In-Sensor Passive Speech Classification with Phononic Metamaterials

Tena Dubček,* Daniel Moreno-Garcia, Thomas Haag, Parisa Omidvar,
Henrik R. Thomsen, Theodor S. Becker, Lars Gebraad, Christoph Bärlocher,
Fredrik Andersson, Sebastian D. Huber, Dirk-Jan van Manen, Luis Guillermo Villanueva,
Johan O.A. Robertsson, and Marc Serra-Garcia*

Mitigating the energy requirements of artificial intelligence requires novel physical substrates for computation. Phononic metamaterials have vanishingly low power dissipation and hence are a prime candidate for green, always-on computers. However, their use in machine learning applications has not been explored due to the complexity of their design process. Current phononic metamaterials are restricted to simple geometries (e.g., periodic and tapered) and hence do not possess sufficient expressivity to encode machine learning tasks. A non-periodic phononic metamaterial, directly from data samples, that can distinguish between pairs of spoken words in the presence of a simple readout nonlinearity is designed and fabricated, hence demonstrating that phononic metamaterials are a viable avenue towards zero-power smart devices.

1. Introduction

The success of deep learning models is based on encoding complex tasks as a combination of large linear transformations and nonlinear activation functions. A variety of technologies, from photonics[1] to memristor crossbar arrays,[2] have been postulated to minimize the energy costs associated with these large linear transformations. Phononic resonators have energy losses that improve on linear passive electronic systems by several orders of magnitude. This is reflected in their quality factors, which quantify the number of periods that oscillations take to decay. Quality factors of several thousands are common in phononic resonators,[3] and can reach billions in specifically optimized devices,[4,5] but are in the tens for electronic circuits.[6] This near-dissipationlessness, combined with the capability of directly processing mechanical signals such as spoken commands without first transducing them into an electronic or photonic domain, makes phononic resonators a prime candidate for zero-power in-sensor edge computing applications. Although these striking advantages have been recognized in the context of both classical[7,8] and quantum[9,10] computing, and the environmental impact of artificial intelligence is increasingly in the spotlight,[11] phononic implementations of machine learning models remain largely unexplored.

A promising application domain for ultra-low-power artificial intelligence is in-sensor computing. Recent advancements in sensor technology allow us to measure a variety of signals, ranging from pressure,[12] magnetic fields,[13] rotations,[14] or even pathogens such as Sudden Acute Respiratory Syndrome

T. Dubček, S. D. Huber
Institute for Theoretical Physics
ETH Zürich
Zürich 8093, Switzerland
E-mail: tena.dubcek@hest.ethz.ch

T. Dubček
Swiss Epilepsy Center
Clinic Lengg
Zürich 8008, Switzerland

T. Dubček
Department of Health Sciences and Technology
ETH Zürich
Zürich 8092, Switzerland

D. Moreno-Garcia, L. G. Villanueva
School of Engineering
EPFL
Lausanne 1015, Switzerland

T. Haag, H. R. Thomsen, T. S. Becker, L. Gebraad, C. Bärlocher,
F. Andersson, D.-J. van Manen, J. O. Robertsson, M. Serra-Garcia
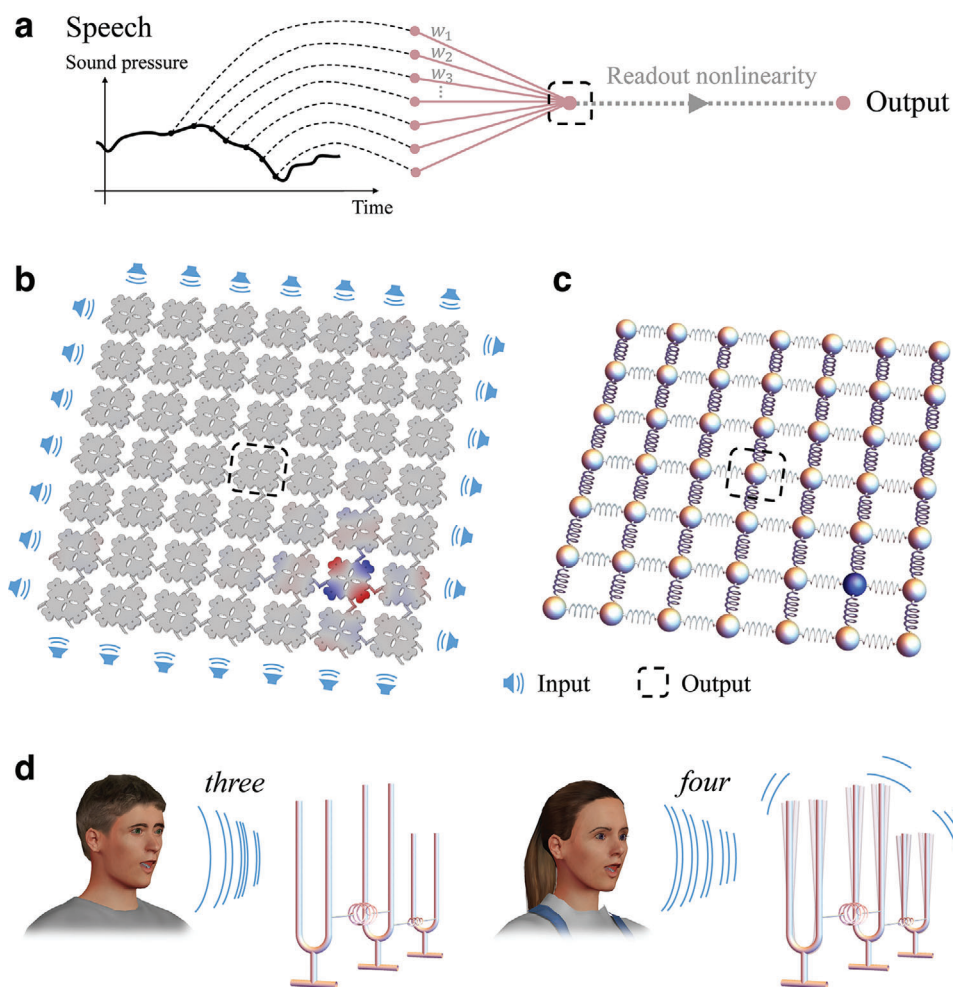Institute for Geophysics
ETH Zürich
Zürich 8092, Switzerland
E-mail: m.serragarcia@amolf.nl

P. Omidvar, M. Serra-Garcia
AMOLF
Science Park 104, Amsterdam 1098 XG, The Netherlands

Figure 1. Passive speech recognition. a) Speech classification by a temporal convolutional network that combines delayed copies of the signal according to a set of weights and then applies a readout nonlinearity. b) We realize a passive instance of such a network by a lattice metamaterial, whose vibrating plates (resonators) are connected by beams. Its geometry (beam locations and hole sizes) is optimized to achieve the desired selective response. c) The structure is modeled as a mass-spring model. Each mass corresponds to a vibration localized at a particular plate. The blue mass corresponds to the displacement represented by the coloring in panel (b). d) The optimized metamaterial can be interpreted as a network of coupled resonators that discriminates between two spoken digits.

CoronaVirus-2[15] through their photonic signatures.[16] However, significant continuous signal processing is still necessary to determine whether a particular event has taken place, even if the event occurs only rarely—a problem known as sparse event detection.[17] In traditional sensing paradigms, information is transferred to a central location, where the measured quantities are analyzed. This results in continuous bandwidth and power consumption, and potential privacy concerns. In-sensor computing[18] is an emergent trend aiming to address these bandwidth, energy consumption, and privacy issues by processing information locally at the sensor, following the emergent trend of edge computing.[19] However, implementing in-sensor signal processing on battery-operated, embedded devices is highly limited by power constraints. This creates a need for low-power or ideally passive, forms of computing. For such tasks, phononic computing is an excellent candidate. While phononic signal processing is significantly slower than electric circuits, a large class of highly relevant signals (e.g., speech commands,[20] bioacoustic

signals,[21] gas concentrations,[22] or intraocular pressure[23]) naturally occur at lower frequencies, and for these in-sensor battery-powered applications, high energetic efficiency is of utmost importance.

However, realizing advanced machine learning functionalities in a phononic device is challenging, as it requires a careful balance between complexity and simplicity. On one hand, the structural design must be expressive enough to encode a complex task such as speech classification (**Figure 1a**); on the other hand, optimizing a mechanical neural network requires simulating a large number of training iterations over a large dataset—hence, the design must be simple enough to be simulated and optimized efficiently. In this work, we demonstrate that phononic metamaterials offer an excellent balance between these two requirements: mode isolation allows for efficient and accurate simulation (with only one degree of freedom per site in the case considered in this work [Figure 1b,c]), while the high sensitivity of metamaterials to the unit cell geometry allows us to cover

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
FUNCTIONAL
MATERIALS**

www.afm-journal.de

a large range of effective properties with a small number of geometric parameters. We illustrate these advantages by designing mechanical metamaterials that perform speech classification tasks, attaining binary-classification accuracies higher than 90% in most tested cases. This capability is experimentally validated by fabricating a metamaterial sample that attains a classification accuracy of 89.6%, close to the simulated value of 91.1%. We then numerically demonstrate that, for words that are not linearly separable, we can achieve good classification performance by constructing deep networks that combine multiple metamaterial elements and commonplace mechanical nonlinearities. Although mechanical metamaterials as a computing platform have gained significant popularity in recent years,[24] for example, in platforms such as buckling elements[25–30] or origami,[31] and wave computing is revolutionizing air acoustics;[8,32] the present paper is, to the best of our knowledge, the first experimental demonstration of a machine-learning task performed by a network of passive phononic resonators—leveraging their unique low-dissipation characteristics.

Speech classification is a widespread application of embedded machine learning, and hence significant efforts have been devoted to minimize its power consumption.[33,34] Therefore, the possibility of passively performing some or all of the associated computations in the elastic domain is highly significant. Electronic approaches to speech classification have traditionally been sum based on convolutional[35] or recurrent[36] architectures (e.g., Long-Short-Term-Memory[37]). Although it is conceivable that both architectures could be realized in mechanical metamaterials—given that there is a known analogy between recurrent networks and wave physics,[8] here we focus on the convolutional approach as it provides a direct interpretation in terms of metamaterial response. In a convolutional neural network, the output is computed by adding together time-shifted copies of the input signal and applying a nonlinear activation function to the resulting signal (Figure 1a). In this work, the phononic metamaterial plays the role of the convolutional filter (encoded in its impulse response) and the nonlinear activation function is given by the measurement of the output energy—as the energy is nonlinearly related to the displacement. The design problem to be solved consists in identifying the metamaterial geometry that encodes a suitable convolutional filter.
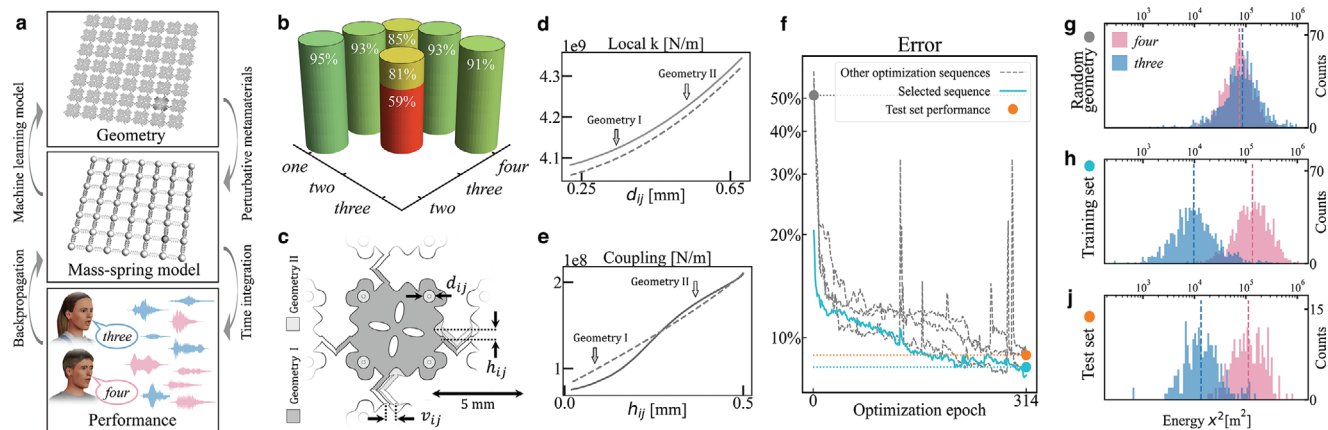
2. Metamaterial Design

We considered a 2D metamaterial consisting of a lattice of 7×7 unit cells. Although the design is based on a repeating unit cell architecture, each site has different geometric parameters (hole radiuses and beam locations). This variability can be understood as a small amount of disorder over a periodic background. Speech signals are applied at the boundary of the metamaterial—by prescribing the vertical displacement of the boundary conditions—and the transmitted energy is measured at the center (output) site (dashed line in [Figure 1b,c]). The choice of output site is arbitrary; once fixed, the optimization algorithm will identify the geometry that maximizes the word classification accuracy for the chosen site. The combination of a metamaterial lattice with an energy measurement can be interpreted as a single-layer neural network. The metamaterial performs the linear transformation, while the energy measurement can be seen as the nonlinear ac-

tivation function—as the energy is proportional to the displacement squared. Intuitively, the task of the metamaterial will be to transmit energy when excited with one word but not another (Figure 1d). Realizing machine-learning models on phononic metamaterials is challenging because machine-learning models tend to be much more complex than traditional periodic lattices: State of the art phononic metamaterials, such as those realizing topological insulators, are described by tight-binding models characterized by a few effective parameters.[3,38–41] In contrast, machine learning models require hundreds to billions of parameters to encode a task.

To bridge this expressibility gap, we devised a multi-step algorithm to efficiently design the sample (**Figure 2a**), which resulted in high classification accuracy (Figure 2b). From the device geometry, we extracted an effective mass-spring model with one degree of freedom per site, following the perturbative metamaterials[42] approach. Perturbative metamaterials implement a Schrieffer–Wolff transformation[43]—a reduction of the lattice dynamics into a low-dimensional, block-diagonal subspace; by projecting the eigenmodes of the metamaterial into a basis of vibrations localized at each site. We then simulated the effective mass-spring model, exciting the sample with utterances of spoken digits from the Google Speech Commands Dataset[44]—composed of recordings from a large and diverse group of speakers under real life conditions. We computed the gradient of the loss function L using backpropagation in time, thus obtaining the gradient of the loss function with respect to the mass-spring values. To obtain the gradient of the classification loss function L with respect to the geometric parameters, we used the chain rule,

$$\frac{\partial L}{\partial (d/h/v)_{ij}} = \sum_{k_{ij-kl}} \frac{\partial L}{\partial k_{ij-kl}} \cdot \frac{\partial k_{ij-kl}}{\partial (d/h/v)_{ij}} \quad (1)$$

where L is the loss function, k_{ij-kl} are the spring constants of the effective mass-spring model connecting site ij with site kl, and d_{ij}, h_{ij}, and v_{ij} are the geometric parameters. The gradient of the mass-spring values with respect to the geometric parameters $\frac{\partial k_{ij-kl}}{\partial (d/h/v)_{ij}}$ was obtained through a surrogate model,[45] a machine learning model that predicts the effective mass and spring constants from the geometry; trained on 5000 full-lattice simulations (see Supporting Information for implementation details). We parameterized the geometry using three geometric parameters per site (Figure 2c): The diameter of holes in the unit cell, d_{ij}, and the horizontal and vertical arm locations h_{ij} and v_{ij}. This choice resulted in a high variability of the effective spring constants, encoded by the fewest possible geometric parameters—achieving the required high expressibility with low model complexity. The effect of each geometric parameter on the mass-spring model can be understood from the unit cell mode shape (Figure 1b). The holes d_{ij} are placed at a modal maximum, and their effect is to increase the effective frequency of the site (Figure 2d) by decreasing the moving mass. The role of the beams is to allow for energy to flow between sites. since in the edge midpoints the local mode has a zero, energy transmission is highly suppressed when the beam is placed in a near-center position ($h_{ij} = 0$ or $v_{ij} = 0$); while the same transmission is enhanced when the beam is placed closer to the maximum. Hence, the position of the beams provides a powerful knob to tune the coupling springs in the

Figure 2. Sample design. a) Training algorithm to determine the metamaterial geometry. b) Speech classification accuracy for all pairs of spoken digits between *one* and *four*. For all but one of the pairs considered, a single layer provides a high classification accuracy. The two–three accuracy can be increased from 59% to 81% with a two-layer network (see generalization section). c) Each unit cell (i, j) contains four holes of equal diameter, d_{ij}. The location of the coupling beams is parameterized by h_{ij} and v_{ij}. d) Local stiffness and e) coupling strength as a function of the hole diameters d_{ij} and beam locations h_{ij}, respectively. The approximation obtained by the machine-learning surrogate model is shown with dashed lines. The coupling is strongly suppressed if the beam is attached where the plate eigenmode has a zero (Figure 1b), making a small beam displacement cause a large shift in the coupling constant. The dark and pale gray arrows denote the corresponding pale and gray configurations in (a). f) Binary classification error rate evolution during training for the *three–four* pair on the training (lines) and test (dots) sets for the selected initial configuration. Training errors for other initial configurations are shown in gray. g) Simulated binary classification performance of a structure before optimization $\{d_{ij}, h_{ij}, v_{ij}\}$ and after optimization for the h) training set (91.8% accuracy) and j) test set (91.1% accuracy).

effective model (Figure 1c), which can be changed by a factor of 3 (Figure 2e). A significant advantage of perturbative metamaterials is that the dependence between effective mass-spring parameters and geometric features is highly local. This allows the surrogate model to precisely predict the local and coupling spring values with a limited number of training samples (Figure 2d,e).

Once the sample design has been parameterized, the design process subsequently consists of identifying the geometry (represented by the values of the parameters d_{ij}, v_{ij}, and h_{ij}) that maximizes the classification accuracy. To train the sample, we defined a sigmoidal loss function of the form

$$L = \sum_{q \notin \text{word}} \left[1 + e^{-\gamma(E_q - E_T)}\right]^{-1} - \sum_{q \in \text{word}} \left[1 + e^{-\gamma(E_q - E_T)}\right]^{-1} \quad (2)$$

where E_q is the energy reaching the central (output) mass (Figure 1c) of the effective mass-spring model when the sample is excited with utterance q, E_T is the threshold energy at which a match is considered to have occurred, and γ is a smoothing parameter to facilitate training. This function captures the training objective to produce a sample that accurately distinguishes between two chosen words. The first trained design consisted of a one-layer model (a metamaterial lattice combined with a square readout function), where the energy E_q is measured at the center site (highlighted with a dotted square in [Figure 1b,c]). Measuring the output energy is equivalent to applying a square activation function as the energy is proportional to the square of the displacement; therefore, this model can be interpreted as a single-layer neural network. Remarkably, such a model attained classification accuracy above 90% for the majority of tested word pairs (Figure 2d).

The design process started with a random configuration of the metamaterial lattice. We then minimized the loss function using

the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm[46] on batches containing the full training dataset—using the gradient computed by the multi-step algorithm described in Figure 2a. The optimization process consisted of 300 iterations and was repeated for 15 different random initial designs. This process is shown in Figure 2f for the *three–four* word pair. Although full-batch BFGS has been associated with overfitting,[47] we observed excellent generalization performance—the degradation was less than 1% between training and test datasets (Figure 2g,h,j). The optimized design that performed best on the training dataset for the *three–four* word pair was selected for fabrication.

3. Experimental Realization

We fabricated the sample (**Figure 3a**) on a 380 μm silicon wafer using standard photolithography and etching techniques (Supporting Information). The equivalence between the full metamaterial and the mass-spring model, provided by the Schrieffer–Wolff transformation, depends on having an isolated phonon band. For materials without local potentials, such as those compatible with our fabrication platform,[3] this requires using a high-order mode (Figure 3c), as the low-frequency spectrum is populated by the three degenerate bands arising from rigid translation modes. To map the broadband speech signal to a high-order band, we modulated the speech on a 10.5 KHz carrier and then increased the playback speed by a factor of 6.8. Such signal transformation would not be necessary for materials with a local support fabricated on multi-layer substrates (see Section 4 and Supporting Information), as the local support can be used to lift the degeneracy between rigid body modes and allows the metamaterial to directly operate on a speech signal.

To impose fixed boundary conditions, the wafer was clamped between two rigid frames and excited uniformly using 28
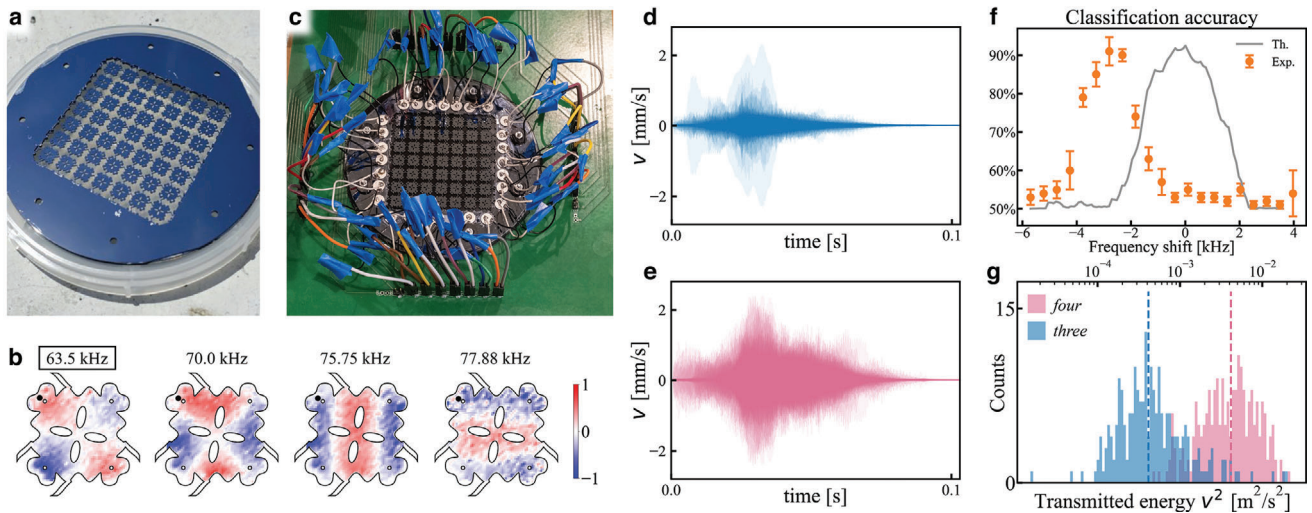
Figure 3. Experimental realization. a) Metamaterial lattice fabricated on a silicon wafer. b) Measured plate vibrations under harmonic excitation at different frequencies. The black dot represents the point where the neural network output is taken. c) Experimental setup (photography by Astrid Robertsson). d) Measurements of the plate vibration at the output point (band-limited to 62.5–74.5 kHz), superimposing the results for the excitation with each of the spoken *three* and e) *four* sound files in the training dataset. The signals corresponding to *three* present a lower vibration amplitude. f) Classification accuracy as a function of modulation frequency. g) Transmitted energy distribution for the test set, calculated from the individual curves in (d,e).

synchronized, thickness-mode piezoelectric actuators (Figure 3b). The large number of actuators allows us to ensure that every boundary site receives a uniform excitation, as these are the conditions that were assumed during optimization. Although samples can be designed to operate under diverse excitation conditions (e.g., with waves applied only at a particular site or boundary), to preserve classification accuracy, experiments must be performed under the same conditions that were assumed during design. We measured the vibration of the output plate using a scanning laser Doppler vibrometer (LDV), band-limited over the range of 62.5–74.5 kHz to minimize the influence of higher-order lattice modes (Figure 3c). The measurements (Figure 3d,e) showed a significantly larger center plate vibration when the lattice was excited by a *four*—even though all excitation signals were normalized to the same mean energy. The optimal classification accuracy was obtained when the modulation frequency was shifted by 2.8 kHz (Figure 3f) with respect to the design value. This deviation can be accounted by the manufacturing tolerance in the thickness of the wafer, which is nominally ±10μm, and can be corrected by combining the theoretical model with physical measurements[48] to trim the sample after fabrication.[49] With the optimal modulation frequency as determined on the training set (Figure 3f), we measured a test-set classification accuracy of 89.6% (Figure 3g), close to the simulated value of 91.1%.

4. Interpretation and Generalization

The full phononic metamaterial is interpreted as a single linear transformation that, when coupled with a nonlinear activation function, implements a layer of a neural network. The action of the metamaterial on the input signal can be understood as a convolution between the speech signal and a kernel encoded in the impulse response of the lattice. Although the lattice contains only nearest-neighbor interactions, the linear transformation effected by the lattice is dense in time, with the weights

for long-range temporal interactions determined by integrating all possible paths that sound waves can take through the lattice with a given signal delay. The effect of the training process is to optimize the weights associated with each delay. Convolution by an impulse response kernel is equivalent to applying a frequency filter with the transfer function, the Fourier transform of the impulse response. This provides a direct interpretation to the classification capabilities of the single lattice. During the design process, the lattice learns to maximize its energy transfer at the frequencies where the difference between words is maximal (**Figure 4a**). The quadratic nonlinearity then rectifies this selectively-transferred signal and computes the mean energy. This mechanism allows the passive metamaterial to distinguish between linearly separable word pairs.

Passive mechanical speech classification can be generalized to word pairs with similar mean spectral contents by assembling deep networks interconnected by nonlinear elements (Figure 4c–e). These nonlinear elements allow the lattice to distinguish the temporal ordering of different frequency components. We optimize a deep network consisting of two 7 × 7 mass-spring lattices interconnected with the nonlinear mechanical element from ref. [50]. This nonlinear element consists of two strings connected to a cantilever. Due to geometric nonlinearity, the vibration of strings results in a dynamic increase of their tension. This is because vibrating strings have, on average, a longer length than stationary strings. The force exerted by the string on the cantilever has the form $F_{sc} = \gamma x_{s2}^2$, where x_{s2} is the second string displacement at the center and γ is the nonlinear constant. This force causes a deflection of the cantilever (Figure 4c), that is proportional to the squared mean amplitude of the string. In turn, the deflection of the cantilever dynamically alters the tension of the first string, shifting its stiffness by a factor $\Delta k_{s1} = 2\gamma x_c$ where k_{s1} is the elastic constant of the first string, and x_c is the displacement of the cantilever. This change in string stiffness Δk_{s1} causes a corresponding shift in the first string resonance frequency (Figure 4d), which induces a
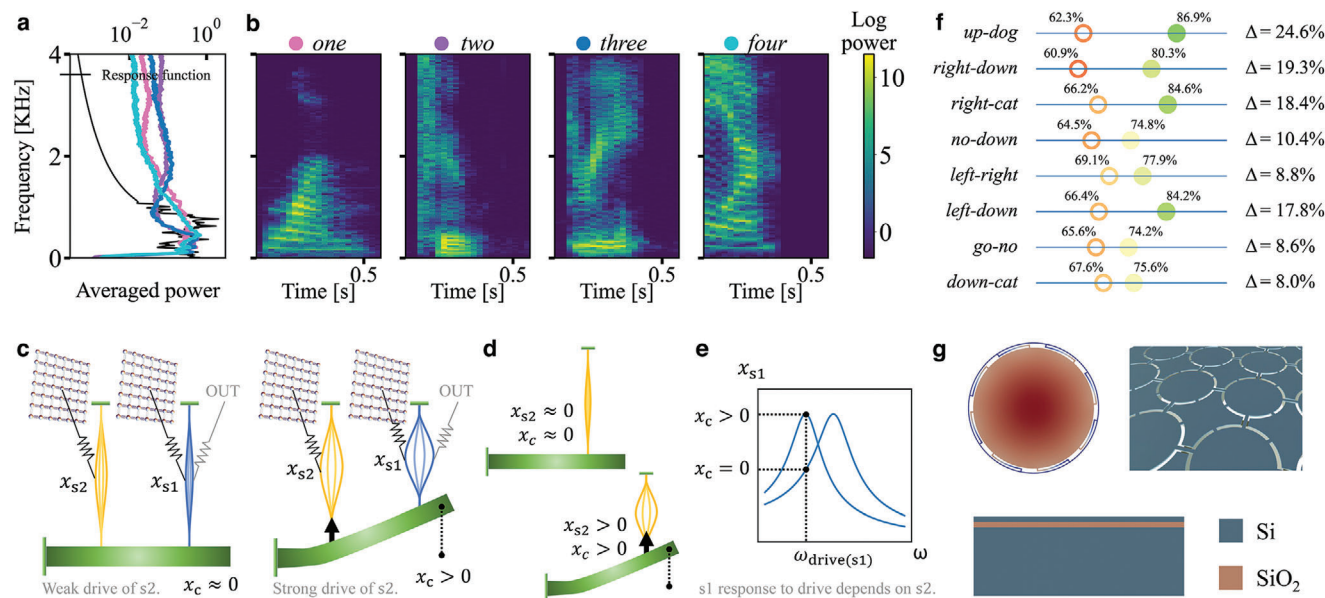
Figure 4. Generalization to other word pairs. a) Mean frequency content of the words *one*, *two*, *three*, and *four* (pink, purple, blue, and magenta, respectively), and transfer function of the linear lattice designed to distinguish between *three* and *four (black)*. Word pairs with more distinct mean frequency contents can be classified more accurately by a single-layer device. b) Example spectrograms for the words *one*, *two*, *three*, and *four*. Word pairs with similar frequency contents can be distinguished from the temporal ordering of the frequency components. This distinction can be mechanically implemented through multi-layer (deep) networks. c) Two layer network implemented by combining two linear transformations interacting through a *mechanical nonlinear activation function*, consisting of two strings (s1, s2) and a cantilever (c), thus realizing an asymmetric quadratic nonlinearity. d) When s2 vibrates with high amplitude, it is on average more curved and hence deflects the cantilever c due to its finite stretching compliance (force denoted by a thick black arrow). e) The time-dependent position of the cantilever c then influences the tension of the string s1, shifting its resonance curve and altering the final output x_{s1}. f) The string-cantilever-based nonlinearity significantly improves the classification accuracies for all tested word pairs with similar spectral content. g) Finite element method simulation of a silicon drum (left) fabricated on a 220 nm silicon-on-insulator wafer (bottom), capable of operating at audio frequencies with no modulation. A concept for an on-chip lattice is shown on the top right.

gating mechanism for elastic waves. When the string resonance frequency is comparable to the lattice it is connected to, energy can flow through the string and reach the output; in contrast, when the string and lattice frequencies are different, energy flow is stopped. This nonlinear mechanism can be interpreted analogously to a gating mechanism in conventional recurrent speech models.[51] A two-layer model more than halved the classification error, from 41% to 19%, for the word pairs *two–three* (see Supporting Information, for training details). Significant improvements were obtained in all tested word pairs with similar spectral content (Figure 4f).

The theorem by Boyd and Chua,[52] guarantees that mechanical systems can theoretically reach accuracies comparable to those of electronic systems, as any fading-memory function can be realized as a combination of linear transfer functions and static nonlinearities. Speech recognition is by definition fading memory—the result cannot depend on signals that took place before the duration of the detected word; arbitrary linear transfer functions can be engineered by branched delay lines; and arbitrary static nonlinearities can be realized by cascading quadratic elements.

In this work, we have shown that it is possible to encode complex information-processing tasks in phononic metamaterials, by taking advantage of their unique balance between wave control flexibility and design simplicity. Although, due to fabrication limitations, our current prototype operates at higher-than-realtime frequencies; thus requiring additional power to convert and modulate the input signal, it is pos-

sible to build micromechanical resonators operating directly at audio frequencies.[53,54] Figure 4g shows an FEM simulation of a silicon drum whose frequency can be tuned over the entire relevant frequency range of 0.5–20 kHz by changing the geometry of the supporting arms. By demonstrating that machine learning tasks can be encoded in the response of phononic metamaterials, together with prior experimental results on passive amplitude activated switches,[55] we illuminate a novel path toward zero-power smart devices that can intelligently respond to events. This capability is out of reach of conventional electronics. State-of-the-art transistors require more than 10^{-18} J to switch.[56] In contrast, phononic resonators can easily go below 10^{-21} J per period of oscillation.[57] This potential for orders-of-magnitude improvement in energy efficiency had already been recognized in the context of conventional digital computing[7] and can now be applied to machine learning problems.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

mechanical computing, metamaterials, physical computing, smart materials, zero-power computing

[1] F. Ashtiani, A. J. Geers, F. Aflatouni, *Nature* **2022**, *606*, 501.
[2] P. Yao, H. Wu, B. Gao, J. Tang, Q. Zhang, W. Zhang, J. J. Yang, H. Qian, *Nature* **2020**, *577*, 641.
[3] M. Serra-Garcia, V. Peri, R. Süsstrunk, O. R. Bilal, T. Larsen, L. G. Villanueva, S. D. Huber, *Nature* **2018**, *555*, 342.
[4] A. H. Ghadimi, S. A. Fedorov, N. J. Engelsen, M. J. Bereyhi, R. Schilling, D. J. Wilson, T. J. Kippenberg, *Science* **2018**, *360*, 764.
[5] A. Beccari, D. A. Visani, S. A. Fedorov, M. J. Bereyhi, V. Boureau, N. J. Engelsen, T. J. Kippenberg, *Nat. Phys.* **2022**, *18*, 436.
[6] T. Helbig, T. Hofmann, S. Imhof, M. Abdelghany, T. Kiessling, L. W. Molenkamp, C. H. Lee, A. Szameit, M. Greiter, R. Thomale, *Nat. Phys.* **2020**, *16*, 747.
[7] S. C. Masmanidis, R. B. Karabalin, I. D. Vlaminck, G. Borghs, M. R. Freeman, M. L. Roukes, *Science* **2007**, *317*, 780.
[8] T. W. Hughes, I. A. D. Williamson, M. Minkov, S. Fan, *Sci. Adv.* **2019**, *5*, eaay6946.
[9] E. A. Wollack, A. Y. Cleland, R. G. Gruenke, Z. Wang, P. Arrangoiz-Arriola, A. H. Safavi-Naeini, *Nature* **2022**, *604*, 463.
[10] U. von Lüpke, Y. Yang, M. Bild, L. Michaud, M. Fadel, Y. Chu, *Nat. Phys.* **2022**, *18*, 794.
[11] P. Dhar, *Nat. Mach. Intell.* **2020**, *2*, 423.
[12] P. Song, Z. Ma, J. Ma, L. Yang, J. Wei, Y. Zhao, M. Zhang, F. Yang, X. Wang, *Micromachines* **2020**, *11*, 56.
[13] A. L. Herrera-May, L. A. Aguilera-Cortés, P. J. García-Ramírez, E. Manjarrez, *Sensors* **2009**, *9*, 7785.
[14] G. Zhanshe, C. Fucheng, L. Boyu, C. Le, L. Chao, S. Ke, *Microsyst. Technol.* **2015**, *21*, 2053.
[15] Z. Chen, J. Li, T. Li, T. Fan, C. Meng, C. Li, J. Kang, L. Chai, Y. Hao, Y. Tang, O. A. Al-Hartomy, S. Wageh, A. G. Al-Sehemi, Z. Luo, J. Yu, Y. Shao, D. Li, S. Feng, W. J. Liu, Y. He, X. Ma, Z. Xie, H. Zhang, *Natl. Sci. Rev.* **2022**, *9*, nwac104.
[16] Y. Hu, M. Wang, L. Hu, Y. Hu, J. Guo, Z. Xie, S. Wei, Y. Wang, Y. Zi, H. Zhang, Q. Wang, W. Huang, *Chem. Eng. J.* **2022**, *450*, 138228.
[17] M. Gazivoda, V. Bilas, *IEEE Sens. J.* **2022**, *22*, 8313.
[18] F. Zhou, Y. Chai, *Nat. Electron.* **2020**, *3*, 664.
[19] W. Shi, J. Cao, Q. Zhang, Y. Li, L. Xu, *IEEE Internet Things J.* **2016**, *3*, 637.
[20] I. B. Crandall, D. MacKenzie, *Bell Syst. Tech. J.* **1922**, *1*, 116.
[21] V. Lostanlen, A. Bernabeu, J.-L. Béchennec, M. Briday, S. Faucou, M. Lagrange, in *Proc. of the 16th Int. Audio Mostly Conf.*, Association for Computing Machinery, New York, NY **2021**, pp. 147–155.
[22] M. I. A. Asri, M. N. Hasan, M. R. A. Fuaad, Y. M. Yunos, M. S. M. Ali, *IEEE Sens. J.* **2021**, *21*, 18381.
[23] K. Gillmann, G. E. Bravetti, L. J. Niegowski, K. Mansouri, *Expert Rev. Ophthalmol.* **2019**, *14*, 263.
[24] H. Yasuda, P. R. Buskohl, A. Gillman, T. D. Murphey, S. Stepney, R. A. Vaia, J. R. Raney, *Nature* **2021**, *598*, 39.
[25] A. Ion, L. Wall, R. Kovacs, P. Baudisch, in *Proc. of the 2017 CHI Conf. on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY **2017**, pp. 977–988.
[26] Y. Song, R. M. Panas, S. Chizari, L. A. Shaw, J. A. Jackson, J. B. Hopkins, A. J. Pascall, *Nat. Commun.* **2019**, *10*, 882.
[27] U. Waheed, C. Myant, S. Dobson, *Extreme Mech. Lett.* **2020**, *40*, 100865.
[28] Y. Jiang, L. M. Korpas, J. R. Raney, *Nat. Commun.* **2019**, *10*, 128.
[29] L. J. Kwakernaak, M. van Hecke, *Phys. Rev. Lett.* **2023**, *130*, 268204.
[30] T. Mei, C. Q. Chen, *Nat. Commun.* **2023**, *14*, 5204.
[31] Z. Meng, W. Chen, T. Mei, Y. Lai, Y. Li, C. Chen, *Extreme Mech. Lett.* **2021**, *43*, 101180.
[32] J. Weng, Y. Ding, C. Hu, X.-F. Zhu, B. Liang, J. Yang, J. Cheng, *Nat. Commun.* **2020**, *11*, 6309.
[33] W. Shan, M. Yang, J. Xu, Y. Lu, S. Zhang, T. Wang, J. Yang, L. Shi, M. Seok, in *2020 IEEE Int. Solid-State Circuits Conference-(ISSCC)*, IEEE, Piscataway, NJ **2020**, pp. 230–232.
[34] G. Cerutti, L. Cavigelli, R. Andri, M. Magno, E. Farella, L. Benini, *IEEE Trans. Circuits Syst. I* **2022**, *69*, 2002.
[35] O. Abdel-Hamid, A.-r. Mohamed, H. Jiang, L. Deng, G. Penn, D. Yu, *IEEE/ACM Trans. Audio, Speech, Lang. Process.* **2014**, *22*, 1533.
[36] A. Graves, A.-r. Mohamed, G. Hinton, in *2013 IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, IEEE, Piscataway, NJ **2013**, pp. 6645–6649.
[37] F. A. Gers, J. Schmidhuber, F. Cummins, *Neural Comput.* **2000**, *12*, 2451.
[38] Z.-K. Lin, Y. Wu, B. Jiang, Y. Liu, S.-Q. Wu, F. Li, J.-H. Jiang, *Nat. Mater.* **2022**, *21*, 430.
[39] H. He, C. Qiu, L. Ye, X. Cai, X. Fan, M. Ke, F. Zhang, Z. Liu, *Nature* **2018**, *560*, 61.
[40] S. Guddala, F. Komissarenko, S. Kiriushechkina, A. Vakulenko, M. Li, V. M. Menon, A. Alú, A. B. Khanikaev, *Science* **2021**, *374*, 225.
[41] J. del Pino, J. J. Slim, E. Verhagen, *Nature* **2022**, *606*, 82.
[42] K. H. Matlack, M. Serra-Garcia, A. Palermo, S. D. Huber, C. Daraio, *Nat. Mater.* **2018**, *17*, 323.
[43] J. R. Schrieffer, P. A. Wolff, *Phys. Rev.* **1966**, *149*, 491.
[44] P. Warden, *arXiv:1804.03209*, **2017**.
[45] D. A. White, W. J. Arrighi, J. Kudo, S. E. Watts, *Comput. Methods Appl. Mech. Eng.* **2019**, *346*, 1118.
[46] R. Battiti, F. Masulli, in *Int. Neural Network Conference*, Springer Netherlands, Dordrecht **1990**, pp. 757–760.
[47] A. Fedorova, O. Glembek, T. Kinnunen, P. Matějka, in *Sixteenth Annual Conf. of the Int. Speech Communication Association*, ISCA, Baixas, France **2015**, pp. 3036–3040.
[48] L. G. Wright, T. Onodera, M. M. Stein, T. Wang, D. T. Schachter, Z. Hu, P. L. McMahon, *Nature* **2022**, *601*, 549.
[49] W. Hsu, A. R. Brown, in *2007 IEEE Int. Frequency Control Symp. Joint with the 21st European Frequency and Time Forum*, IEEE, Piscataway, NJ **2007**, pp. 1088–1091.
[50] M. Serra-Garcia, A. Foehr, M. Molerón, J. Lydon, C. Chong, C. Daraio, *Phys. Rev. Lett.* **2016**, *117*, 010602.

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
FUNCTIONAL
MATERIALS**

www.afm-journal.de

[51] J. Chung, C. Gulcehre, K. Cho, Y. Bengio, *arXiv:1412.3555*, **2014**.

[52] S. Boyd, L. Chua, *IEEE Trans. Circuits Syst.* **1985**, *32*, 1150.

[53] J. J. Neumann, K. J. Gabriel, *Sens. Actuators, A* **2002**, *95*, 175.

[54] D. Scheibner, J. Mehner, D. Reuter, T. Gessner, W. Dötzel, *Sens. Actuators, A* **2005**, *123–124*, 63.

[55] J. J. Bernstein, M. G. Bancu, E. H. Cook, A. E. Duwel, R. D. Elliott, D. A. Gauthier, S. L. Golmon, J. J. LeBlanc, M. J. Tomaino-Iannucci, J. S. Ung, M. S. Weinberg, *J. Microelectromech. Syst.* **2018**, *27*, 625.

[56] S. Datta, W. Chakraborty, M. Radosavljevic, *Science* **2022**, *378*, 733.

[57] V. Sazonova, Y. Yaish, H. Üstünel, D. Roundy, T. A. Arias, P. L. McEuen, *Nature* **2004**, *431*, 284.

2311877 (8 of 8)