confirmed directional emission from 500 Hz to 10 kHz. Furthermore, numerical analysis demonstrates the potential for maintaining a directional beam at frequencies as low as 100 Hz, which would be unattainable using conventional linear regime-based wavefront control methods.
- **Beam pattern consistency**: To facilitate a direct comparison with the beam pattern contours (Fig. 3f[R28]) of the referenced work, we provide our results in Fig. R4 both on logarithmic and linear frequency scales.

Response 5. (*continued*)

4. **Directivity index comparison**

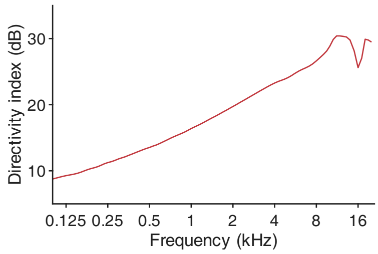


Fig. R5 | Directivity index of MiPAL, calculated at a distance of 0.5 m from the device surface.

The precision of directivity can be quantitatively evaluated using the directivity index. A comparison of the numerical values reveals a significant performance gap between the two technologies:

- **Prior work**: Exhibits a directivity index below 3 dB at 2 kHz and under 13 dB at 10 kHz (Fig. S6[R28]).
- **This work**: Achieves a directivity index of approximately 20 dB at 2 kHz and 29 dB at 10 kHz, as shown in Fig. R5.

5. **Spatial distribution and propagation integrity**
Finally, while the cited study focuses on spatial distribution in the θ direction, it lacks data regarding the acoustic field in the propagation direction. Passive wavefront manipulation can often introduce unintended focusing points or spatial artifacts[R31]. As demonstrated in Fig. 4b, the MiPAL provides a uniform and high-intensity sound beam along the propagation axis, ensuring a clean audio field for the listener.

Reviewer 3

Comment 1. The work reports the design of a novel ultrasonic transducer that leverages artificial structures—including an acoustic metasurface and elastic meta-units—to achieve directional sound. The topic is of clear interest to the acoustic wave engineering community, with promising potential for practical applications. The work presents a comprehensive study combining numerical simulations and experimental validation, including audio tests that demonstrate enhanced directivity. Overall, the paper is well organized and clearly written. However, several points require clarification or further discussion before the work can be fully assessed. My specific comments are listed follow.

Response 1. We sincerely thank the referee for your high recognition for our work. We have substantially revised the manuscript with improved clarity, providing additional analyses, simulations, experiments, and discussions to address your concern.

Comment 2. The abstract does not mention the elastic meta-unit design, despite this component being discussed extensively in the main text and playing a key role in the system's performance. Including this element in the abstract would improve consistency and better reflect the core contributions of the work.

Response 2. We sincerely thank the reviewer for this insightful suggestion. We entirely agree that the **elastic meta-units** are a critical cornerstone of our MiPAL system, and their omission from the abstract may lead to an incomplete representation of our work's core contributions. As the reviewer acknowledged, the integration of elastic meta-units is not merely an incremental addition but a key innovation that addresses a critical bottleneck in existing works. In addition, while conventional studies on locally resonant elastic metamaterials have been largely confined to idealized, infinite media (e.g., infinite elastic plates) for wave attenuation, our work represents a **pioneering transition** from conceptual demonstration to real-device adoption, for the first time. By successfully harnessing these structures within a compact, operational transducer platform, we successfully demonstrate that the utility of elastic metamaterials extends beyond fundamental physics into substantive engineering applications, offering a robust solution.

By **interweaving dual-domain metamaterials within the transducer architecture**, we achieve simultaneous control over both elastic and acoustic wave physics. This transition from isolated passive elements to a **monolithic, multi-physics transduction platform** is what makes our work **categorically distinct** from prior studies.

In accordance with the reviewer's suggestion, we have thoroughly refined the manuscript to clearly articulate the significance of interweaving dual-domain metamaterial structures—addressing both elastic and acoustic regimes—within the compact transducer platform. This revision ensures that the unique contributions of our work are accurately reflected across all key components—from the initial framing in the title and abstract to the in-depth analysis in the introduction and discussion.

Specifically, we have implemented the following changes:

- Revisit the **title** and **abstract**: The manuscript title has been revised to better reflect the core novelty and broader scientific impact of our work. As you noted, the elastic meta-unit-based engineering represents one of the key contributions. The abstract has also been revised to emphasize the main contribution and impact of our work.
- Enhanced **Introduction** and **discussion**: We have substantially expanded these sections to articulate the **categorical distinction** of our work from prior studies. The revised text clarifies how the co-design of acoustic and elastic domains enables unprecedented control over wave manipulation, bridging the gap between theoretical conceptualization and high-level functional engineering.

Response 2. (*continued*)

Title	
Original: Metasurface-integrated compact ultrasonic transducer for highly directional audio generation	**Revised:** Dual-domain metamaterials interwoven with a compact ultrasonic transducer for highly directional audio generation

Abstract	
Original: ... Here, we report an alternative form factor: a compact, single-body ultrasonic transducer augmented with an acoustic metasurface, termed metasurface-integrated parametric array loudspeakers (MiPALs). Unlike conventional PALs, the MiPAL utilizes a single piezoelectric driver combined with a co-designed metasurface, yielding a simple and cost-effective architecture that achieves compact directional sound systems. To realize this, we exploit a unified design strategy that integrates theoretical modeling, transducer-metasurface co-design, and comprehensive quantitative analyses of both linear and non-linear acoustic fields. ... These findings underscore the potential of metasurface-transducer integrated platforms for advanced audio systems, laying the foundation for developments in immersive audio, spatial sound delivery, and extended reality.	**Revised:** ... Here, we report an alternative form factor: a compact, single-body ultrasonic transducer integrated with dual-domain metamaterials, termed metamaterials-integrated parametric array loudspeakers (MiPALs). Unlike conventional PALs, the MiPAL utilizes a single piezoelectric driver interwoven with dual-domain metamaterials—consisting of a co-designed acoustic metasurface and elastic meta-units—to yield a simple, cost-effective, and highly integrated architecture. To realize this, we exploit a unified design strategy that integrates theoretical modeling, transducer-metamaterials co-design, and comprehensive quantitative analyses of both linear and non-linear acoustic fields. ... These findings underscore the potential of metamaterial-transducer integrated platforms for advanced audio systems, laying the foundation for developments in immersive audio, spatial sound delivery, and extended reality.

Response 2. (*continued*)

In the Introduction (Lines 47–58)

Original:

... In this paper, we propose a new form factor for PALs—the metasurface-integrated PAL (MiPAL)—which combines a modular, single-body piezoelectric ultrasonic transducer with a co-designed acoustic metasurface. The integrated metasurface features subwavelength spatial resolution and high design flexibility, both considered well-known advantages of metamaterials[25–32,41], and effectively redirects the non-directional ultrasound generated by the transducer's flexural vibration into highly collimated and directional wavefront. As a result, directional audible sound is produced through DFW generation via nonlinear interactions of high-intensity ultrasonic waves. Furthermore, embedding locally resonant elastic meta-units on the radiating plate effectively suppresses the omnidirectional structure-borne sound associated with the combination-resonance (CR) mode—an undesired by-product nonlinearly induced through intermodulation[24] and one of the key drawbacks of SPPALs.

Revised:

... In this paper, we propose a new form factor for PALs—the metamaterials-integrated PAL (MiPAL)—which consists of a modular, single-body piezoelectric ultrasonic transducer interwoven with dual-domain metamaterial architectures. Within this unified system, the acoustic metasurface serves as the primary functional layer; by leveraging subwavelength spatial resolution and high design flexibility, both considered well-known advantages of metamaterials[25–32], it effectively redirects the non-directional ultrasound generated by the transducer's flexural vibration into highly collimated and directional wavefront. As a result, directional audible sound is produced through DFW generation via nonlinear interactions of high-intensity ultrasonic waves. Simultaneously, the elastic meta-units act as a critical compensatory layer. By embedding deep-subwavelength tiny locally resonant structures[33] onto the transducer's radiating plate, we effectively suppress the omnidirectional sound leakage associated with the CR mode that undermines the core functionality of the PAL by compromising its defining high-directivity.

In the Introduction (Lines 61–64)

Original:

... This framework comprehensively encompasses the design of transducer components, geometries of metamaterial structures, and rigorous full-field analysis of both linear and nonlinear acoustic behavior.

Revised:

... This framework encompasses the comprehensive design of the MiPALs, integrating electro-mechanical transduction with dual-domain wave manipulation through theoretical investigation and rigorous full-field analyses of both linear and nonlinear behaviors in acoustic and elastic regimes.

Response 2. (*continued*)

In the Discussion (Lines 297–308)

Original:	**Revised:**
... In this work, we introduce the MiPAL, which fundamentally redefines the architecture for personal audio systems. By integrating an acoustic metasurface with a compact single-body ultrasonic transducer, the MiPAL achieves broadband directional sound generation without relying on costly multi-emitter arrays constructed from acoustically oversized off-the-shelf ultrasonic emitters or the complex phase-control electronics[1,19]. The metasurface redirects the incoherent flexural-mode radiation into collimated wavefront, while embedded elastic meta-units effectively suppress parasitic vibration modes that otherwise cause undesired acoustic leakage. As a result, the MiPAL demonstrates directional audio across an ultra-broad frequency range from 500 Hz to 10 kHz—covering over four octaves. This compact platform overcomes the cost, complexity, and scalability limitations of conventional PALs.	... In this work, we introduce the MiPAL, which fundamentally redefines the architecture for advanced personal audio systems through the dual-domain integration of metamaterials. By synergistically combining an acoustic metasurface and elastic meta-units with a compact single-body ultrasonic transducer, the MiPAL achieves broadband directional sound generation without relying on costly multi-emitter arrays or the complex phase-control electronics typically required in conventional systems[R1,R19]. The metasurface redirects the incoherent flexural-mode radiation into collimated wavefronts, while embedded elastic meta-units effectively suppress parasitic vibration modes that otherwise cause undesired omnidirectional leakage. As a result, the MiPAL demonstrates directional audio performance across an ultra-broad frequency range from 500 Hz to 10 kHz—covering over four octaves within a physically compact platform. Our approach overcomes the cost, complexity, and scalability limitations of conventional PALs, paving the way for widespread adoption in diverse applications.

In the Discussion (Lines 319–323)

Added:
... Furthermore, we foresee that the interweaving of architected metamaterials with real-world functional platforms will catalyze a paradigm shift across diverse fields, including biomedicine, non-destructive testing, and energy harvesting, by enabling unprecedented control over wave-matter interactions.

Comment 3. Figure 2a shows multiple geometric parameters associated with both space-coiling channels and Helmholtz resonators in the metasurface design. Since several physical mechanisms may contribute to wave collimation, it is not clear which mechanism dominates the observed directional propagation. A more detailed discussion or quantitative analysis clarifying the relative contributions of these mechanisms would strengthen the interpretation.

Response 3. We thank the reviewer for this important point. In the original manuscript, specific dimensional details were indeed insufficiently described. We apologize for any confusion this may have caused. We clarify that the primary design variables of the acoustic metasurface are the parameters t_1 and t_2, as indicated in Fig. 2a, while all other geometric parameters were held fixed throughout the design process. To achieve wave collimation, these parameters were strategically chosen to ensure a full 2π phase range with maximum transmittance, allowing the metasurface to compensate for the incoherent radiation originating from the radiating plate's structural vibrations.

To address this confusion, we have now added complete dimensional details of both the acoustic and elastic unit cells in the revised manuscript:

The caption of Fig. 2a

Original:	**Revised:**
Acoustic metasurface and its unit cells. Schematic of the metasurface colored according to the target phase distribution (left), consistent with (b) and (c). Unit-cell configuration comprising a central waveguide and Helmholtz resonators connected through a space-coiling channel (right), where the geometric parameters (t, w, $t_0 - t_2$, $w_1 - w_3$) define the acoustic response.	Acoustic metasurface and its unit cells. Schematic of the metasurface colored according to the target phase distribution (left), consistent with (b) and (c). Unit-cell configuration comprising a central waveguide and Helmholtz resonators connected through a space-coiling channel (right), where the geometric parameters ($t = 2.43$ mm, $w = 8.58$ mm, $t_0 = 2.86$ mm, $w_1 = 0.3$ mm, $w_2 = 0.9$ mm, t_1, t_2) define the acoustic response.

The caption of Fig. 2e

Original:	**Revised:**
High-order flexural-mode plate and its elastic meta-units. Comparison of the radiating plate without and with attached elastic meta-units (left). Configuration of an elastic meta-unit (right), defined by the geometrical parameters (w_{f}, h_{f}, t_{s}, l_{s}, w_{m}, d_{m}, h_{m}).	High-order flexural-mode plate and its elastic meta-units. Comparison of the radiating plate without and with attached elastic meta-units (left). Configuration of an elastic meta-unit (right), defined by the geometrical parameters ($w_{\mathrm{f}} = 3$ mm, $h_{\mathrm{f}} = 1$ mm, $d_{\mathrm{m}} = 3$ mm, t_{s}, l_{s}, w_{m}, h_{m}, h_{m}). The foot and spring segments are integrated with a 1 mm longitudinal overlap.

Comment 4. In Fig. 2c, the transmittance appears to remain strong for $t_1 \gtrsim 1.5$ mm, while the influence of t_2 is relatively weak. Given that t_1 may also be constrained by other geometric parameters (e.g., total thickness t), it is unclear whether this trend persists over a broader parameter space. Please clarify the explored parameter ranges and discuss the robustness of this observation.

Response 4. We thank the referee for this comment. As discussed, the desired phase coverage and transmittance of the proposed metasurface were achieved by varying the parameters t_1 and t_2. From the perspective of meta-atom design, t_2 governs the degree of space-coiling, and the phase delay changes according to this degree. Similarly, for the Helmholtz resonators, the reactance varies depending on the cavity volume (controlled by t_1), which also affects the phase delay. Therefore, t_1 and t_2 work complementarily to modulate the phase delay of the unit cell. In Supplementary Fig. 7a, we illustrate the locations of the unit cells constituting the metasurface by overlaying them on the transmittance and phase map. Indeed, the unit cells constituting the metasurface are not primarily located in the region where $t_1 \geq 1.5$ mm. This specific region corresponds to cases where the HR volume is very small. Furthermore, when t_2 is less than 0.6 mm—meaning there is minimal coiling by the strips—the structure becomes close to a pure waveguide. In such cases, the phase delay is nearly constant within this range.

Following the reviewer's constructive feedback, we have expanded the Supplementary Information to include a detailed description of the acoustic metasurface. The newly included Supplementary Note 5 now offers comprehensive structural details to ensure better readership and a thorough grasp of our manuscript:

In Supplementary Note 5

Added:

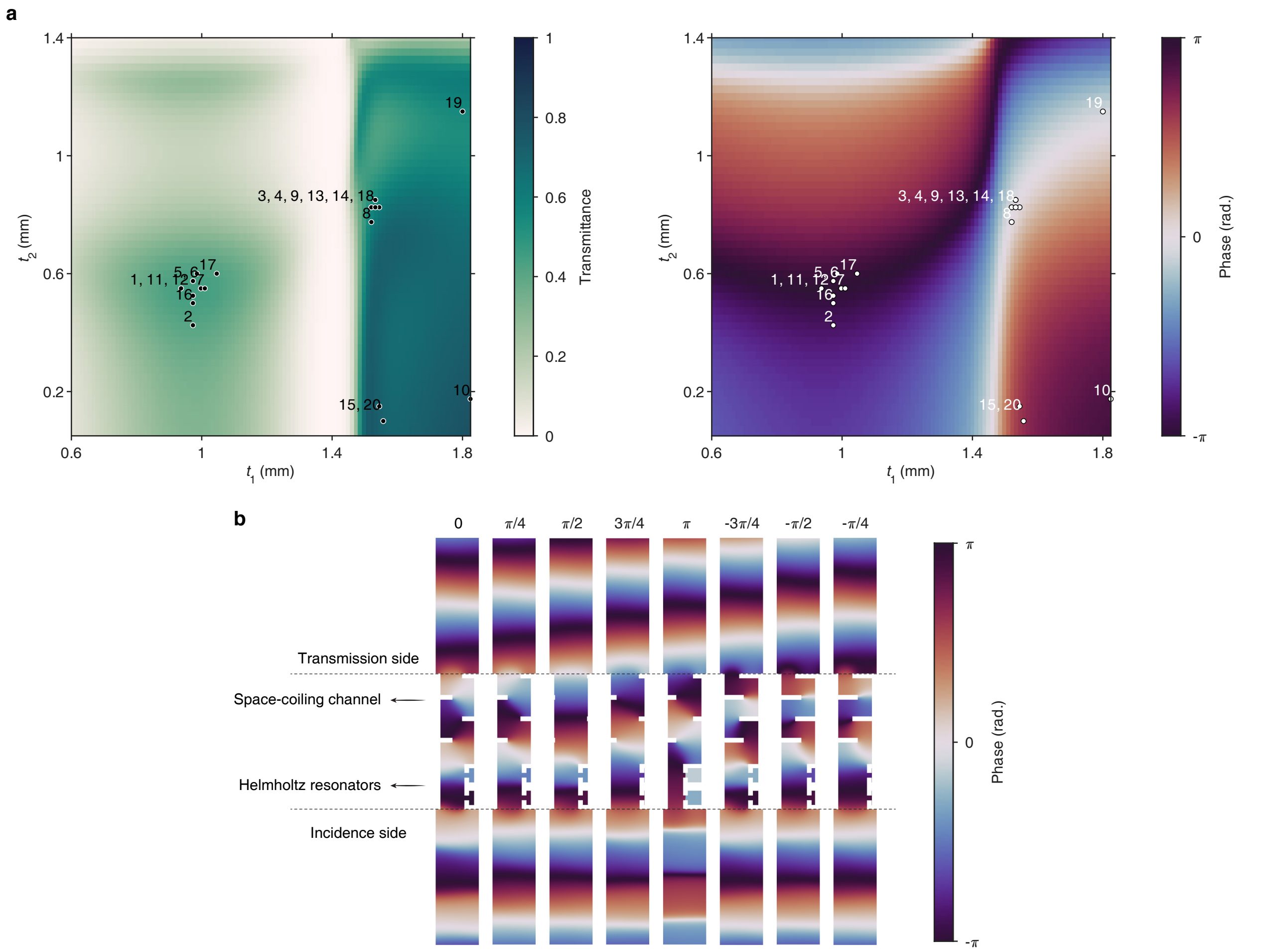
Here, we provide additional details on the design and characterization of the acoustic metasurface integrated into the MiPAL system. The metasurface is engineered to impose a specific phase profile on the ultrasonic carrier wave, tailored to match the incoherent phase profile resulting from the plate's mode shape. This spatial phase modulation compensates for the structural vibrations induced acoustic radiation, thereby enabling wavefront collimation.

As discussed, the desired phase coverage and transmittance of the proposed metasurface are achieved by varying the design parameters t_1 and t_2. In terms of straightforward meta-atom design, t_2 governs the degree of space-coiling, and the phase delay changes according to this degree. Similarly, for the Helmholtz resonators, the acoustic reactance varies depending on the cavity volume (controlled by t_1), which also affects the phase delay. Therefore, t_1 and t_2 work complementarily to modulate the phase delay of the unit cell. For clearer illustration, representative acoustic unit cells achieving full 2π phase modulation are presented in Supplementary Fig. 7. Also, the geometric dimensions of the designed acoustic metasurface are summarized in Supplementary Table 2. The first meta-atom, designed as a pure waveguide, is excluded from the table.

Response 4. *(continued)*

In Supplementary Note 5

Added:



Supplementary Fig. 7 | Further details on the acoustic metasurface design. **a,** Design space of the acoustic metasurface showing transmittance and phase distributions as functions of geometric parameters. Scattered points represent the actual meta-atoms employed in the metasurface design, with numbers indicating different meta-atom configurations. **b,** Representative acoustic unit cells achieving full 2π phase modulation with equally spaced phase shifts of $\pi/4$.

In Supplementary Note 5

Added:

Supplementary Table 2 | Geometric dimensions of the designed acoustic metasurface.

Index	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
t_1 (mm)	0.94	0.97	1.54	1.52	0.99	1.00	1.01	1.52	1.52	1.82	0.97	0.97	1.53	1.53	1.54	0.97	1.05	1.53	1.80	1.56
t_2 (mm)	0.55	0.43	0.83	0.83	0.60	0.55	0.55	0.78	0.83	0.18	0.53	0.58	0.83	0.85	0.15	0.50	0.60	0.85	1.15	0.10

Comment 5. In Fig. 2g, it is not immediately clear which curve corresponds to the case with elastic meta-units. Adding explicit labels or annotations directly in the figure would improve readability and help readers quickly identify the key results.

Response 5. We thank the reviewer for this helpful suggestion. We agree that a clear visual distinction between the configurations is essential for the proper interpretation of the results. To address this concern, we have implemented the following updates:

- Legend addition: We have updated Fig. 2g to include an explicit legend.
- Marker definitions: We revised the marker styles in this figure to ensure consistency with the audio beam patterns presented in Fig. 4f. Specifically, the circular marker (○) now represents the configuration with elastic meta-units, while the triangular marker (△) is used for the configuration without elastic meta-units.
- Cross-domain synchronization: By synchronizing these markers across the different physical domains (vibration and acoustics), we have addressed the overlapping concerns regarding figure clarity raised in Comment 10 by Reviewer 1 and Comment 5 by Reviewer 3. We believe the updated manuscript allows for a much more immediate and intuitive identification of the key results.

Comment 6. In Supplementary Fig. 11a, the reduced wavenumber appears to be incorrectly associated with the real-space axis. This labeling issue should be corrected.

Response 6. We thank the Reviewer for identifying this labeling error. We agree that the original association of the reduced wavenumber with the real-space axis was incorrect. To resolve this and provide a more rigorous physical context, we have updated Supplementary Fig. 11a and its corresponding caption as follows:

- **Inclusion of Brillouin zone contour**: We have added a schematic of the irreducible Brillouin zone (IBZ) to Supplementary Fig. 11a to clarify the path along which the dispersion relations are calculated.
- **Caption refinement**: The figure caption has been rewritten to explicitly detail the transition from the geometric schematic to the resulting band-gap physics:

The caption of Supplementary Fig. 11a	
Original: Geometrical schematic of the elastic meta-unit and the periodic plate used for band-diagram simulations (left).	**Revised:** Geometrical schematic of the elastic meta-unit and the periodic plate used for band-structure simulations, including the defined irreducible Brillouin zone (IBZ) contour used for reduced wavenumber calculation (Left).

Comment 7. The distinction between modes 1^a and 1^b, as well as 2^a and 2^b, is not clearly explained. In addition, the color bar in Fig. 2h does not seem to correspond consistently to the modes involving elastic meta-units.

Response 7. We thank the reviewer for this insightful comment. We agree that the nomenclature and the specific visualization choices for the meta-units required further clarification for technical rigor.

Regarding the nomenclature, we have added an explicit explanation in the manuscript (indicated in red) to clarify that the superscripts a and b denote the lower- and higher-frequency split modes, respectively, within each modal family. These split modes arise from the strong dynamic coupling between the plate's structural eigenmode and the local resonators.

Concerning the color bar consistency, the reviewer is correct that the meta-units in the original Fig. 2h were rendered with enhanced contrast to ensure visibility given the figure's small scale. To address this concern and provide a more scientifically rigorous representation, we have provided an enlarged and more detailed figure, Supplementary Fig. 14. In this detailed version, both the plate surface and the elastic meta-units are rendered using a normalization scale and a consistent color bar for each modal family.

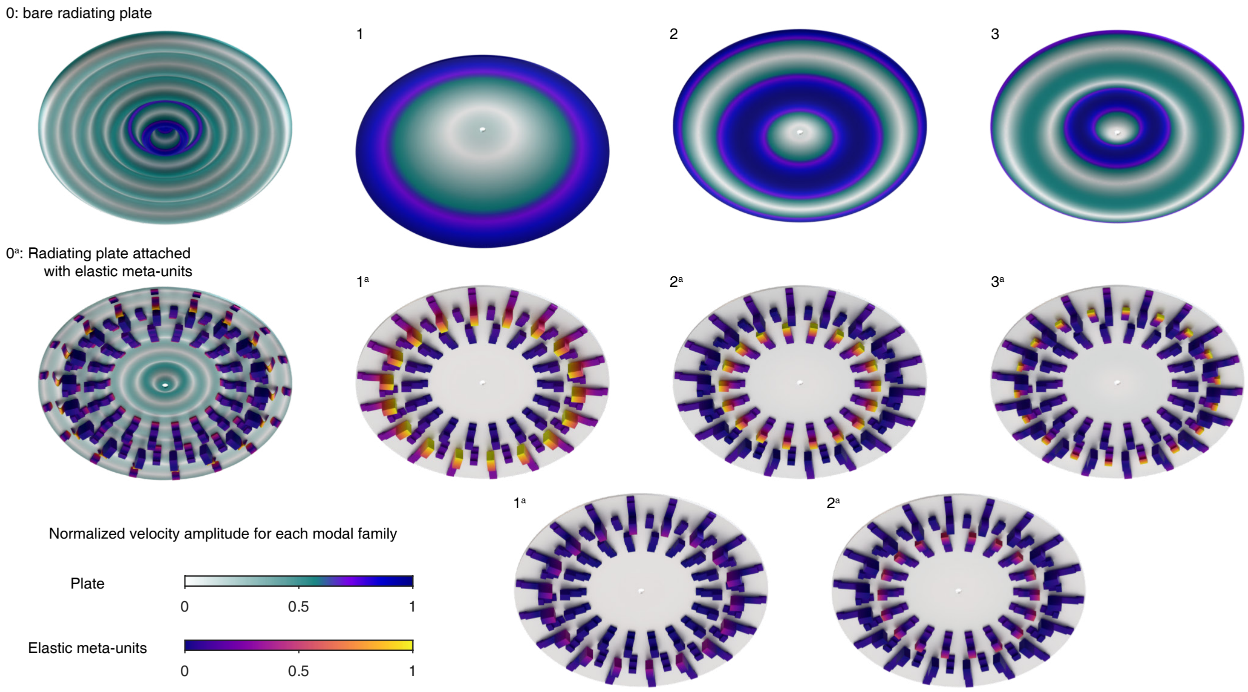
In the subsection 'Design, characterization, and fabrication of the MiPALs' (Lines 169–172)

Original:	**Revised:**
In the audible band (right panel), the lower-order resonances exhibit pronounced peak reduction and broader bandwidth ($1 \rightarrow 1^a, 1^b$; $2 \rightarrow 2^a, 2^b$; $3 \rightarrow 3^a$), indicating mode splitting and increased damping. The plate's vibration-field maps further visualize the effect of the elastic meta-units (Fig. 2h).	In the audible band (right panel), the lower-order resonances exhibit pronounced peak reduction and broader bandwidth ($1 \rightarrow 1^a, 1^b$; $2 \rightarrow 2^a, 2^b$; $3 \rightarrow 3^a$), indicating mode splitting and increased damping, where the superscripts a and b denote the lower- and higher-frequency split modes, respectively. The plate's vibration-field maps (Fig. 2h) further visualize the effect of the elastic meta-units, with an enlarged and more detailed representation provided in Supplementary Fig. 14.

Response 7. *(continued)*

In Supplementary Note 7

Added:



Supplementary Fig. 14 | Vibration-field distributions of the plate and elastic meta-units. A unified normalization scale is applied within each modal family (e.g., 1, 1^{a}, 1^{b}), while scales differ across different modal orders for visual clarity.

Comment 8. In Fig. 3b, a pronounced valley is observed around 0.23 m for different audio modulation offsets. Please provide a physical explanation or discussion of the origin of this feature.

Response 8. We thank the reviewer for this insightful observation. The local minimum (valley) observed at approximately 0.23 m is a characteristic diffraction feature of the primary ultrasonic field.

- **Physical origin**: As detailed in the revised Supplementary Note 3, this valley corresponds to the first local minimum, $r_{\min}(1)$, of a finite-sized circular radiator. This feature arises from the destructive interference between the direct geometric wave and the diffraction waves originating from the radiator's boundary, often referred to as **edge waves**[S1,S9]. For a circular aperture of radius a, this interference null occurs when the path-length difference between the center and the edge of the source is exactly one wavelength (λ) of the ultrasonic frequency.

- **Mathematical validation**: Using the analytical expressions for axial extrema provided in our revised Supplementary Note 3 (Eq. (S3)), the position of the first minimum is determined by:

$$r_{\min}(1) = \left(\frac{a}{2\lambda} - \frac{\lambda}{2a}\right) a \approx 0.22 \text{ m} \tag{R2}$$

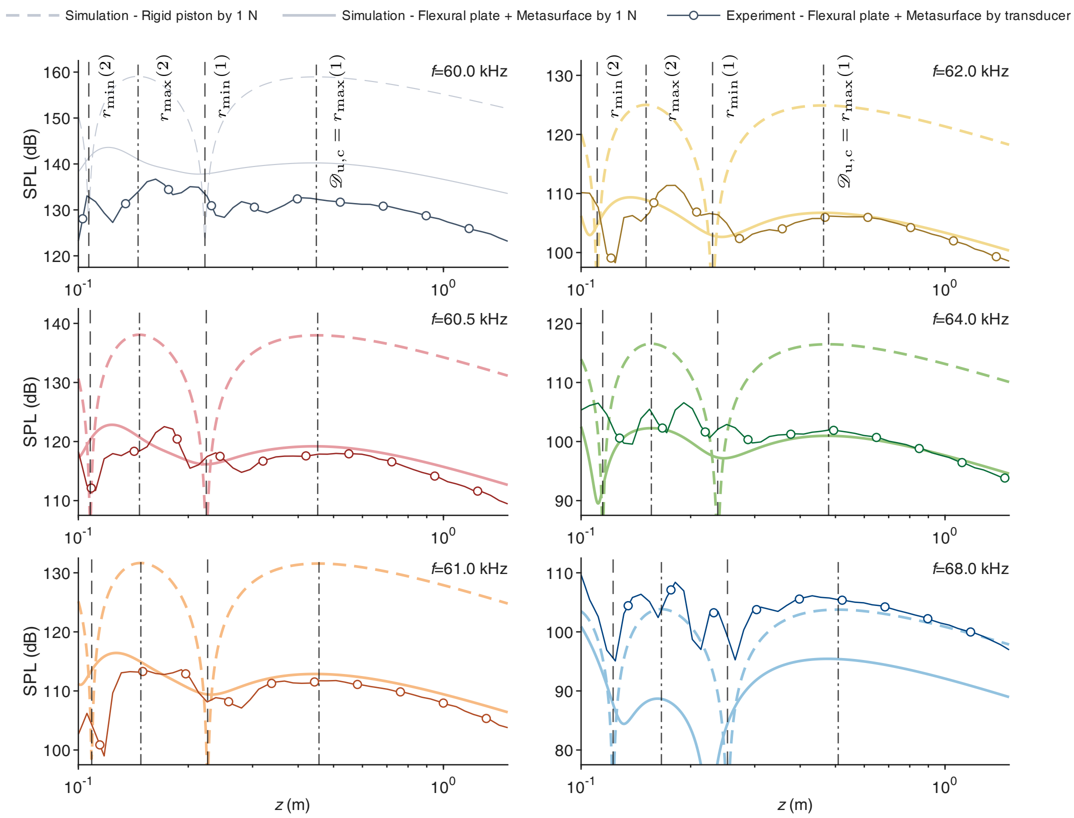
  where $a = 50$ mm, $\lambda = c/f$, $c = 343$ m/s, $f = 60$ kHz. This analytical result closely matches the observed valley position in Fig. 3b. In addition, Supplementary Fig. 4 compares the axial propagation curves of the velocity-mapped rigid piston, the simulated integrated flexural plate with the acoustic metasurface, and the experimental transducer-driven MiPAL. The precise alignment of the local extrema ($r_{\max}, r_{\min}$) across these configurations confirms that the complex structural system effectively mimics the radiation characteristics of a rigid piston.

- **Physical significance of the extrema on parametric acoustic array**:
  - The axial position of the first local maximum, $\mathscr{D}_{\mathrm{u,c}}$, represents the Westervelt far-field transition distance, which is a fundamental benchmark in parametric array design. Beyond this critical distance, the contribution of the Lagrangian density to the cumulative nonlinear interaction can be effectively ignored for a rigid piston radiator, simplifying the prediction of the secondary field[S2].
  - Furthermore, this ultrasonic critical distance is directly coupled to the audio critical distance, $\mathscr{D}_{\mathrm{a,c}}$, which we define as the target listening distance for application[S6]. By validating that our structural MiPAL assembly mimics these precise transition points of a rigid piston, we establish a reliable design bridge: the complex structural parameters of the metasurface and flexural plate are mapped onto a baseline that guarantees the desired audible output and beam characteristics at the intended listening position.

Response 8. (*continued*) We have reflected the reviewer's helpful comment and, to enhance clarity for a broad future readership, we have updated as:

In Supplementary Note 3

Added:



Supplementary Fig. 4 | Acoustic field resemblance and experimental validation of the MiPAL system. Axial propagation curves are compared for the velocity-mapped rigid piston, the simulated integrated flexural plate integrated with the acoustic metasurface, and the experimental transducer-driven MiPAL. The precise alignment of the local extrema ($r_{\max}, r_{\min}$) demonstrates that the complex structural system effectively mimics the radiation characteristics of a rigid piston.

The part of Supplementary Note 3

Added:

Furthermore, the validation of our design rationale is reinforced by comparing the theoretical propagation of the velocity-mapped rigid piston with simulated and experimental results for the fully integrated MiPAL system, as illustrated in Supplementary Fig. 4. Notably, the simulated response of the metasurface-integrated flexural plate and the corresponding experimental data—where the plate is driven by the physical transducer—demonstrate exceptional agreement in their spatial characteristics. While the similarity in absolute amplitude between the simulation and experiment is partly attributable to differences in excitation conditions—specifically, the simulation utilizes a 1 N central force while the experiment utilizes the actual transducer—the alignment of the field's spatial characteristics is highly significant. Specifically, the local extrema of the primary ultrasonic field, including the local maxima ($r_{\max}$) and the local minima ($r_{\min}$), show precise matching across all three cases.

Response 8. (*continued*)

This consistency provides strong evidence that the complex combination of a flexural circular plate and an acoustic metasurface effectively mimics the radiation characteristics of an ideal rigid piston.

$$r_{\max}(n) = \left(\frac{a}{m\lambda} - \frac{m\lambda}{4a}\right)a, \qquad m = 2n - 1 \tag{S2}$$

$$r_{\min}(n) = \left(\frac{a}{m\lambda} - \frac{m\lambda}{4a}\right)a, \qquad m = 2n \tag{S3}$$

where $n = 1, 2, 3, ...$ indicates the order of the local extrema, a is the radius of the circular piston, and λ is the wavelength of the ultrasonic frequency[S2]. The observed peaks and valleys arise from the superposition of the direct axial wave and the phase-shifted edge waves originating at the piston's boundary, a classic characteristic of finite-aperture diffraction[S1,S9].

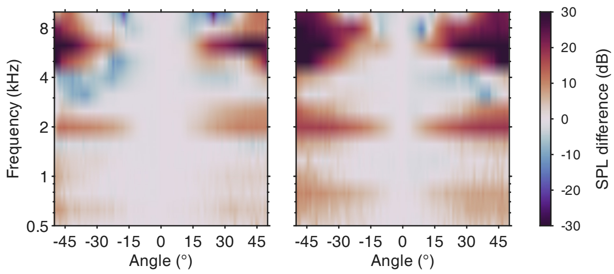
Comment 9. Figure 4d indicates that elastic meta-units enhance performance mainly within specific low-frequency ranges (approximately 0.5−0.8 kHz, ∼2 kHz, and ∼6.2 kHz). Outside these ranges, it is less clear that the configuration with elastic meta-units consistently outperforms the one without them. If the primary role of the elastic meta-units is to suppress combination-resonance modes, as suggested in Fig. 1g, then the acoustic metasurface appears to be the dominant contributor to high directivity. This point deserves further discussion. Additionally, it would be helpful to include supplementary audio test comparisons—similar to Fig. 5d—directly contrasting configurations with and without the acoustic metasurface.

Response 9. We agree with the reviewer's assessment that the acoustic metasurface is the primary driver of high directivity. As correctly noted, the two components of our design serve distinct yet synergistic roles: the acoustic metasurface addresses large plate breakup and phase cancellation (Fig. 1f), while the elastic meta-units are specifically engineered to suppress the combination resonance (CR) modes (Fig. 1g).

To clarify the effectiveness of the elastic meta-units in improving audio beam patterns, we have added Supplementary Fig. 18. This figure illustrates the discrepancy between experimental data and ideal simulation results (which exclude CR effects). The near-white regions in the left panel demonstrate that the elastic meta-units successfully align the physical device's performance with theoretical predictions by suppressing CR modes. Conversely, in the configuration without these units (right panel), significant discrepancies emerge—particularly near the plate's second-order resonance frequency (∼2 kHz). While the current suppression is effective, further optimization of manufacturing tolerances, meta-unit density, or individual unit mass could potentially enhance this effect[R1], as elaborated in Supplementary Note 7.

In Supplementary Note 9

Added:



Supplementary Fig. 18 | Discrepancies between experimental data and simulation results for audio beam patterns (Fig. 4d), comparing configurations with and without the elastic meta-units. A divergent colormap is employed, where white indicates zero difference. The configuration with elastic meta-units (left), demonstrating strong agreement with simulation results. In contrast, the case without the meta-units (right), revealing significant discrepancies—particularly at the plate's second-order resonance frequency—arising from combination resonance effects.

In the subsection 'Experimental characterization of the MiPAL' (Lines 246–248)

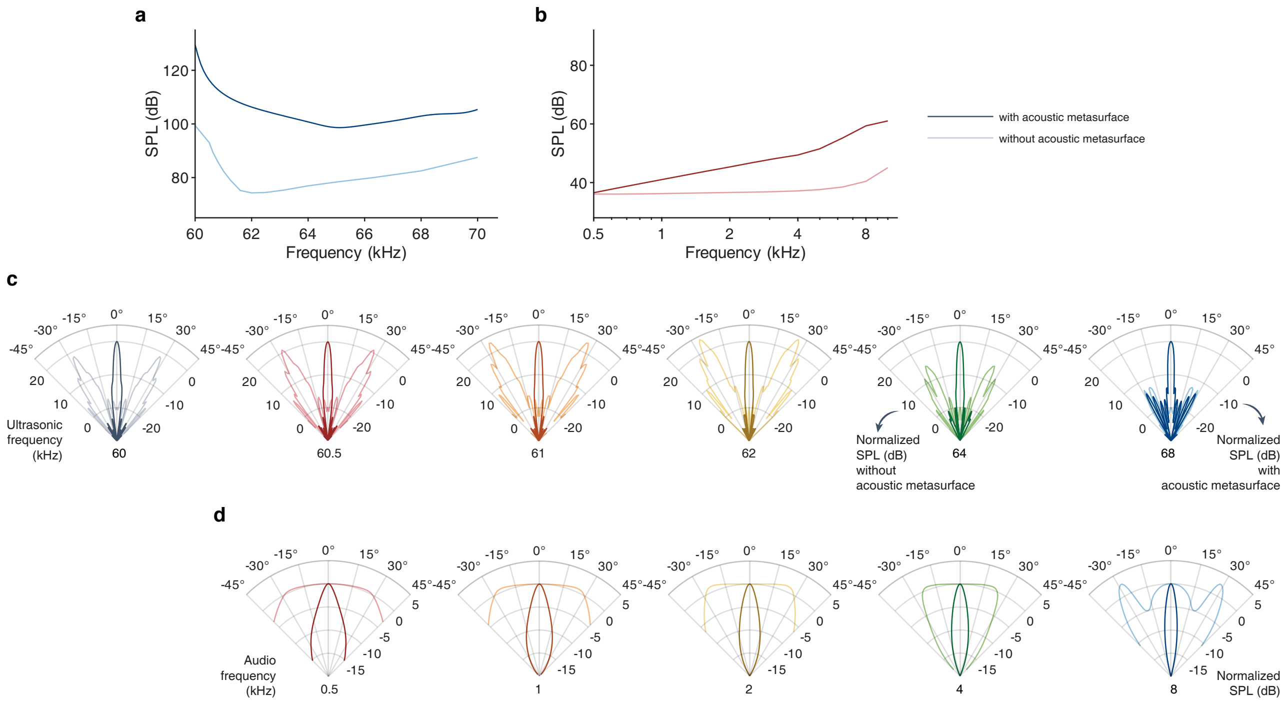
Added:

... The quantified discrepancy in Supplementary Fig. 18 further demonstrates that the meta-units effectively align experimental performance with theoretical predictions by suppressing parasitic CR modes. ...

Response 9. (*continued*) Secondly, to substantiate the claim that the acoustic metasurface is the dominant contributor to high directivity, we have included a comparative analysis in Supplementary Note 5 (Supplementary Fig. 8). These results confirm that without the acoustic metasurface, the plate undergoes significant breakup, causing out-of-phase pressure cancellation that diminishes ultrasonic output and destroys directional integrity. The integration of the acoustic metasurface recovers a coherent, high-intensity ultrasonic field, providing the foundation for a highly directional parametric array.

In Supplementary Note 5

Added:



Supplementary Fig. 8 | Comparative analysis of frequency responses and directional beam characteristics with and without the acoustic metasurface. **a, b,** Simulated ultrasonic and audio frequency responses at 0.5 m along the acoustical axis, comparing the metasurface-integrated plate with a bare plate configuration. **c,** Simulated representative ultrasonic beam patterns at the carrier and sideband frequencies (60, 60.5, 61, 62, 64, and 68 kHz) at 0.5 m, normalized to the on-axis intensity. **d,** Simulated representative audio beam patterns at corresponding frequencies (0.5, 1, 2, 4, and 8 kHz) at 0.5 m, normalized to the on-axis intensity.

In Supplementary Note 5

Added:

Functional impact of the acoustic metasurface

The functional necessity of the acoustic metasurface is demonstrated by comparing its performance against a bare plate configuration. Without the metasurface, the plate undergoes breakup—as illustrated in Fig. 1f—where mechanical bending wavelengths become comparable to the diaphragm's dimensions. This causes different regions of the plate to vibrate in opposite directions, creating out-of-phase pressures that cancel each other in the air. Consequently, as shown in Supplementary Figs. 8a

Response 9. (*continued*)

> and 8b, the bare plate cannot generate the high-intensity ultrasonic carrier required for efficient parametric interaction, resulting in diminished ultrasonic output and a loss of directional integrity. The acoustic metasurface recovers a coherent, high-intensity ultrasonic field that effectively beams the power directly in front of the transducer. This is evident in Supplementary Fig. 8c, where the integrated system maintains narrow, high-directivity ultrasonic beam patterns. Furthermore, the audio beam patterns in Supplementary Fig. 8d confirm the effectiveness of the acoustic metasurface in translating this directional intensity into the audible range.

Lastly, we would like to thank the Reviewer once more for careful evaluation of our manuscript and for taking the time to review the Supplementary Videos. We fully understand your point regarding the supplementary audio test comparisons. However, we would like to respectfully clarify the purposes of our audio demonstrations. The primary objective of the current Supplementary Videos is to highlight the improvement in the actual listening experience enabled by the elastic meta-units, specifically the suppression of the severe "ringing effect" in the time domain induced by the parastic CR mode.

On the other hand, the primary function of the acoustic metasurface is to achieve high directivity. We believe the spatial compensation effects of the acoustic metasurface are already rigorously quantified and directly contrasted in the main text through the experimental directivity patterns in Figs. 3d and 3e. If we were to provide an audio comparison with and without the acoustic metasurface, the difference would primarily manifest as a variation in SPL. In fact, the directional sound experience enabled by the acoustic metasurface can be captured in the current Supplementary Videos; as the recording angle moves away from the on-axis, the SPL noticeably decreases, effectively demonstrating the high directivity.

Therefore, we believe the current combination of the directivity patterns provided in the main text and the Supplementary Videos sufficiently conveys the performance of the MiPALs.

References

[R1] Sugino, C., Xia, Y., Leadenham, S., Ruzzene, M. & Erturk, A. A general theory for bandgap estimation in locally resonant metastructures. *J. Sound Vibr.* **406**, 104–123 (2017).

[R2] Van Belle, L., Claeys, C., Deckers, E. & Desmet, W. The impact of damping on the sound transmission loss of locally resonant metamaterial plates. *J. Sound Vibr.* **461**, 114909 (2019).

[R3] de Melo Filho, N. G. R., Van Belle, L., Claeys, C., Deckers, E. & Desmet, W. Dynamic mass based sound transmission loss prediction of vibro-acoustic metamaterial double panels applied to the mass-air-mass resonance. *J. Sound Vibr.* **442**, 28–44 (2019).

[R4] Whitwell, T. *The Balanced Mode Radiator (BMR)*. Tectonic Audio Labs, Inc., Woodinville, WA (2019). Revision 2.0. Available at: https://www.hifi-selbstbau.de/images/stories/chassis/tectonic_TEBM46C20N-4B/Tectonic_BMRWhitePaper_Rev2.0_2019-1.pdf, accessed: 2026-01-28.

[R5] Bank, G. & Harris, N. Acoustic device and method of making acoustic device. U.S. Patent 7,916,878 B2 (2011). Assigned to Tectonic Audio Labs, Inc. Available at: https://patents.google.com/patent/US7916878B2, accessed: 2026-01-28.

[R6] Duda, R. O. & Martens, W. L. Range dependence of the response of a spherical head model. *J. Acoust. Soc. Am.* **104**, 3048–3058 (1998).

[R7] Zhong, J., Kirby, R., Karimi, M., Zou, H. & Qiu, X. Scattering by a rigid sphere of audio sound generated by a parametric array loudspeaker. *J. Acoust. Soc. Am.* **151**, 1615–1626 (2022).

[R8] Burkhard, M. & Sachs, R. Anthropometric manikin for acoustic research. *J. Acoust. Soc. Am.* **58**, 214–222 (1975).

[R9] Vezzá, C., Moraes, E. & Ooki, F. Study of acoustic absorption of seat foams and fabrics. Tech. Rep., SAE Technical Paper (2015).

[R10] McMullan, A. & Mealman, M. An investigation of automotive seat fabric sound absorption. Tech. Rep., SAE Technical Paper (2001).

[R11] Ward, G. *et al.* Boundary-layer effects on acoustic transmission through narrow slit cavities. *Phys. Rev. Lett.* **115**, 044302 (2015).

[R12] Jiang, X., Li, Y. & Zhang, L. Thermoviscous effects on sound transmission through a metasurface of hybrid resonances. *J. Acoust. Soc. Am.* **141**, EL363–EL368 (2017).

[R13] Beyer, R. T. in *The parameter B/A* (eds Hamilton, M. F. & Blackstock, D. T.) *Nonlinear Acoustics* 23–37 (Springer Nature Switzerland, Cham, 2024).

[R14] Li, M., Zhong, J., Jing, Y. & Lu, J. A note on the audio sound power generated by a parametric array loudspeaker. *J. Acoust. Soc. Am.* **154**, 3899–3905 (2023).

[R15] Worthington, D. A. *et al.* Comparing two methods to measure preferred listening levels of personal listening devices. *J. Acoust. Soc. Am.* **125**, 3733–3741 (2009).

[R16] Kim, C., Hwang, Y., Shin, K. & Moon, W. A parametric array loudspeaker featuring a langevin transducer, a circular aluminum plate, and composite polymer steps. *Sens. Actuator A-Phys.* **338**, 113466 (2022).

[R17] Oh, B., Kim, C., Lee, D., Rho, J. & Moon, W. An improved analytic model for designing the polymer-composite stepped-plate transducer using the modified mindlin plate theory. *Ultrasonics* **131**, 106933 (2023).

[R18] Kim, W., Oh, B., Kim, C. & Moon, W. Design, analysis, and experimental validation of a stepped plate parametric array loudspeaker. *J. Acoust. Soc. Am.* **158**, 2561–2576 (2025).

[R19] Barone, A. & Gallego Juarez, J. A. Flexural vibrating free-edge plates with stepped thicknesses for generating high directional ultrasonic radiation. *J. Acoust. Soc. Am.* **51**, 953–959 (1972).

[R20] Je, Y., Lee, H., Park, J. & Moon, W. A stepped-plate bi-frequency source for generating a difference frequency sound with a parametric array. *J. Acoust. Soc. Am.* **127**, 3494–3502 (2010).

[R21] Cheng, Z.-q., Chen, A., Yang, J., Liang, B. & Cheng, J.-c. Broadband achromatic acoustic vortex generator based on integrated-resonant meta-atoms. *Appl. Phys. Lett.* **122**, 183506 (2023).

[R22] Cummer, S. A., Christensen, J. & Alù, A. Controlling sound with acoustic metamaterials. *Nat. Rev. Mater.* **1**, 16001 (2016).

[R23] Ma, G. & Sheng, P. Acoustic metamaterials: From local resonances to broad horizons. *Sci. Adv.* **2**, e1501595 (2016).

[R24] Assouar, B. *et al.* Acoustic metasurfaces. *Nat. Rev. Mater.* **3**, 460–472 (2018).

[R25] Wang, S. *et al.* A broadband achromatic metalens in the visible. *Nat. Nanotechnol.* **13**, 227–232 (2018).

[R26] Chen, W. T. *et al.* A broadband achromatic metalens for focusing and imaging in the visible. *Nat. Nanotechnol.* **13**, 220–226 (2018).

[R27] Dong, H.-W. *et al.* Achromatic metasurfaces by dispersion customization for ultra-broadband acoustic beam engineering. *Natl. Sci. Rev.* **9**, nwac030 (2022).

[R28] Peng, X., Hernandez, G., Cummer, S. A., Hoffmann, F.-M. & Li, J. Ultra-broadband acoustic metamaterial lens for loudspeaker directivity control. *Appl. Phys. Lett.* **127** (2025).

[R29] Fastl, H. & Zwicker, E. *Critical Bands and Excitation*, 149–173 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2007).

[R30] Westervelt, P. J. Parametric acoustic array. *J. Acoust. Soc. Am.* **35**, 535–537 (1963).

[R31] Kim, S., Kim, J., Kim, K., Jeong, M. & Rho, J. Anti-aliased metasurfaces beyond the Nyquist limit. *Nat. Commun.* **16**, 411 (2025).

Response to Reviewers

Dual-domain metamaterials co-integrated with a compact ultrasonic transducer for highly directional audio generation (NCOMMS-25-100737A-Z)

Woongji Kim[1,2†], Beomseok Oh[1†], Wonkyu Moon[1*], Junsuk Rho[1,3,4,5*]

[1]Department of Mechanical Engineering, Pohang University of Science and Technology (POSTECH), Cheongam-ro 77, Pohang-si, 37683, Gyeongsangbuk-do, Republic of Korea. [2]Department of Mechanical Engineering, Ulsan National Institute of Science and Technology (UNIST), UNIST-gil 50, Ulsan-si, 44919, Republic of Korea. [3]Department of Chemical Engineering, Pohang University of Science and Technology (POSTECH), Cheongam-ro 77, Pohang-si, 37683, Gyeongsangbuk-do, Republic of Korea. [4]Department of Electrical Engineering, Pohang University of Science and Technology (POSTECH), Cheongam-ro 77, Pohang-si, 37683, Gyeongsangbuk-do, Republic of Korea. [5]POSCO-POSTECH-RIST Convergence Research Center for Flat Optics and Metaphotonics, Pohang University of Science and Technology, Cheongam-ro 77, Pohang-si, 37683, Gyeongsangbuk-do, Republic of Korea.

*Corresponding author(s). E-mail(s): wkmoon@postech.ac.kr; jsrho@postech.ac.kr; Contributing authors: wj.kim@postech.ac.kr; bs.oh@postech.ac.kr; †These authors contributed equally to this work.

General Response

We would like to express our sincere gratitude to the Editor and the Reviewers for their rigorous and insightful evaluations of our manuscript now retitled "**Dual-domain metamaterials co-integrated with a compact ultrasonic transducer for highly directional audio generation**," submitted to *Nature Communications* (NCOMMS-25-100737A-Z). Their suggestions have been instrumental in refining the conceptual clarity and enhancing the academic rigor of our work.

In the detailed point-by-point response that follows, we have revised the manuscript and Supplementary Information accordingly and added further explanations and supporting content to address each concern. We once again thank the Reviewers for their invaluable time and constructive evaluation, and we hope that our revised manuscript will now be deemed suitable for publication in *Nature Communications*.

Reviewer 1

Comment 1. The authors have fully addressed my concerns, and I am happy to recommend the manuscript for publication.

Response 1. We are grateful to the reviewer for recommending our manuscript for publication. Thanks to the reviewer's rigorous and insightful comments, we believe the manuscript has been significantly improved and is now substantially stronger than the original submission.

Reviewer 2

Comment 1. The authors have responded point-by-point to initial comments, yet the core flaws—insufficient substantial novelty and inadequate elucidation of fundamental physical mechanisms—remain unresolved. This manuscript does not meet the rigorous publication standards of Nature Communications and it may be more suitable for a specialized journal such as Communications Engineering, where its practical implementation and platform-specific contributions would be better appreciated.

Response 1. We thank the reviewer for revisiting and re-evaluating our work. While we fully understand and respect the reviewer's perspective, we firmly believe that our work meets the standards for publication in *Nature Communications*. In what follows, we provide a more comprehensive elaboration to address reviewer's specific concerns point by point.

Comment 2. The authors claim to have expanded discussions on the MiPAL's novelty and physical mechanisms, but only conducted superficial textual rephrasing (e.g., rhetorical emphasis on "Dual-domain metamaterials interwoven") with no core physical mechanism analysis added. The so-called "dual-domain metamaterial integration" is merely a conceptual description, which fails to reveal the multi-physical field coupling laws between the acoustic metasurface and elastic meta-units, nor explain how this coupling achieves performance breakthroughs unavailable with single-domain technologies. Their argument of "dual-domain interweaving rather than simple integration" is only conceptual. No specific theoretical framework or engineering methodology for the proposed "co-design" paradigm is provided, leaving it without practical support.

Response 2. We thank the reviewer for this comment and would like to clarify what we believe to be a misunderstanding of the physical structure of MiPAL. We are concerned that some elements of our previous rebuttal may not have been fully reflected in the present evaluation. Here, we once again summarize the representative breakthroughs that the dual-domain integration enables and that previous approaches alone **cannot achieve**, together with their original locations in the submitted manuscript and Supplementary Information. We respectfully ask the reviewer to consider the present response in conjunction with our previous rebuttal.

First of all, we would like to clarify what appears to be a misunderstanding regarding the term "coupling" in the context of MiPAL.

- **Not a field-to-field coupling law:** The acoustic metasurface and the elastic meta-units do not directly couple acoustic and elastic waves, nor do they constitute a new physical coupling law between the two metamaterial components. Rather, our claim concerns a co-design strategy involving dual-domain metamaterials: a system-level integrated design methodology in which the transducer, acoustic metasurface, and elastic meta-units are jointly designed through the shared radiating plate to realize the target device functionality, not a physical interaction law between the two domains.
- **Different operating regimes:** The acoustic metasurface operates in the ultrasonic acoustic regime near ~ 60 kHz, where it compensates for the phase profile generated by the high-order flexural vibration of the radiating plate. The elastic meta-units operate in the audible structural-vibration regime in the kHz range, where they suppress lower-order plate modes responsible for CR-induced omnidirectional leakage.
- **System-level integration:** From this perspective, "coupling" refers to system-level integration of two domain-specific controls within one compact PAL architecture, rather than direct physical coupling between the acoustic metasurface and the elastic meta-units.
- **Core contributions:** Acoustic-domain control alone cannot suppress CR-induced audible leakage, while elastic-domain control alone cannot generate the coherent ultrasonic primary field required for directional audio generation. These two domains play complementary roles, and MiPAL therefore requires simultaneous, transducer-mediated consideration of both. To the best of our knowledge, this device-level dual-domain co-design strategy has not been previously demonstrated.

To make this evidence easy to locate, the quantitative demonstrations already provided across the manuscript and Supplementary Information are summarized in Table R1. Several of these—in particular the bare-plate functional-impact analysis and the stiffness-detuning robustness study—were added during the previous round of revision specifically to substantiate the dual-domain necessity claim.

Response 2. *(continued)*

Table R1 | Quantitative evidence that single-domain control is insufficient for the targeted device functionality. Each row pairs a domain-specific limitation with the corresponding simulation–experiment evidence already reported in the manuscript or Supplementary Information.

Single-domain limitation	Evidence	Figure
Acoustic-domain only (no elastic meta-units): CR-induced audible leakage masks the directional DFW beam	Audio beam pattern at 2 kHz and CR-band frequencies, w/wo meta-units	Figs. 4d and 4f
	Plate center-velocity FR showing CR mode suppression on attaching meta-units	Fig. 2g; Supplementary Fig. 14d
	Audio beam-pattern simulation–experiment discrepancy map (CR contribution isolated)	Supplementary Fig. 21
	Robustness of CR suppression and leakage levels under ±20% stiffness detuning	Supplementary Fig. 16
Elastic-domain only (no acoustic metasurface): incoherent flexural radiation precludes the coherent ultrasonic primaries required by PA	Bare-plate ultrasonic on-axis pressure level at 0.5 m	Supplementary Fig. 8a
	Bare-plate ultrasonic beam patterns over 60–70 kHz	Figs. 3d and 3e; Supplementary Fig. 8c
	Demodulated audio level and beam pattern with the bare plate	Supplementary Figs. 8b and 8d

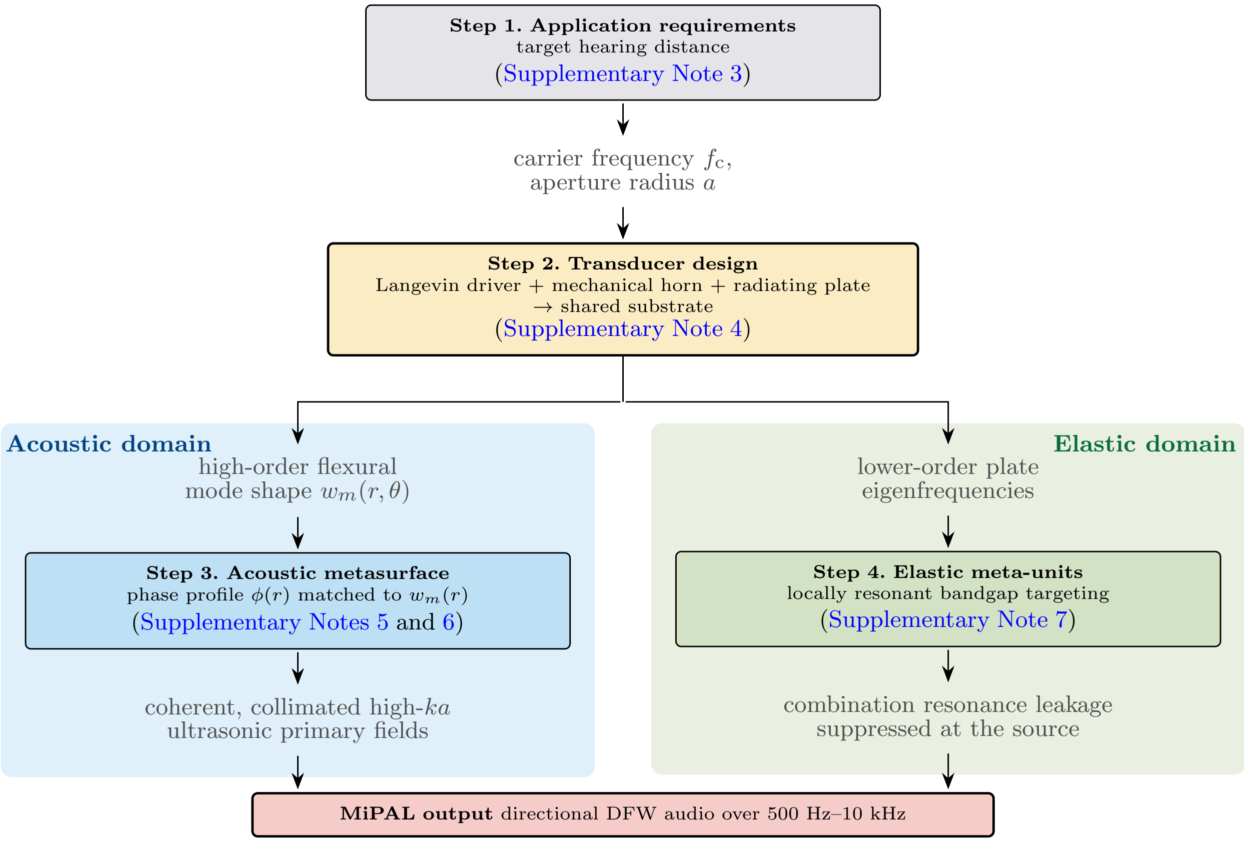


Fig. R1 | Sequential co-design workflow of the MiPAL. Each stage's quantitative output feeds the next, defining a reproducible engineering pipeline rather than a heuristic concept. Step 1 fixes the application target (target hearing distance), yielding the carrier frequency f_c and aperture radius a that drive Step 2. The radiating plate is the shared substrate (yellow) from which the workflow branches into two physically decoupled domains—acoustic (blue) and elastic (green)—integrated through the plate rather than through field coupling. Each domain resolves a distinct device-level bottleneck of the compact PAL architecture, and the combined output (red) is directional DFW audio.

Response 2. (*continued*)
Secondly, we respectfully disagree with the statement that "No specific theoretical framework or engineering methodology for the proposed co-design paradigm is provided, leaving it without practical support." This statement does not reflect the methodological contents included in the manuscript and Supplementary Information. As shown in Fig. R1, we have provided a sequential design framework in which each stage has a specific physical basis, input, and output.

In summary, the dual-domain integration of MiPAL is not a textual reframing but a co-design strategy substantiated by single-domain insufficiency demonstrations (Table R1) and a reproducible sequential design framework (Fig. R1). We respectfully hope this consolidation allows the reviewer to reassess our contribution.

Comment 3. For broadband dispersion control, the supplementary upper-sideband modulation selection and phase-matching RMSE analysis are merely simple numerical simulations. They fail to uncover the physical nature of the coupling between the transducer's structural dispersion and the acoustic metasurface's dispersion, and no quantitative theoretical model for dispersion control is established. Thus, generalizability and effectiveness of the strategy cannot be verified.

Response 3. We thank the reviewer for this point and use this opportunity to clarify the dispersion-control strategy of our work and the basis on which its generalizability and effectiveness rest. Our approach is a "dispersion-aware" design direction that occupies a position between *no dispersion treatment* and *full dispersion engineering*. Rather than designing a fully dispersion-controlled, optimized metasurface such as an achromatic focusing metalens[R1], we adopt a single physically anchored design rule: from the structural dispersion of the high-order flexural plate, we identify the sideband over which the phase slopes of the transducer and the metasurface are best matched, and accordingly adopt that band for modulation (Supplementary Note 6). Indeed, the broadband performance achieved both numerically and experimentally (Figs. 4 and 5 in the main text) demonstrates the validity of this dispersion-aware design choice, confirming that the upper-sideband selection rule alone is sufficient for the present system. We of course recognize that more sophisticated dispersion-engineering approaches well established in prior studies[R1–R6] could further enhance wavefront quality; however, this does not detract from the novelty of the proposed form factor. Rather, the same rule applies to other compact PAL design cases of comparable architectural configuration, with full dispersion-engineering techniques available as a natural escalation when target platforms demand stronger treatment—a direction already indicated in the previous revised manuscript.

This generalizability was further demonstrated in our previous response to Comment 3 by Reviewer 2 through parametric validation across various design splits ($\mathscr{D}_{\mathrm{u,c}} \in \{0.40, 0.45, 0.50\}$ m $\times$ $m \in \{7, 8, 9\}$), where the same design rule consistently produced robust collimated ultrasound beams and high-directivity audio fields well-matched to the rigid-piston based theoretical predictions described in Supplementary Note 3. For the benefit of future readers, we have now incorporated this analysis into the revised Supplementary Information (Supplementary Note 6), so that the evidence supporting the generalizability of the design rule is demonstrated directly (please refer to the changes below).

From this perspective, on the request for a quantitative theoretical model for dispersion control, we respectfully note that developing such a model is not the scope of the present work. The generalizability and effectiveness of our strategy do not rest on theoretical-model-based dispersion engineering; rather, they rest on demonstrating that the "newly" proposed metamaterials-integrated platform delivers the desired compact PAL performance more effectively than conventional approaches—in particular, by being substantially more compact and cost-effective than array-based PALs (Supplementary Note 2) while simultaneously outperforming existing single-driver-based PALs. This quantitative head-to-head comparison was already added in the previous round of revision (Supplementary Note 16). More broadly, we believe this design approach can offer useful insight that extends beyond PAL applications—to other **metamaterials-integrated devices** that require comprehensive wave control at the device level. We therefore hope that the reviewer will reconsider our work from this broader perspective—as a contribution that opens new directions for metamaterials-augmented functional device design.

Response 3. (*continued*)

In Supplementary Note 6

Added:

Generalizability of the upper-sideband design rule across radiator dimensions

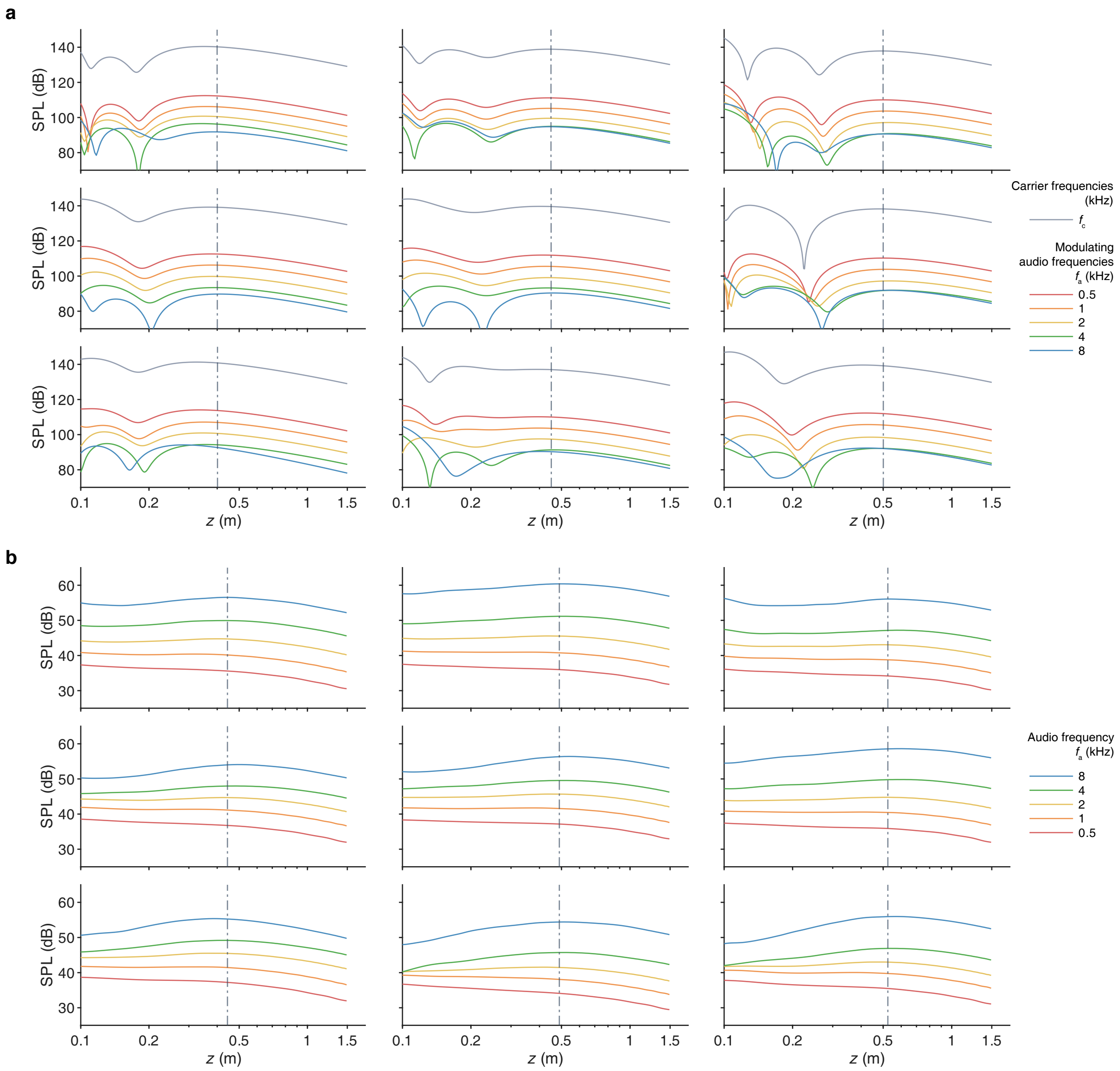
To establish that the upper-sideband design rule introduced in the preceding subsections generalizes beyond the reference configuration ($\mathscr{D}_{\mathrm{u,c}} = 0.45$ m, $m = 8$), parametric simulations were carried out across a range of radiator dimensions. Specifically, the design space was sampled across $\mathscr{D}_{\mathrm{u,c}} \in \{0.40, 0.45, 0.50\}$ m $\times$ $m \in \{7, 8, 9\}$, defining nine design cases. For each combination, the transducer, plate, and acoustic metasurface were redesigned following the methodology summarized in Supplementary Note 3, and the upper-sideband modulation band was selected following the dispersion-aware phase-matching criterion quantified by Supplementary Fig. 10 and Eq. (S4). The resulting acoustic fields are reported in Supplementary Figs. 11 to 13.

The propagation curves in Supplementary Fig. 11 show that the on-axis ultrasonic and demodulated audio fields obtained for every case closely follow the rigid-piston-based predictions detailed in Supplementary Note 3, with the predicted critical distances $\mathscr{D}_{\mathrm{u,c}}$ and $\mathscr{D}_{\mathrm{a,c}}$ recovered consistently across all nine $(\mathscr{D}_{\mathrm{u,c}}, m)$ combinations. The frequency responses in Supplementary Fig. 12 confirm that the upper-sideband modulation strategy preserves broadband fidelity for both the ultrasonic carrier and the demodulated audio band, with no abrupt phase-induced distortions emerging at any of the cases considered. The directivity patterns in Supplementary Fig. 13 further indicate that wavefront control by the metasurface and high-directivity audio beam formation are maintained across the design space. Together, these results substantiate that the design rule developed in this note—comprising the rigid-piston-based parameter selection (Supplementary Note 3), the dispersion-aware upper-sideband modulation choice (Supplementary Fig. 10 and Eq. (S4)), and the metasurface phase compensation (Supplementary Fig. 9)—is generalizable to MiPAL configurations of comparable architecture across the considered range of radiator scales.

Response 3. (*continued*)

In Supplementary Note 6

Added:

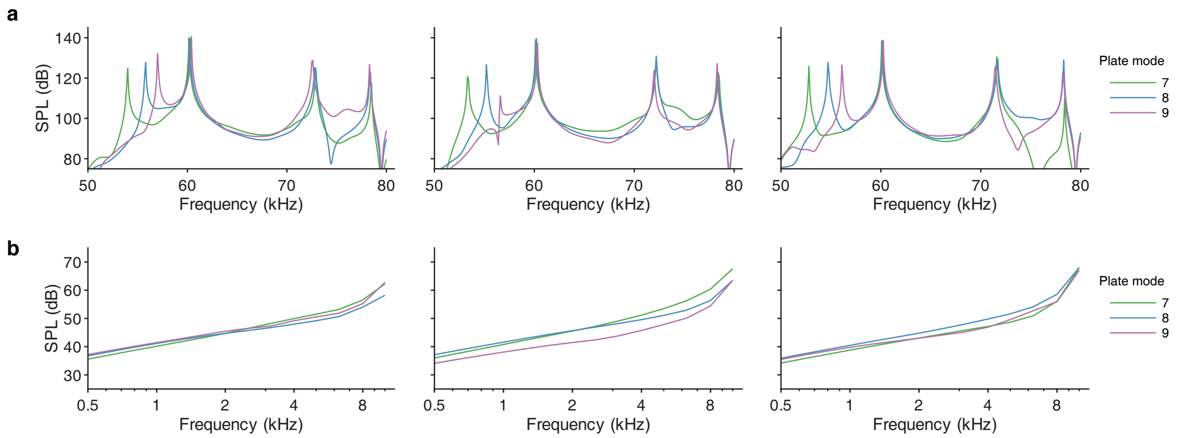


Supplementary Fig. 11 | Propagation curves of the MiPAL across diverse radiator dimensions. **a,** Ultrasound propagation curves for the nine design cases ($\mathscr{D}_{\mathrm{u,c}} \in \{0.40, 0.45, 0.50\}$ m $\times$ $m \in \{7, 8, 9\}$). Vertical dash-dotted lines indicate the predicted critical distance for each case. **b,** Audio sound propagation curves; vertical dash-dotted lines indicate the corresponding audio critical distance ($\mathscr{D}_{\mathrm{a,c}}$) for the demodulated signal.

Response 3. *(continued)*

In Supplementary Note 6

Added:

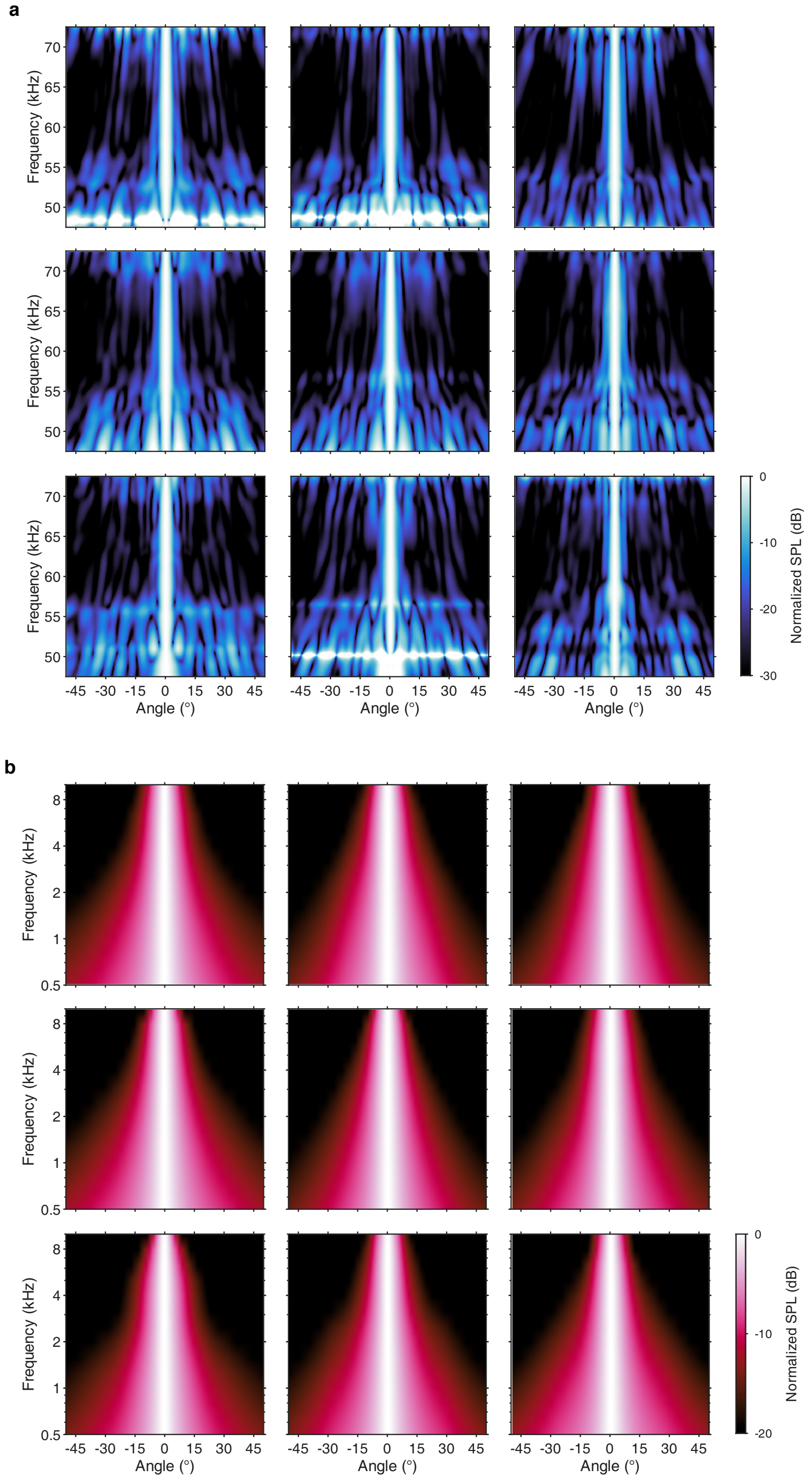


Supplementary Fig. 12 | Frequency responses of the MiPAL across diverse radiator dimensions. **a,** Ultrasound frequency responses evaluated at the ultrasound critical distance $\mathscr{D}_{\mathrm{u,c}}$ of each design case. **b,** Audio sound frequency responses evaluated at the audio critical distance $\mathscr{D}_{\mathrm{a,c}}$ of each design case.

Response 3. *(continued)*

In Supplementary Note 6

Added:



Supplementary Fig. 13 | Directivity patterns of the MiPAL across diverse radiator dimensions. **a,** Ultrasound directivity patterns (top) at the ultrasound critical distance $\mathscr{D}_{\mathrm{u,c}}$, showing consistent wavefront control across geometries. **b,** Audio sound directivity patterns (bottom) at the audio critical distance $\mathscr{D}_{\mathrm{a,c}}$, demonstrating preserved high directivity across the parametric design space.

Comment 4. The difference-frequency wave (DFW) generation mechanism has no essential difference from existing studies. The authors' claim of "systematic innovation via dual-domain integration" (as opposed to previous "local nonlinear effects") completely evades the core issue: the manuscript still lacks substantive physical analysis and experimental validation to prove that this dual-domain integration brings fundamental improvements to DFW generation compared with state-of-the-art technologies.

Response 4. We thank the reviewer for raising this important point. We respectfully note that the main elements of this comment were already addressed in our previous responses: Comment 4 by Reviewer 2 clarified that MiPAL does not claim a new DFW/self-demodulation law, and Comment 5 by Reviewer 2 discussed the appropriate scope of comparison with the cited passive acoustic metamaterial lens. Nevertheless, because the present comment combines these issues with the question of physical analysis and experimental validation, we restate and consolidate our response below in five points.

- **Scope of our claim**

  MiPAL does not claim a new DFW self-demodulation law. We respectfully note that this issue was already addressed in our previous response to Comment 4 by Reviewer 2, where we explicitly stated that the fundamental principle of DFW generation through self-demodulation is a well-established physical phenomenon originating from Westervelt's parametric-array theory. Therefore, we do not claim that MiPAL introduces a new nonlinear self-demodulation mechanism, a new nonlinear propagation law, or a new DFW source term.

  This point is also reflected in the manuscript itself. The DFW process is introduced as the standard PA process in which two ultrasonic primary waves generate an audible difference-frequency component through air nonlinearity, and the Methods section models this process using the quasi-linear Westervelt formulation. In particular, Eq. (3) shows that the generated audio field is driven by the product of the two primary ultrasonic fields, $p_1^* p_2$.

  Thus, the relevant question is not whether MiPAL changes the intrinsic self-demodulation law. It does not, and we have not claimed so. The scope of our claim is device-level: how this established PA mechanism can be realized in a compact, single-body source architecture.

- **Physical role of the dual-domain integration.**

  The dual-domain integration is not presented as a modification of the intrinsic DFW law. Rather, it provides two physically distinct device-level functions required for realizing a clean directional PA audio beam from a compact source platform.

  First, the acoustic metasurface prepares the coherent, high-ka, collimated ultrasonic primary fields that enter the DFW source term. Since the generated DFW field is proportional to $p_1^* p_2$, controlling the phase, coherence, and directivity of the primary ultrasonic fields directly affects the DFW-driving source distribution. This is the role illustrated in Fig. 1f and validated through the ultrasonic field measurements in Fig. 3.

  Second, the elastic meta-units do not increase the intrinsic nonlinear conversion coefficient of air. Their role is to suppress a separate parasitic radiation channel: CR-induced audible leakage caused by lower-order structural modes of the single-body plate. Without this suppression, nearly omnidirectional audible leakage can mask the intended directional DFW beam. This is the role illustrated in Fig. 1g and validated through the vibration and audio-field measurements in Figs. 2, 4 and 5.

  Therefore, if "fundamental improvements to DFW generation" is interpreted as a new nonlinear propagation law, this is not our claim. The contribution of MiPAL is instead a device-level improvement in the realization of directional PA audio, achieved by primary ultrasonic field engineering and CR-induced leakage suppression.

Response 4. (*continued*)

- **Contrast with previous local-nonlinearity-enabled DFW studies**[3,34]**.** We also clarify that our contrast with the cited "local nonlinear effects" studies is not intended to claim novelty in the local-nonlinearity physics itself. In nonlinear acoustics, local nonlinear effects associated with the Lagrangian density are an established modeling regime, commonly distinguished from the cumulative nonlinear effects. Classical nonlinear acoustic theory and PA theory have long distinguished such local terms from propagating cumulative parametric fields[7,8 R9].

  The cited local-nonlinearity-enabled DFW studies[3,34] exploit this local nonlinear effect for localized wave control, such as deep-subwavelength DFW focal spots or self-bending beams[R10]. MiPAL addresses a different problem. It does not claim novelty in local-nonlinearity physics; instead, it uses the established cumulative PA/self-demodulation process and focuses on compact PA-source implementation. In this sense, both the cited local-nonlinearity studies and MiPAL are device-level implementations built on established nonlinear-acoustic frameworks, but they solve different device problems.

- **Physical analysis and experimental validation.**

  Our study does not rely solely on a qualitative architectural description. It provides physical analysis and experimental validation of the two device-level functions described above. The nonlinear audio field is modeled using the quasi-linear Westervelt framework described in Methods and Supplementary Note 8, covering the full chain from the electrical input to the resulting nonlinear acoustic field. To make this validation evidence easy to locate, every simulation–experiment comparison reported in the manuscript and Supplementary Information that supports the device-level functions of the co-integrated acoustic and elastic metamaterials is summarized in Table R2 and Table R3.

- **Comparison with the passive acoustic metamaterial lens**[R11]**.**

  The cited lens performs a passive linear wavefront/aperture transformation of a conventional loudspeaker output, whereas MiPAL is a nonlinear PA source whose audio output is generated from ultrasonic primary fields via DFW generation. The two systems therefore belong to different device categories, and the cited lens is not a direct benchmark for DFW or PA-source architecture. A point-by-point comparison over the overlapping directional band—octave span, Bark-scale perceptual coverage, directivity index, and propagation-direction field integrity—was already provided in our previous response to Comment 5 by Reviewer 2, and we do not repeat it here. The relevant PA-source benchmark is the APAL/SPPAL family, for which a head-to-head comparison under matched transducer and radiating-plate geometry is given in Supplementary Note 16.

In summary, the contribution of MiPAL lies in the device-level realization of directional PA audio: the acoustic metasurface prepares the coherent high-ka primary ultrasonic fields that enter the DFW source term, and the elastic meta-units suppress CR-induced audible leakage that would otherwise obscure the intended directional DFW beam. Therefore, the dual-domain integration does not evade the physical issue; it directly addresses the source-level requirements for realizing a compact directional PA audio system.

Response 4. (*continued*)

Table R2 | Acoustic-field validation across primary ultrasound and demodulated audio bands. Each row reports a simulation–experiment comparison drawn from the manuscript or Supplementary Information. Setup: Supplementary Note 9.

Band	Objective	Figure	Remark
Ultrasound	SPL frequency response	Fig. 3c	over the operating frequency range
	Propagation curve at selected frequencies	Fig. 3b	carrier and sidebands (60.5, 61, 62, 64, 68 kHz); SPL peaks near 0.45 m
	Propagation curve, frequency-resolved	Supplementary Fig. 20	2D (f, z) contour over ultrasound band
	Beam pattern at 0.45 m	Figs. 3d and 3e	2D (f, θ) contour over ultrasound band; w/wo acoustic metasurface; off-axis → collimated 60–70 kHz beam
	Beam patterns at selected frequencies	Supplementary Fig. 22	w/wo acoustic metasurface; carrier and sidebands (60, 60.5, 61, 62, 64, 68 kHz)
Audio sound	SPL frequency response	Fig. 4c	broadband audio over 0.5–10 kHz
	Propagation curve at selected frequencies	Fig. 4b	5 audio freqs (0.5, 1, 2, 4, 8 kHz); SPL peaks near 0.5 m
	Beam pattern at 0.5 m	Fig. 4d	2D (f, θ) contour over ultrasound band; w/wo elastic meta-units
	HPBW frequency response	Fig. 4e	confirms high directivity over operating band
	Beam patterns at selected frequencies	Fig. 4f	5 audio freqs; 2 kHz panel adds w/wo elastic meta-units
	Beam pattern discrepancy map	Supplementary Fig. 21	2D (f, θ) contour over ultrasound band; w/wo elastic meta-units; isolates CR contribution

Table R3 | Vibration validation across the radiating plate and elastic meta-units. Each row reports a simulation–experiment comparison drawn from the manuscript or Supplementary Information. Setup: Supplementary Note 11.

Component	Objective	Figure	Remark
Plate	Velocity frequency response at the plate center driven by the transducer	Fig. 2g, Supplementary Figs. 6d and 14d	w/wo elastic meta-units, which selectively suppress audio-band CR vibration
	Mechanical impedance seen by the plate's center	Supplementary Figs. 6d and 14c	sim.: assembled-transducer baseline; exp.: w/wo 18 attached elastic meta-units
Elastic meta-unit	Velocity transmissibility	Supplementary Fig. 14b	3 groups with different spring lengths l_s; sim.: thick lines ; exp.: thin/light lines

Reviewer 3

Comment 1. I believe the authors have addressed my concerns and I would like to recommend the manuscript for publication.

Response 1. We thank the reviewer for the positive assessment and recommendation for publication. We also appreciate the reviewer's thoughtful comments in the previous round, which helped us improve the clarity and presentation of our work.

Comment 2. My only minor suggestion concerns the revised title. I am not sure if the term "dual-domain metamaterials" accurately captures the integration of the elastic and acoustic design aspects presented in the work. Besides, the metamaterial design itself appears to be the key element enabling directional audio. It may therefore be worth reconsidering whether the phrase "dual-domain metamaterials interwoven with a compact ultrasonic transducer" is suitable, or if a more concise title could better highlight the central contribution.

Response 2. We thank the reviewer for this careful and constructive suggestion. As the reviewer you correctly noted, the metamaterial design itself is the key element that enables directional sound generation in our proposed system.

After carefully reconsidering the title, we have modified the wording to make the central contribution clearer and more technically precise. In particular, we recognized that the term "interwoven" could be read as somewhat metaphorical, whereas "co-integrated" more directly reflects the intended meaning: the acoustic metasurface, elastic meta-units, and compact ultrasonic transducer are functionally integrated at the device level to realize highly directional audio.

We also considered replacing "dual-domain metamaterials" with a more explicit phrase such as "acoustic and elastic metamaterials". While this alternative is descriptive, we found it less concise for the title. We therefore decided to retain "dual-domain metamaterials", where "dual-domain" refers to the simultaneous engineering of the acoustic and elastic wave regimes through the shared transducer architecture. This phrase compactly captures the two metamaterial functions central to the MiPAL architecture: acoustic-domain wavefront shaping and elastic-domain suppression of parasitic vibration modes. We believe this wording appropriately reflects both the central role of the metamaterial design and its co-integration with the compact ultrasonic transducer.

Accordingly, we have changed the title as follows:

Title	
Original: Dual-domain metamaterials interwoven with a compact ultrasonic transducer for highly directional audio generation	**Revised:** Dual-domain metamaterials co-integrated with a compact ultrasonic transducer for highly directional audio generation

We thank the reviewer again for this helpful suggestion, which has improved the clarity and precision of the title.

References

[R1] Shrestha, S., Overvig, A. C., Lu, M., Stein, A. & Yu, N. Broadband achromatic dielectric metalenses. *Light: Science & Applications* **7**, 85 (2018).

[R2] Wang, S. *et al.* A broadband achromatic metalens in the visible. *Nat. Nanotechnol.* **13**, 227–232 (2018).

[R3] Chen, W. T. *et al.* A broadband achromatic metalens for focusing and imaging in the visible. *Nat. Nanotechnol.* **13**, 220–226 (2018).

[R4] Ren, H. *et al.* An achromatic metafiber for focusing and imaging across the entire telecommunication range. *Nature Communications* **13**, 4183 (2022).

[R5] Chen, W. T. *et al.* Dispersion-engineered metasurfaces reaching broadband 90% relative diffraction efficiency. *Nature Communications* **14**, 2544 (2023).

[R6] Chang, S. *et al.* Achromatic metalenses for full visible spectrum with extended group delay control via dispersion-matched layers. *Nature Communications* **15**, 9627 (2024).

[R7] Červenka, M. & Bednařík, M. An algebraic correction for the westervelt equation to account for the local nonlinear effects in parametric acoustic array. *J. Acoust. Soc. Am.* **151**, 4046–4052 (2022).

[R8] Zhong, J. & Qiu, X. *Acoustic Waves Generated by Parametric Array Loudspeakers* 1st edn (CRC Press, 2024).

[R9] Hamilton, M. F. & Morfey, C. L. in *Model equations* (eds Hamilton, M. F. & Blackstock, D. T.) *Nonlinear Acoustics* 39–61 (Springer Nature Switzerland, Cham, 2024).

[R10] Zhang, P. *et al.* Generation of acoustic self-bending and bottle beams by phase engineering. *Nature Communications* **5**, 4316 (2014).

[R11] Peng, X., Hernandez, G., Cummer, S. A., Hoffmann, F.-M. & Li, J. Ultra-broadband acoustic metamaterial lens for loudspeaker directivity control. *Appl. Phys. Lett.* **127** (2025).