Science Advances

MAAAS

Supplementary Materials for

Wave physics as an analog recurrent neural network

Tyler W. Hughes, Ian A. D. Williamson, Momchil Minkov, Shanhui Fan*

*Corresponding author. Email: shanhui@stanford.edu

This PDF file includes:

Section S1. Derivation of the wave equation update relationship

In the main text, we specified that the dynamics of the scalar field distribution, $u = u(x, y, z, t)$, are governed by the wave equation

$$\frac{\partial^2 u}{\partial t^2} - c^2 \cdot \nabla^2 u = f \tag{S1}$$

where $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplacian operator. $c = c(x, y, z)$ is the spatial distribution of the wave speed and $f = f(x, y, z, t)$ is a source term. As discussed in the main text, nonlinear materials have a wave speed which depends on the wave amplitude. Equation S1 can be discretized in time using centered finite differences with a temporal step size of Δt, after which it becomes

$$\frac{u_{t+1} - 2u_t + u_{t-1}}{\Delta t^2} - c^2 \cdot \nabla^2 u_t = f_t \tag{S2}$$

Here, the subscript t is used to indicate the value of a scalar field at a given time step. To connect Eq. S2 to the RNN update equations from Eq. 1 and 2, we exress this in matrix form as

$$\begin{bmatrix} u_{t+1} \\ u_t \end{bmatrix} = \begin{bmatrix} 2 + \Delta t^2 \cdot c^2 \cdot \nabla^2 & -1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} u_t \\ u_{t-1} \end{bmatrix} + \Delta t^2 \cdot \begin{bmatrix} f_t \\ 0 \end{bmatrix} \tag{S3}$$

Then, the update equation for the wave equation defined by Eq. S3 can be rewritten as

$$\boldsymbol{h}_t = \boldsymbol{A}(\boldsymbol{h}_{t-1}) \cdot \boldsymbol{h}_{t-1} + \boldsymbol{P}^{(i)} \cdot \boldsymbol{x}_t \tag{S4}$$

$$\boldsymbol{y}_t = \left| \boldsymbol{P}^{(o)} \cdot \boldsymbol{h}_t \right|^2 \tag{S5}$$

where we have defined $\boldsymbol{A}$ as the matrix appearing in Eq. (S3). The nonlinear dependence on $\boldsymbol{h}_{t-1}$ is defined by the nonlinear wave speed described above.

An absorbing region is introduced to approximate an open boundary condition (*35*), corresponding to the grey region in Fig. 2B. This region is defined by a damping coefficient, $b(x, y)$, which has a cubic dependence on the distance from the interior boundary of the layer. The scalar wave equation with damping is defined by the inhomogeneous partial differential equation (*36*)

$$\frac{\partial^2 u}{\partial t^2} + 2b \cdot \frac{\partial u}{\partial t} = c^2 \cdot \nabla^2 u + f \tag{S6}$$

where u is the unknown scalar field, b is the damping coefficient. Here, we assume that b can be spatially varying but is frequency-independent. For a time step indexed by t, Eq. S6 is discretized using *centered* finite differences in time to give

$$\frac{u_{t+1} - 2u_t + u_{t-1}}{\Delta t^2} + 2b \frac{u_{t+1} - u_{t-1}}{2\Delta t} = c^2 \nabla^2 u_t + f_t \tag{S7}$$

From Eq. S7, we may form a recurrence relation in terms of u_{t+1}, which leads to the following

update equation

$$\left(\frac{1}{\Delta t^2} + \frac{b}{\Delta t}\right) u_{t+1} - \frac{2}{\Delta t^2} u_t + \left(\frac{1}{\Delta t^2} - \frac{b}{\Delta t}\right) u_{t-1} = c^2 \cdot \nabla^2 u_t + f_t$$

$$\left(\frac{1}{\Delta t^2} + \frac{b}{\Delta t}\right) u_{t+1} = \frac{2}{\Delta t^2} u_t - \left(\frac{1}{\Delta t^2} - \frac{b}{\Delta t}\right) u_{t-1} + c^2 \cdot \nabla^2 u_t + f_t$$

$$u_{t+1} = \left(\frac{1}{\Delta t^2} + \frac{b}{\Delta t}\right)^{-1} \left[\frac{2}{\Delta t^2} u_t - \left(\frac{1}{\Delta t^2} - \frac{b}{\Delta t}\right) u_{t-1} + c^2 \cdot \nabla^2 u_t + f_t\right] \tag{S8}$$

Equation S8 therefore represents the discretized update equation for the scalar wave equation with damping. In matrix form, we may express Eq. S8 as

$$\begin{bmatrix} u_{t+1} \\ u_t \end{bmatrix} = \begin{bmatrix} \frac{2 + \Delta t^2 \cdot c^2 \cdot \nabla^2}{1 + \Delta t \cdot b} & \frac{-1 - \Delta t \cdot b}{1 + \Delta t \cdot b} \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} u_t \\ u_{t-1} \end{bmatrix} + \Delta t^2 \cdot \begin{bmatrix} f_t \\ 0 \end{bmatrix} \tag{S9}$$

which also has the same form as in Eq. 5 and 6 of the main text.

Section S2. Realistic physical platforms and nonlinearities

The results presented in the main text were obtained using a very general physical model for the wave dynamics, where the field is assumed to be scalar and the hidden state to hidden state nonlinearity originates from an intensity-dependent wave speed. In this section, we discuss in more detail how the scalar wave results can be translated into various practical platforms, in both optics and acoustics, using various materials as well as different forms of nonlinearity. The key experimental considerations for a practical realization of the analog wave RNN are (1) achieving a compact (wavelength-scale) physical footprint, (2) a physical medium into which the physical weights can be patterned, and (3) a nonlinear material response which can be achieved at reasonable signal energies, without requiring high power sources. We emphasize that the introduction of a nonlinearity into the wave dynamics is critical to realizing the complex information processing capabilities akin to a conventional RNN, allowing the wave dynamics to extend beyond what is achievable with linear time-invariant system theory.

Optics

The most straightforward realization of the nonlinear wave speed in optical platforms is using Kerr nonlinearity. Silicon (Si) and the family of chalcogenide glasses (e.g. As_2S_3) are two widely used nonlinear optical materials for integrated platforms, with chalcogenide having one of the highest damage thresholds. Such a high threshold allows for processing of sub-picosecond laser pulses with peak powers on the order of 10-50 MW (*37*). A longer pulse duration in such nonlinear materials will lead to irreversible damage and imposes an upper bound on the pulse length which can be processed by the analog RNN. Such an ultra-fast optical analog RNN with Kerr nonlinearity may be useful in a number of scientifically-relevant applications, such as diagnostics and processing for ultrafast pulses in nonlinear spectroscopy and X-ray free electron lasers (*38*). In such applications, the integration of conventional electronic processors is extremely challenging due to the arrival rate of information as well as environmental factors.

Additionally, a potential advantage of sub-ps optical analog RNNs is that the training process for the RNN will not need to access regions of the device parameter space with ultra-narrowband (high-Q) spectral features because the relative signal bandwidths of sub-ps pulses are large. In contrast, an optical analog RNN for processing optical carriers broadened by GHz-rate electro-optic modulators will likely be required to operate in the parameter space associated with sub-GHz, or even sub-MHz, spectral features in order to effectively learn features found in such narrowband signals.

An alternative optical nonlinearity which could be used to construct an analog RNN is saturable absorption. Such a nonlinear response consists of an intensity-dependent absorption/damping, which is mathematically defined as

$$b(u) = \frac{b_0}{1 + \left(\frac{u}{u_{\text{th}}}\right)^2} \tag{S10}$$

where b_0 and u_{th} are the saturable absorption strength and threshold, respectively. An example of the saturable absorption response is plotted in Fig. S1. One potential realization of this effect involves the patterning of graphene or other absorptive 2D materials on top of the linear optical circuit etched into a dielectric such as silicon. An advantage of saturable absorption over the Kerr effect is that the resulting nonlinear response can be observed at input powers on the mW scale (*39*), making saturable absorption a promising candidate for use in optical analog RNNs. On the other hand, the main disadvantage of saturable absorption is that it inevitably introduces power loss into the system, which could potentially degrade the signal to noise ratio (SNR) at the detectors in a large-scale analog RNN. Because the complexity of the analog RNN is directly related to its physical footprint, this may limit the expressive capability of an RNN using this form of nonlinearity. However, if we assume a saturable absorption threshold intensity of 0.5 MW/cm² for graphene (*39*), a structure that is 10λ wide [the y-extent of the structure in Fig. 2(D)] at an operating wavelength of $\lambda = 1550$ nm, and an out-of-plane thickness of 1 μm, this situation would require an input power of 77 mW to achieve the intensity threshold across the input facet of the analog RNN. This nonlinear threshold is orders of magnitude lower than that of the Kerr effect, making saturable absorption an appealing nonlinearity for the analog RNN.

As an example, we numerically demonstrate a version of the wave RNN with saturable absorption. The training results for this system are shown in Fig. S2(A) and (B), where we observe that this form of nonlinearity can also perform well on the vowel classification task from the main text, achieving a training and testing classification accuracy of 95.5% ± 1.4% and 90.3% ± 6.4%, respectively. These accuracies are comparable to those of the analog RNN with a nonlinear wave speed as shown in Fig. 3 of the main text. We note however, that the saturable absorption nonlinearity results in a larger variance in the accuracy on the testing set over the 5 cross validated training runs. The increase in variance could be due to variations in the peak signal amplitudes of the various vowel recordings, even though we normalize all recordings to have equal time-integrated power. Essentially, some vowel samples may not be "loud" enough to overcome the damping and, thus, are nearly completely absorbed before they reach the detectors.

In terms of the device's physical footprint, although we have considered a 2D simulation,
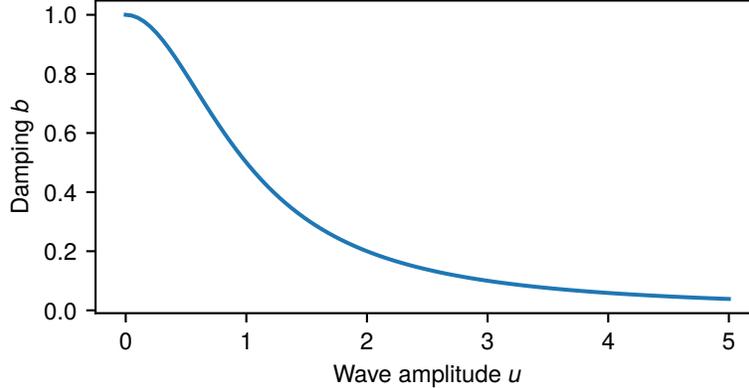
Fig. S1. Saturable absorption response, for the parameters $b_0 = 1.0$ and $u_{\text{th}} = 1.0$ in Eq. S10, indicating the nonlinear dependence of the damping parameter, b on the wave amplitude, u.

the physics observed in such systems translates well to planar integrated optical circuits. In such 3D devices, confinement of light in the out-of-plane direction can be achieved using index confinement. We note that similar planar optical circuits produced via inverse-design techniques have been previously demonstrated experimentally (*24*).

Acoustics

An acoustic or elastic wave implementation of the wave RNN has several advantages over optical platforms in terms of the availability of off-the-shelf components and nonlinear material responses, which can be much stronger. Moreover, the typical operating frequencies used in acoustic signal processing are also orders of magnitude smaller than typical optical carrier frequencies. This, in turn, naturally leads to much larger relative signal bandwidths and eliminates the requirement of realizing ultra-narrowband spectral features in the parameter space of the analog RNN. An acoustic version of the analog RNN could utilize off-the-shelf freespace ultrasonic transducers and receivers in conjunction with a 3D-printed or laser-cut polymer. Many polymers can exhibit a *slower* sound speed than in air (*22*), as is the case in the system demonstrated in the main text. However, we emphasize that a lower sound speed in the printed material is *not* a fundamental requirement for the RNN. In principle, so-called *hard* material sidewalls could also be utilized, similarly to the structures of previously demonstrated acoustic metamaterials (*40*), where $c > 2000$ m/s. It's also worth pointing out that sub-wavelength 3D-printed inclusions could be treated via an effective medium theory in order to achieve a range of sound speeds. This approach would offer interesting opportunities for non-binarized implementations of the analog RNN.

In terms of a nonlinear material response, many fluids, particularly those with embedded gas bubbles (e.g. carbonated water), exhibit a strongly nonlinear response. This effect is captured through a Taylor expansion of the fluid's equation of state, which defines the relationship between pressure, density, and entropy (*20*). A common approach for modeling such effects is using the nonlinear acoustic Westervelt equation (*41*), which includes several terms in addition to the ones
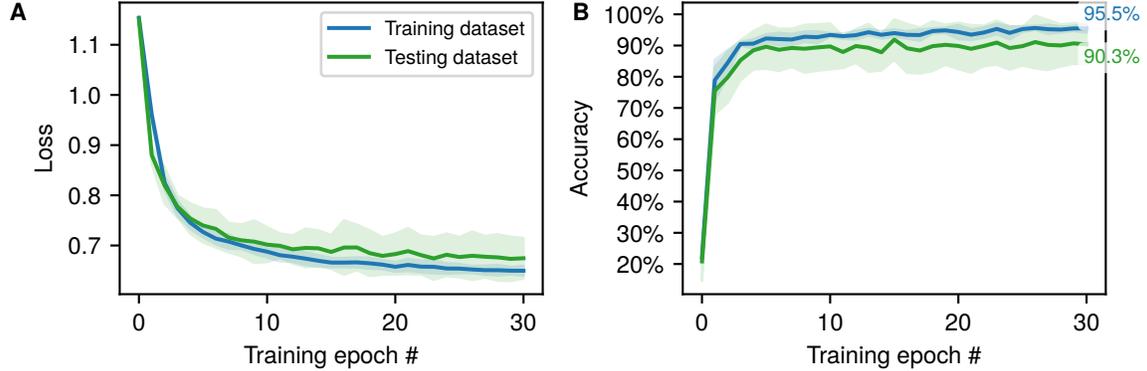
Fig. S2. Cross-validated training results for an RNN with a saturable absorption nonlinearity. The mean (sold line) and standard deviation (shaded region) of the (**A**) cross entropy loss and (**B**) prediction accuracy over 30 training epochs and 5 folds of the dataset, which consists of a total of 279 total vowel samples of male and female speakers. The parameters used for the saturable absorption (Eq. S10) in these results are $b_0 = 0.1$ and $u_{\text{th}} = 2.5 \times 10^{-4}$, while a batch size of 15 samples was used during training.

in the linear wave dynamics defined by Eq. S6. Typically, nonlinear fluids exhibit a second-order nonlinear response (20), where $c(u) \sim u$, which is different from the third-order nonlinearity we use in the main text, but is qualitatively similar. The Westervelt equation also includes a term that accounts for thermoviscous damping, which introduces a frequency-dependent attenuation. One potential route towards including such nonlinearities into an experimental realization of the RNN would be to infiltrate a 3D-printed linear structure with a highly nonlinear fluid. Although, this approach would be different from the system considered in the main text, in terms of which material includes a nonlinear response, it would represent a practical realization of the analog RNN. A non-freespace platform for implementing the RNN would be to use elastic Lamb waves or surface waves on patterned slabs. The features of such systems could be defined lithographically and nonlinearities could again be achieved by infiltrating the patterned linear material with a nonlinear fluid.

In summary, this section shows that there are several realistic pathways for implementing the wave-based RNN described in the main text, including the nonlinearities.

Section S3. Input and output connection matrices

In this section we discuss, in detail, the linear operators, $\boldsymbol{P}^{(\text{i})}$ and $\boldsymbol{P}^{(\text{o})}$, that define the injection and measurement locations within the domain of the wave equation. We start from the vectors $\boldsymbol{u}_t$ and $\boldsymbol{f}_t$, that are discretized and flattened vectors from the field distribution u_t and f_t. Then, we define the linear operators, $\boldsymbol{M}^{(\text{i})}$ and $\boldsymbol{M}^{(\text{o})}$, each column of which define the respective spatial distributions of the injection and measurement points in this flattened basis. With this, we can write the injection of the input vector, $\boldsymbol{x}_t$ as a matrix-vector multiplication

$$\Delta t^2 \boldsymbol{f}_t \equiv \boldsymbol{M}^{(\text{i})} \cdot \boldsymbol{x}_t \tag{S11}$$

Similarly, as the output of the RNN at each time step is given by an intensity measurement of the scalar fields, we may express this in terms of the flattened scalar field as

$$\boldsymbol{y}_t = \boldsymbol{M}^{(\mathrm{o})T} \cdot \boldsymbol{u}_t{}^2 \tag{S12}$$

As the wave equation *hidden state*, $\boldsymbol{h}_t$ is defined as the concatenation of $\boldsymbol{u}_t$ and $\boldsymbol{u}_{t-1}$, we define the following matrices for convenience, as they only act on the $\boldsymbol{u}_t$ portion of $\boldsymbol{h}_t$

$$\boldsymbol{P}^{(\mathrm{i})} \equiv \begin{bmatrix} \boldsymbol{M}^{(\mathrm{i})} \\ \mathbf{o} \end{bmatrix} \tag{S13}$$

$$\boldsymbol{P}^{(\mathrm{o})} \equiv [\boldsymbol{M}^{(\mathrm{o})T}, \ \mathbf{o}] \tag{S14}$$

where $\mathbf{o}$ is a matrix of all zeros. These matrices are used in the injection and measurement stages of the scalar wave update equations of the main text and thus serve a similar role to the $\boldsymbol{W}^{(x)}$ and $\boldsymbol{W}^{(y)}$ matrices of the traditional RNN in Eqs. 1 and 2. However, unlike $\boldsymbol{W}^{(x)}$ and $\boldsymbol{W}^{(y)}$, these matrices are fixed and not trainable parameters. In our numerical implementation, the operation of ∇^2 on a spatially discretized wave field, u_t is computed using the convolution operation, defined mathematically as

$$\nabla^2 u_t = \frac{1}{h^2} \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} * u_t \tag{S15}$$

where h is the step size of the spatial grid.

Section S4. Comparison of wave RNN and conventional RNN

In this section, we compare the performance of the wave RNN to that of a conventional RNN, as defined by Eq. 1 and Eq. 2. In the conventional RNN, the number of trainable parameters is determined by the size of the hidden state, N_h, where the model is parameterized by three matrices $\boldsymbol{W}^{(x)}, \boldsymbol{W}^{(h)}$, and $\boldsymbol{W}^{(y)}$ of size $[N_h \times 1]$, $[N_h \times N_h]$, and $[3 \times N_h]$, respectively. We selected $N_h = 70$ and $N_h = 100$, which correspond to a total number of free parameters in the RNN of 5250 and 10500, respectively. This RNN model was implemented and trained using the `pytorch` framework. In Table S1 we compare the final prediction accuracy of the conventional RNN on the vowel recognition task to the wave RNN, where we observe that the conventional RNN can achieve a performance comparable to that of the wave RNN. However, the conventional RNN is very sensitive to the total number of trainable parameters. For a similar number of trainable parameters to that of the wave RNN, the conventional RNN achieves approximately 6% lower classification accuracy. However, when the number of free parameters is increased to about twice that of the wave RNN, the accuracy is higher by approximately 3 %. We note that it may be possible to achieve higher accuracy in more advanced recurrent models such as the long short-term memory (LSTM) (*42*) or gated recurrent unit (GRU) (*43*) architectures. However, a detailed exploration of these models is outside the scope of the current study.

The conventional RNN and the wave RNN have a number of qualitative differences which we now discuss in more detail. First, in the conventional RNN, the trainable parameters are given by the elements of the weight matrices. In the case of the wave RNN, we choose to use the wave

Table S1. Comparison of a scalar wave model and a conventional RNN on a vowel recognition task. The saturable damping nonlinearity is described in detail in the next section.

Model	Nonlinearity	# parameters	Accuracy Training	Testing
Wave Equation	linear wave speed	4200	93.1%	86.6%
	nonlinear wave speed	4200	92.6%	86.3%
	saturable damping	4200	95.5%	90.3%
Conventional RNN	linear	5250	78.8%	79.4%
	leaky ReLU	5250	82.6%	80.2%
	linear	10500	88.9%	88.2%
	leaky ReLU	10500	89.4%	89.4%

speed, $c(x, y, z)$, defined on a discretized grid to define the set of trainable parameters, because a specific distribution of c can be physically implemented after the training process. Moreover, while the free parameters of the conventional RNN define a matrix which is multiplied by the input, output, and hidden state vectors, in the wave RNN, the free parameters are multiplied element-wise with the hidden state, which limits the influence of each individual parameter over the full dynamics of information within the hidden state. For a given amount of expressive power, the size of the hidden state in the wave equation must arguably be much larger than that of the conventional RNN. The reason for this is that the amount of information which can be encoded into the spatial distribution of u_t is constrained by the diffraction limit for waves. Thus, it follows that a single element from the hidden state of a conventional RNN may be analogous to several grid cells in the scalar wave equation. Furthermore, the discretized wave equation update matrix, A, is sparse and only contains non-zero values around its main diagonal, which physically corresponds to a neighbor-to-neighbor coupling between spatial grid cells (through the Laplacian operator). Due to this form of coupling, information in a given cell of u_t will take many time steps before interacting with information stored in distant cells, as determined by the wave velocity and the physical distance between them. The presence of this form of causality practically means that one must wait longer for a full 'mixing' of information between cells in the domain, suggesting that in the our numerical simulations, a larger number of time steps may be needed as compared to the typical RNN.

Finally, the form of nonlinearity used in the wave RNN is conceptually distinct from that used in the conventional RNN, which involves the application of the nonlinear function, $\sigma^{(h)}(\cdot)$, as in Eq. 1. In the wave RNN, nonlinearity is provided by making the wave velocity, c, or damping, b, to be dependent on the instantaneous wave intensity u_t^2, i.e. $c = c(u_t^2)$, or $b = b(u_t^2)$. With this addition, the update matrix of Eq. 5, $\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{h}_{t-1})$, becomes a function of the solution at that time step, making the dynamics nonlinear. Nonlinearity is introduced into the output of the wave system ($\boldsymbol{y}_t$) through a measurement the wave intensity, which involves a squaring operation. The practical realization of material nonlinearities is discussed in detail in the supplementary materials section 2.

Section S5. Binarization of the wave speed distribution

In this section we discuss how we create realistic distributions of material with a binarized $c(x, y)$ distribution using filtering and projection schemes during our optimization. Rather than updating the wave speed distribution directly, we instead update a design density $\rho(x, y)$, which describes the *density* of material in each pixel. To create a structure with larger feature sizes, a low pass spatial filter can be applied to $\rho(x, y)$ to created a filtered density

$$\tilde{\rho}(x, y) = \begin{bmatrix} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{bmatrix} * \rho(x, y) \tag{S16}$$

For binarization of the structure, a projection scheme is used to recreate the final wave speed from the filtered density. We define $\bar{\rho}(x, y)$ as the projected density, which is created from $\tilde{\rho}(x, y)$ as

$$\bar{\rho}_i = \frac{\tanh\left(\beta\eta\right) + \tanh\left(\beta\left[\tilde{\rho}_i - \eta\right]\right)}{\tanh\left(\beta\eta\right) + \tanh\left(\beta\left[1 - \eta\right]\right)} \tag{S17}$$

Here, η is a parameter between a value of 0 and a value of 1 that controls the mid-point of the projection, typically 0.5, and β controls the strength of the projection, typically around 100. The distribution of $\bar{\rho}$ varies between 0 and 1. Finally, the wave speed can be determined from $\bar{\rho}$ as

$$c(x, y) = (c_1(x, y) - c_0(x, y))\bar{\rho} + c_0(x, y) \tag{S18}$$

where c_0 and c_1 are the background and optimized material wave speed, respectively.