Multiphysics Modeling of Voice Coil Actuators with Recurrent Neural Network

Tianjian Lu, *Member, IEEE,* Michael Smedegaard, *Member, IEEE,* and Ken Wu, *Member, IEEE*

***Abstract*—In order to accurately model the behaviors of a voice coil actuator (VCA), the three-dimensional (3-D) method is preferred over a lumped model. However, building a 3-D model of a VCA is often very computationally expensive. The computation efficiency can be limited by the spatial discretization, the multiphysics nature, and the nonlinearities of the VCA. In this work, we propose incorporating the recurrent neural network (RNN) into the multiphysics simulation to enhance its computation efficiency. In the proposed approach, the multiphysics problem is first solved with the finite element method (FEM) at full 3-D accuracy within a portion of the required time steps. A RNN is then trained and validated with the obtained transient solutions. Once the training completes, the RNN can make predictions on the transient behaviors of the VCA in the remaining portion of the required time steps. With the proposed approach, it avoids solving the 3-D multiphysics problem at all time steps such that a significant reduction of computation time can be achieved. The training cost of the RNN model can be amortized when a longer duration of transient behaviors is required. A loudspeaker example is used to demonstrate the enhancement of the computation efficiency by using RNN in the multiphysics modeling. Various structures of neural networks and tunable parameters are investigated with the numerical example in order to optimize the performance of the RNN model.**

***Index Terms*—Finite element method, gated recurrent unit, long short-term memory, machine learning, multiphysics, recurrent neural network, voice coil actuator.**

I. Introduction

Voice coil actuators (VCAs) utilize the interaction between the current going through the voice coil and the magnetic field from a permanent magnet to produce force. The working principle of VCAs first appears in loudspeakers and has thereafter been extended to many applications. For example, voice coils are adopted in digital cameras to achieve shorter focusing time [1]. Voice coils are also designed as critical components in positioning systems with precision control in microscopy, lithography, and alignment [2], [3]. There are also consumer electronics that integrate voice coils into the touch input [4], [5]. The prevalence of voice coils leads to an increasing demand on accurately modeling VCAs for both design and optimization purposes. The types of approaches of modeling VCAs falls into two major categories, namely, lumped and three-dimensional (3-D) models. In general, building a lumped model is computationally efficient. One significant drawback of a lumped model is its accuracy [6]. For example, in a lumped model of a loudspeaker where a set of parameters are used to describe the stiffness, mass, and damping, the accuracy of the model can only be retained up to a certain frequency beyond which the motion pattern can no longer be assumed as piston. In order to achieve adequate accuracy, the model has to consider the physical quantities of all three dimensions, known as 3-D models. Owing to its unmatched capabilities in modeling complex geometries and material properties, the finite element method (FEM) [7], [8] has become one of the most popular approaches in 3-D modeling.

Tianjian Lu and Ken Wu are with Google Inc., 1600 Amphitheatre Pkwy, Mountain View, CA 94043, USA, e-mail: {tianjianlu, kenzwu}@google.com.

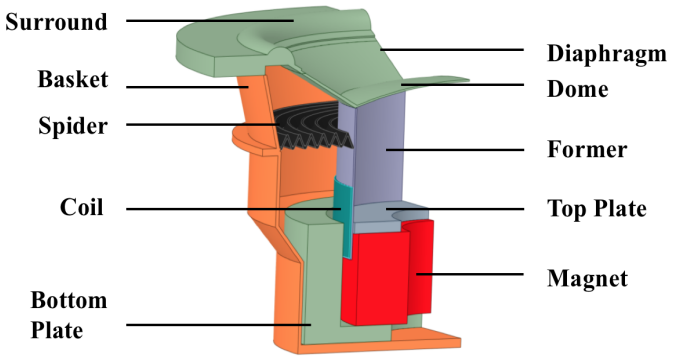

Fig. 1: The parts in a loudspeaker.

It is well known that generating 3-D models can be very computationally expensive. The computation burden primarily results from the spatial discretization and is further exacerbated by the multiphysics nature of and the nonlinearities involved in VCAs. In this work, we describe the formulation of the 3-D FEM-based multiphysics modeling approach through a loudspeaker example. As shown in Fig. 1, the voice coil is connected to the diaphragm through the former. When a time-varying current is applied to the voice coil, the coil moves up and down due to its interactions with the magnetic field of the permanent magnet. Consequently, the former and the diaphragm moves up and down, and the later pushes the air and creates sound. The 3-D FEM-based multiphysics modeling approach consists of two parts: one electromagnetic analysis and one mechanical analysis. The electromagnetic analysis solves for the Lorentz force on the voice coil, which is taken as the excitation in the mechanical analysis. The mechanical analysis solves for the deformation, the velocity, and the acceleration of the diaphragm. The coupling between the electromagnetic and the mechanical analyses is one way through the Lorentz force. At higher amplitude the loudspeaker behaves nonlinearly, generating signal components that do not exist in the input signal [8]. The nonlinearities are caused by the transducer principle and are directly related to, the voice coil, the suspension, and the material properties [9]. Therefore,

the 3-D FEM-based multiphysics modeling approach has to be formulated in the transient regime such that the nonlinearities can be properly addressed. In order to address the nonlinearities, Newton-like iterations are required at every time step in the transient simulation.

There are efforts in developing novel numerical techniques to enhance the computation efficiency while dealing with this multiphysics problem. In previous work, domain decomposition and parallel computing have been incorporated into the multiphysics simulations, which significantly reduces the computation time [10]–[12]. In this work, we take a different route and use machine learning methods, to be specific, the recurrent neural networks (RNNs) [13] to improve the computation efficiency. RNNs have been successfully applied to tasks related to time-sequence modeling such as speech recognition and natural language processing [14], [15]. We incorporate RNN into the multiphysics modeling and use it to predict the transient behaviors of VCAs. Instead of solving the multiphysics problem at all required time steps, we first obtain the results from the 3-D FEM-based multiphysics simulation within only a portion of the required time steps, with which a RNN is trained and validated. Once the training completes, the RNN can start making predictions on the transient behaviors of VCAs in the remaining time steps. There are several advantages of employing RNN in the multiphysics modeling: with RNN, one does not have to solve the nonlinear system associated with the multiphysics problem repeatedly at all time steps, which leads to a significant reduction of computation time; the training cost can be amortized if transient behaviors at additional time steps are in need; the simulation with RNN does not require specific domain knowledge as the predictions are simply achieved through inference. It is worth mentioning that the proposed multiphysics simulation methodology with RNN can be easily mitigated to broader applications related to transient electromechanical analysis [12], [16] beside the voice coil actuators demonstrated in this work.

It is known that a vanilla RNN cannot remember things for very long durations due to the issue of vanishing gradients. One remedy to this problem is to design a more sophisticated activation function with gating units, for example, the long short-term memory (LSTM) unit [17], [18] and the gated recurrent unit (GRU) [19]. Through the loudspeaker example, we compare the vanilla RNN, LSTM network, and GRU network on their performances in the multiphysics modeling. It is found that the vanilla RNN does not achieve adequate accuracy in the multiphysics modeling due to the limitation imposed by the temporal dependency. We also investigate the performance of the LSTM network in predicting transient behaviors of VCAs with respect to various tunable parameters such as the optimization methods and activation functions.

II. 3-D FEM-Based Multiphysics Modeling

In this section, we describe the detailed formulation of the 3-D multiphysics modeling approach based on the finite element method. We also illustrate the causes of nonlinearities in the loudspeaker example shown in Fig. 1 including the nonuniform distribution of magnetic field over the gap and the nonlinear material properties.

A. Electromagnetic Analysis

The governing equation describing the interaction of the current-carrying voice coil with the magnetic field of a permanent magnet is derived from Maxwell's equations

$$\nabla \times \boldsymbol{H} = \boldsymbol{J} \tag{1}$$

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \tag{2}$$

$$\nabla \cdot \boldsymbol{B} = 0, \tag{3}$$

where $\boldsymbol{E}$ is the electric field, $\boldsymbol{H}$ is the magnetic field, $\boldsymbol{B}$ denotes the magnetic flux density, and $\boldsymbol{J}$ represents the conduction current. The derivation of the governing equation also relies on the constitutive relations

$$\boldsymbol{B} = \overleftrightarrow{\mu} \cdot \boldsymbol{H} \tag{4}$$

$$\boldsymbol{J} = \sigma \boldsymbol{E}, \tag{5}$$

where $\overleftrightarrow{\mu}$ and σ are the permeability tensor and the conductivity of the material, respectively. The current density $\boldsymbol{J}$ can be represented by the auxiliary vector potential $\boldsymbol{T}$ as

$$\boldsymbol{J} = \nabla \times \boldsymbol{T}. \tag{6}$$

Consequently, the magnetic field $\boldsymbol{H}$ can be written as

$$\boldsymbol{H} = \boldsymbol{T} + \nabla\Omega + \boldsymbol{H_s}, \tag{7}$$

where Ω is the magnetic scalar potential and the source field $\boldsymbol{H_s}$ is due to known current densities in the conductors. Therefore, the governing equation of the electromagnetic analysis can be derived as

$$\nabla \cdot \left(\overleftrightarrow{\mu} \cdot \nabla\Omega\right) = -\nabla \cdot \left(\overleftrightarrow{\mu} \cdot \boldsymbol{H_s}\right) \tag{8}$$

in the nonconducting regions and

$$\nabla \times \left(\frac{1}{\sigma}\nabla \times \boldsymbol{T}\right) + \overleftrightarrow{\mu} \cdot \frac{\partial}{\partial t}\left(\boldsymbol{T} + \nabla\Omega\right) = -\overleftrightarrow{\mu} \cdot \frac{\partial}{\partial t}\boldsymbol{H_s} \tag{9}$$

in the conducting regions. It can be seen that the $\boldsymbol{T}$-Ω formulation employs vector potential only within conductors.

The Lorentz force acting on the voice coil can be calculated by

$$\boldsymbol{F} = \int_{\Omega_e} \boldsymbol{J} \times \boldsymbol{B} \ \mathrm{dV},$$

where Ω_e represents the region in which the excitation current exists.

B. Mechanical Analysis

The finite element discretization to the transient mechanical analysis leads to the following system

$$[M]\{\ddot{u}\} + [C]\{\dot{u}\} + [K]\{u\} = \{F\}, \tag{10}$$

where $[M]$, $[C]$, and $[K]$ are the mass, damping, and stiffness matrices, respectively, and $\{u\}$, $\{\dot{u}\}$, and $\{\ddot{u}\}$ are vectors containing the displacement, velocity, and acceleration fields, respectively. The temporal discretization follows the Newmark-Beta method [20] in the following form

$$\dot{u}_{n+1} = \dot{u}_n + (1-\gamma)\,\Delta t\ddot{u}_n + \gamma\Delta t\ddot{u}_{n+1} \tag{11}$$

$$u_{n+1} = u_n + \Delta t\dot{u}_n + \left(\frac{1}{2} - \beta\right)\Delta t^2\ddot{u}_n + \beta\Delta t^2\ddot{u}_{n+1}, \tag{12}$$

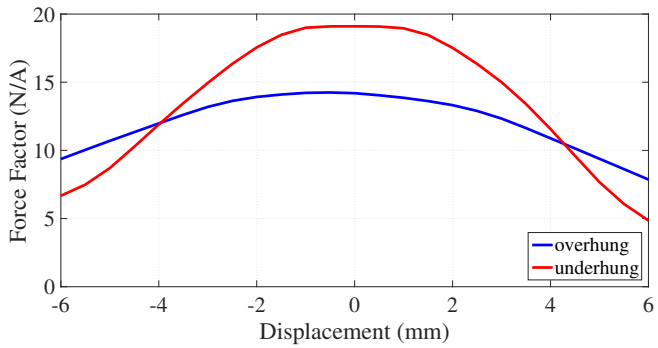

Fig. 2: The force factor of the overhang and underhung coil-gap configuration.

in which $\gamma = 0.5$ and $\beta = 0.25$ are typically used, Δt is the time step, and the subscript n denotes the n^{th} time step. Once the displacement field is obtained, one can use Equations (11) and (12) to update the corresponding velocity and acceleration fields, respectively.

C. Nonlinearities of VCAs

One major nonlinearity of the loudspeaker shown in Fig. 1 arises from the nonuniform magnetic flux density over the gap region. The force factor often used in a lumped model well explains the nonlinearity of this kind. The force factor is calculated by integrating the magnetic flux density over the length of the voice coil. The force factor in both overhung and underhung configurations are depicted in Fig. 2. It can be seen from Fig. 2 that the force factor is displacement-dependent, so as the magnetic flux density in the gap region. The shape of the force factor depends on the geometry of the coil-gap configuration [9]. For example, the force factor in the underhung case where the height of the coil is less than the gap decreases with displacement without a constant region at low amplitude.

Another major nonlinearity of the loudspeaker shown in Fig. 1 results from the stiffness of the suspension system. As shown in Fig. 1, the suspension system consisting of a spider and a surround is able to restore the coil back to the reset position, keeping the coil in the gap. The suspension system acts as a string, restricting the motion along the axial direction by suppressing the rocking modes. However, when the displacement becomes larger than a certain threshold, the stiffness of the suspension becomes displacement-dependent such that the restoring force contains nonlinear terms of the displacement. The stiffness is also frequency-dependent due to the visco-elastic behavior of the suspension system.

In addition to the displacement-dependent magnetic flux density and the stiffness of the suspension, material properties such as the nonlinear relationship between the magnetic field strength and the flux density described in Equation (4) can also cause nonlinearities in the loudspeaker shown in Fig. 1.

III. Recurrent Neural Network

Figure 3(a) shows a unit of a vanilla RNN [13], in which the activation of the hidden state $\{h^t\}$ at time t depends on

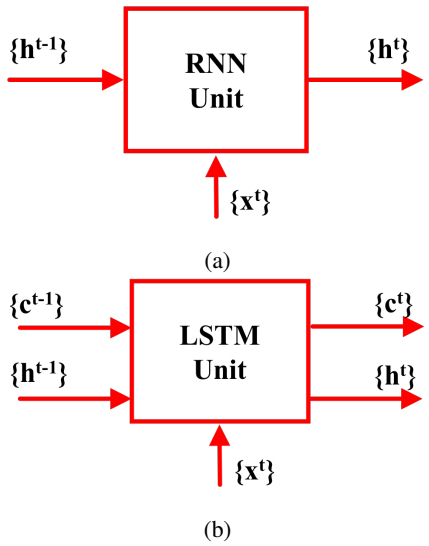

Fig. 3: A unit of (a) a vanilla RNN and (b) a LSTM network.

both the input of the same time step and the hidden state of the previous time step. In a vanilla RNN, the output is usually the same as the hidden state and the update of the later follows

$$\{h^t\} = \phi\left([W]\{x^t\} + [U]\{h^{t-1}\}\right), \tag{13}$$

where ϕ is a smooth, bounded function, for example, a sigmoid function. Unfortunately, it has been found out that as the span of the temporal dependencies increases, the gradient-based optimization becomes increasingly inefficient as the gradients tend to vanish or explode [21]. In order to handle this problem, a more sophisticated activation function with gating units has been developed, for example, the long short-term memory (LSTM) unit as shown in Fig. 3(b). It can be seen from Fig. 3(b) that there is a second state $\{c\}$ named the memory state in additional to the hidden state $\{h\}$. The hidden state $\{h\}$ is a gated version of the memory state $\{c\}$.

There are three gates within a LSTM unit, namely, the input gate $\{i\}$, the forget gate $\{f\}$, and the output gate $\{o\}$, which can be written as

$$\{i^t\} = \sigma\left([W_i]\{x^t\} + [U_i]\{h^{t-1}\}\right) \tag{14}$$
$$\{f^t\} = \sigma\left([W_f]\{x^t\} + [U_f]\{h^{t-1}\}\right) \tag{15}$$
$$\{o^t\} = \sigma\left([W_o]\{x^t\} + [U_o]\{h^{t-1}\}\right), \tag{16}$$

where σ denotes the sigmoid function with the expression

$$\sigma(z) = \frac{1}{1+e^{-z}}. \tag{17}$$

All these three gates make decisions by using the input and the previous hidden state. The input gates operates on the memory state candidate and determines whether the later can be used to update the memory state. The memory state candidate has the expression of

$$\{\tilde{c}^t\} = \tanh\left([W_c]\{x^t\} + [U_c]\{h^{t-1}\}\right), \tag{18}$$

where the hyperbolic tangent function is expressed as

$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}. \tag{19}$$

Similar to the input gate, the forget gate makes an assessment on whether the previous memory state is useful in computing the current memory state. With both the input and the forget gates, the memory state can be updated with the following

$$\{c^t\} = \{f^t\} \circ \{c^{t-1}\} + \{i^t\} \circ \{\tilde{c}^t\}, \tag{20}$$

where $\circ$ represents an element-wise multiplication. Once the memory state is updated, one can write it to the hidden state by using the output gate

$$\{h^t\} = \{o^t\} \circ \tanh\left(\{c^t\}\right). \tag{21}$$

The GRU [19] is a popular variant of the LSTM unit. A GRU keeps the gating functions but merges the memory state into the hidden state. A GRU looks similar to a vanilla RNN unit as shown in Fig. 3(a). Therefore, instead of having three gates as how a LSTM unit is constructed, a GRU has only two gating functions, namely, a reset gate $\{r\}$ and an update gate $\{z\}$, both of which depend on the input and the previous hidden state. The update of the hidden state candidate can be achieved through

$$\left\{\tilde{h}^t\right\} = \tanh\left([W_h]\{x^t\} + [U_h]\left(\{r^t\} \circ \{h^{t-1}\}\right)\right). \tag{22}$$

After that, the hidden state can be updated through

$$\{h^t\} = \left(\{1\} - \{z^t\}\right) \circ \{h^{t-1}\} + \{z^t\} \circ \left\{\tilde{h}^t\right\}. \tag{23}$$

Because of the increased complexity of the LSTM unit and the GRU, it takes longer computation time for the information to propagate through the network comparing to a vanilla RNN. The data for training, validation, and test purposes in building the neural network models are obtained with ANSYS Workbench [22].

IV. Numerical Example

In this section, we use the loudspeaker example to demonstrate the computation efficiency enhancement of the multiphysics simulation through the incorporation of RNN. In the numerical example, a two-tone current is applied to the voice coil of the loudspeaker, one tone at 50 Hz and the other one at 1 kHz. The transient behaviors of the loudspeaker in terms of the current, the induced voltage, and the force on the voice coil, and the deformation, the velocity, and the acceleration of the diaphragm can all be captured with the multiphysics simulation. The multiphysics simulation can also capture the nonlinearities associated with the loudspeaker. It can be seen from Fig. 4(a) that the input current has only nonzero components at 50 Hz and 1 kHz. Because of the nonlinearities, both harmonic and intermodulation distortions take place, which is demonstrated through the spectrum of the induced voltage and the Lorentz force shown in Fig. 4(b) and (c), respectively. Figure 5 depicts the distribution of the acceleration field captured by the FEM-based multiphysics simulation.

In this example, the total number of time steps is 20,000 as required. The multiphysics simulation is performed on the first 14,000 time steps for training and validation purpose. When the training completes, the LSTM network takes over the multiphysics simulation and predicts the transient behaviors of the loudspeaker in the remaining 6,000 time steps. For the electromagnetic analysis based on FEM, it takes 21 iterations within each time step to achieve the convergence in solving the nonlinear problem where the residual is set to be 10^{-4}. It takes 34 seconds for the electromagnetic analysis at each time step to complete solving the nonlinear system with 50,602 degrees of freedom (DOFs). The computation is performed with Intel Xeon CPU E5-2680 v2 @2.8GHz.

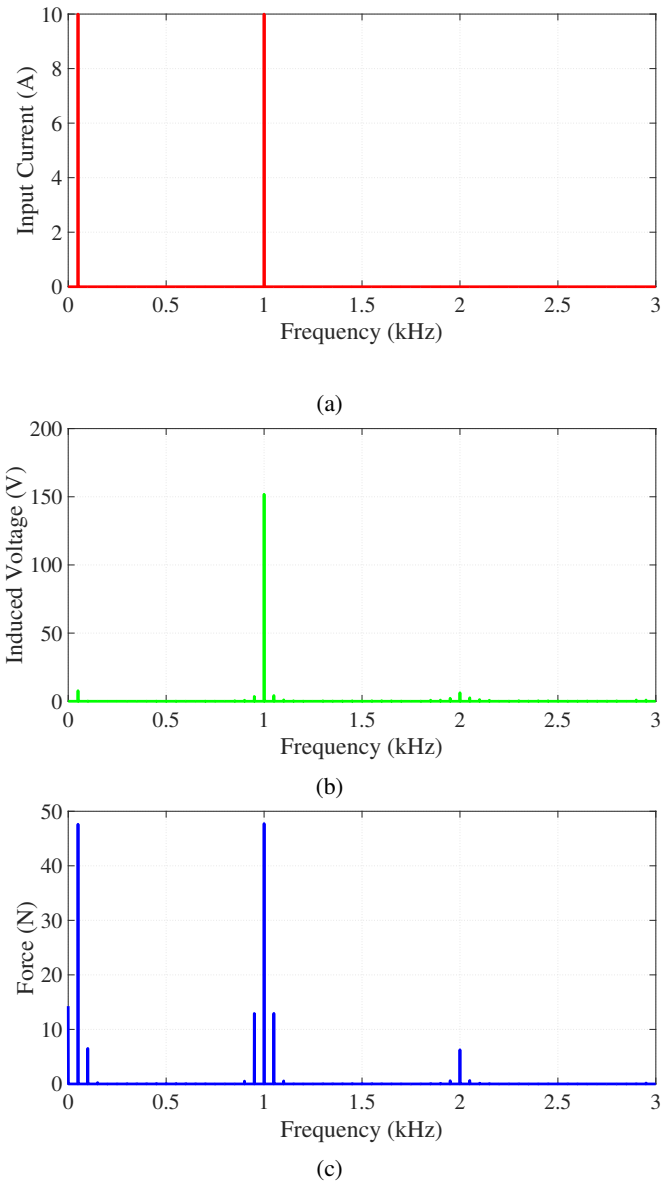

Fig. 4: The spectrum of (a) the input current, (b) the induced voltage, and (c) the Lorentz force on the voice coil of the loudspeaker.

There are four hidden layers in the LSTM network and each layer has the width of 20, 30, 10, and 10, respectively. The training takes 5 epochs and 61 seconds in total for the residual to drop below 10^{-4}. The training cost is shared by the 6,000 time steps in the test set. It takes 0.003 seconds for the LSTM network to make the prediction at one time step

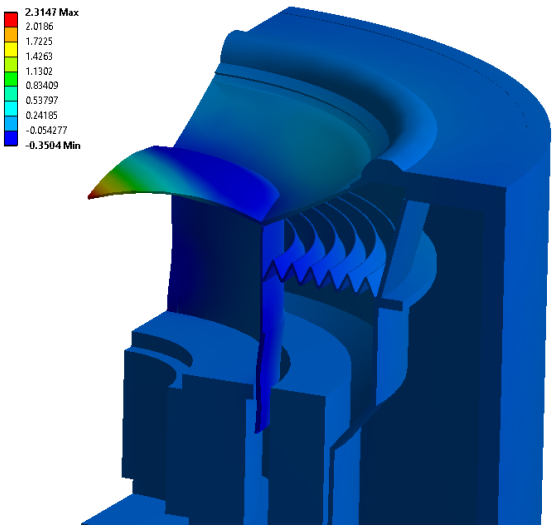

Fig. 5: Distribution of the acceleration field ($\mathrm{m/s^2}$) of the loudspeaker captured by the FEM-based multiphysics simulation.

through inference. Considering the cost of both training and inference, it takes about 0.02 seconds for the LSTM network to produce results at one time step, which is a lot shorter than the 34 seconds with the FEM-based electromagnetic analysis in the 3-D simulation.

The prediction by the LSTM network achieves good accuracy. A portion of the predicted waveforms for the input current, the induced voltage, and the Lorentz force on the voice coil is provided in Fig. 6(a), (b), and (c), respectively. The predictions in terms of the mechanical behaviors including the deformation, the velocity, and the acceleration on the diaphragm are depicted in Fig. 7(a), (b), and (c), respectively. It can be seen that the predictions agree very well with the test set. For example, the root-mean-square error (RMSE) is 32 nm in terms of the predicted deformation through the LSTM network whereas the actual range of the deformation is from -701 to 660 nm. The RMSE is approximately 2.4% of the range of the deformation in the test set.

The comparison among three different structures of neural networks in terms of the prediction accuracy is provided in Table I. The comparison is performed based on the RMSE of the predicted deformation of the loudspeaker diaphragm. It is worth mentioning that the length of the input sequence, denoted by n in Table I, measures the number of time steps used in predicting the behavior at the future time step. As shown in Table I, when the length of the input sequence increases from two to ten, the prediction accuracy of the vanilla RNN improves significantly. It is understood that when more prior data are included, the temporal dependency is reduced such that the vanilla RNN achieves better prediction accuracy. It can also be seen in Table I that LSTM and GRU networks achieve better prediction accuracy than the vanilla RNN. The prediction accuracy of LSTM and GRU networks does not vary significantly with respect to the length of input sequence because the sophisticated activation with gating units handles the temporal dependency very well.

Through the numerical example, we investigate the performance of the neural network models with various optimization methods. With the fixed maximum number of epochs, the convergence speeds achieved by the different optimization methods often translate to different prediction accuracies. We perform the numerical experiment on three variants of the stochastic gradient descent (SGD) method, namely, the plain SGD (`GradientDescentOptimizer` in TensorFlow), the Adagrad variant (`AdagradOptimizer` in TensorFlow), and the Adam variant (`AdamOptimizer` in TensorFlow) [23]. As can be seen from Fig. 8, Adagrad achieves significantly better convergence speed than SGD and Adam while predicting deformations. The convergence speed is measured by the mean-square error (MSE) with respect to the number of epochs. Note that the data preprocessing is applied prior to the training.

Activation function is another important tunable parameter of neural networks. In a LSTM network, the activation

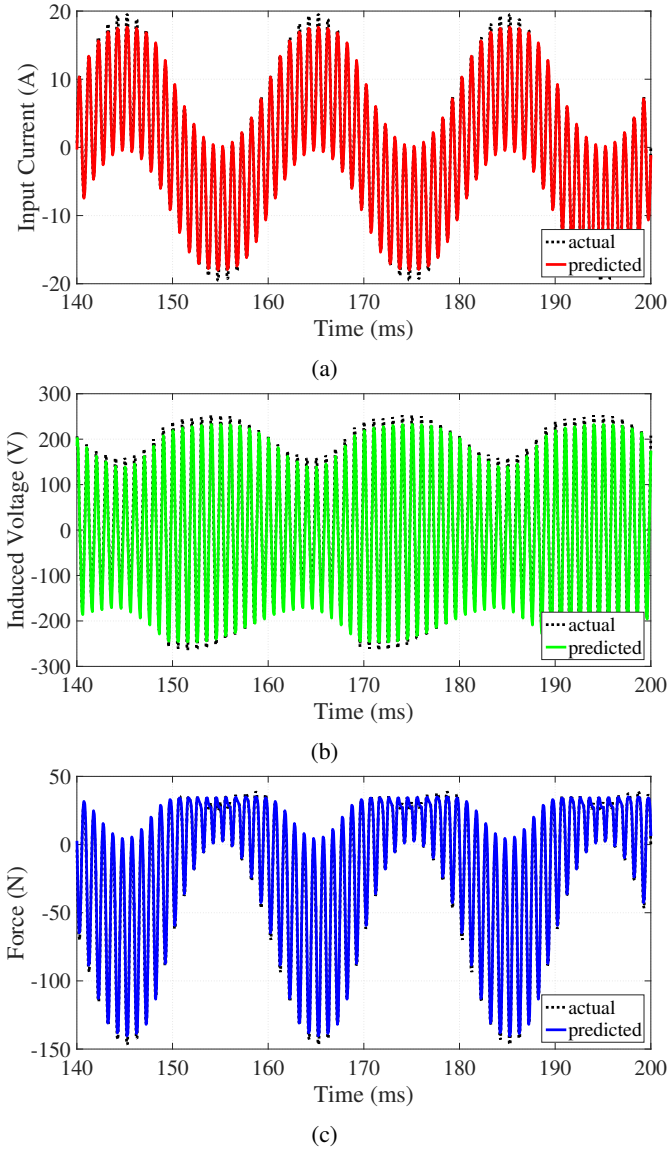

Fig. 6: The predicted (a) input current, (b) induced voltage, and (c) Lorentz force on the voice coil of the loudspeaker. The prediction is made through a LSTM network on the test set.

TABLE I: The RMSE of the predicted deformation on the diaphragm with vanilla RNN, GRU network, and LSTM network with varying lengths of input sequence. The length of input sequence, denoted by n, measures the number of time steps used in predicting the behavior at the future time step.

	n = 2	n = 10	n = 50
RNN	146 nm	78 nm	76 nm
GRU	43 nm	33 nm	31 nm
LSTM	41 nm	35 nm	32 nm

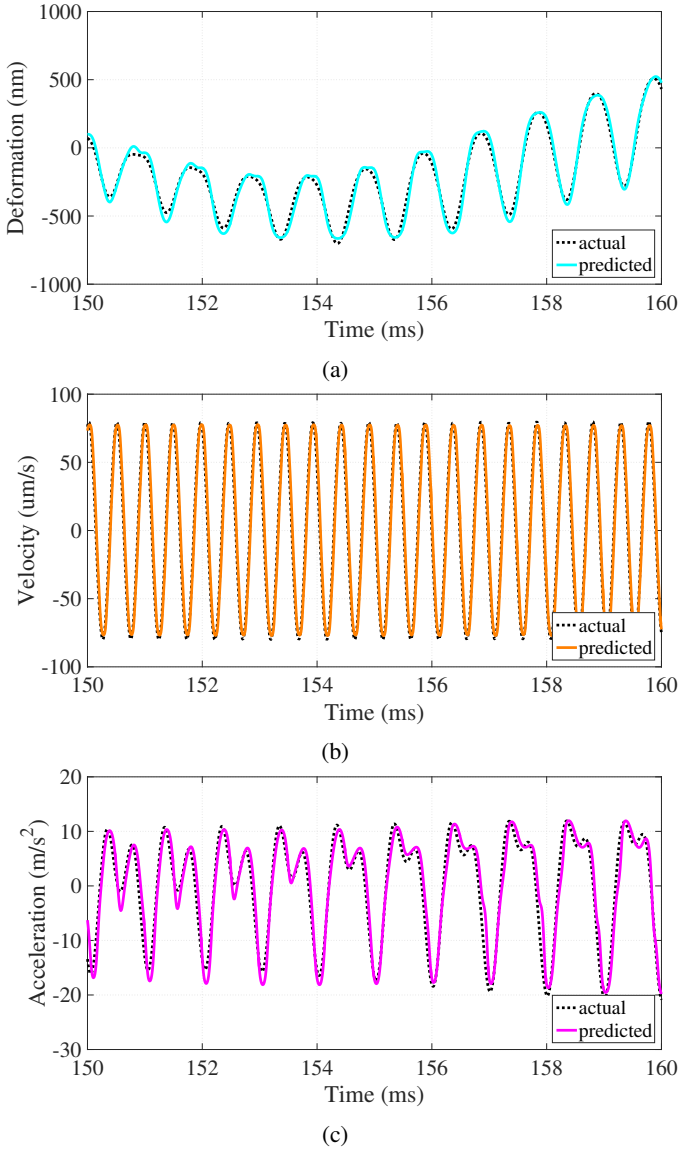

Fig. 7: The predicted (a) deformation, (b) velocity, and (c) acceleration field of the diaphragm in the loudspeaker. The prediction is made through a LSTM network on the test set.

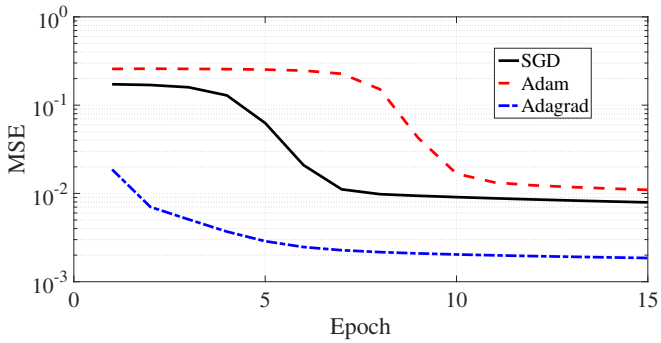

Fig. 8: The convergence speeds achieved by three different optimization methods, namely, SGD, Adam, and Adagrad. The convergence speed is measured by the mean-square error of deformations with respect to the number of epochs.

functions associated with the input, the forget, and the output gates are fixed as sigmoid functions. It is because the sigmoid function squashes the signal into the range from 0 to 1, with 0 denoting no flow and 1 representing complete flow of the information through the gate. The hyperbolic tangent function cannot be used in this case as it squashes the signal into the range from -1 to 1. Similarly, the rectified linear unit (ReLU) cannot be used for the gates as it may amplify the signal when it is passing through the gate. The selection of the activation function is further verified by the results shown in Fig. 9. When the hyperbolic tangent function or ReLU is used for the recurrent activation, the training process does not converge at all.

The activation function applied to the memory candidate and hidden states has to be hyperbolic tangent function in this example as all the physical quantities describing the loudspeaker behaviors have both negative and positive values. It shows in Fig. 10 that when ReLU is used as the activation function on both the memory candidate and the hidden states, only the positive values of deformations remain after the prediction. It can be seen from the numerical results in Figs. 9 and 10 that selecting the appropriate activation functions is critical in making meaningful and accurate predictions for the loudspeaker behaviors with neural networks.

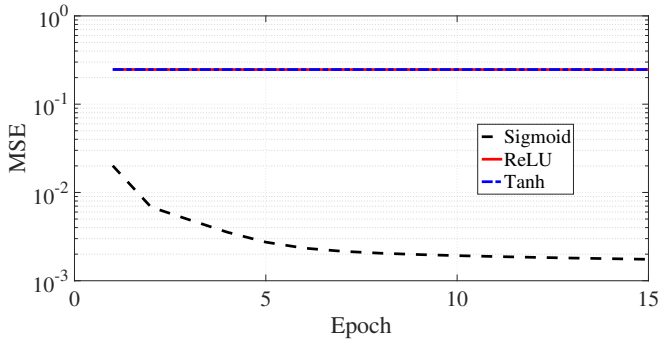

Fig. 9: The convergence speeds achieved by three activation functions for the gates in LSTM, namely, the sigmoid function, the hyperbolic tangent function, and the rectified linear unit. The convergence speed is measured by the mean-square error of deformations with respect to the number of epochs.

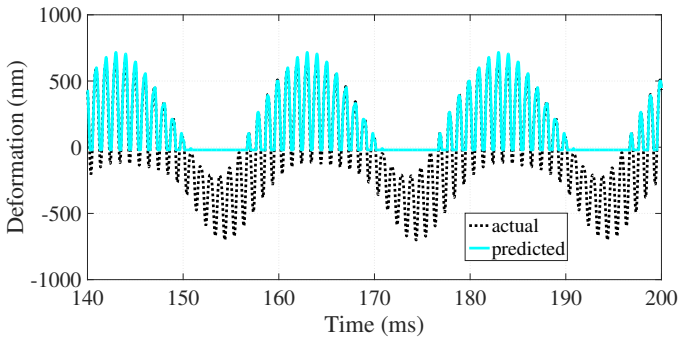

Fig. 10: The predicted deformation on the diaphragm of a loudspeaker with the LSTM network where the rectified linear unit is used as the activation function for both the memory candidate and the hidden states.

V. Conclusion and Discussion

In this work, we describe the 3-D FEM-based multiphysics modeling approach for VCAs. The FEM-based multiphysics simulation integrates an electromagnetic analysis and a mechanical analysis, in which the coupling is one-way through the Lorentz force on the voice coil. The FEM-based multiphysics simulation is formulated in the transient regime such that it can address the nonlinearities arising from both the geometry and the material properties with adequate accuracy. Through the FEM-based multiphysics simulation, one can obtain the current, the voltage, and the force on the voice coil as well as the deformation, the velocity, and the acceleration on the diaphragm.

In order to improve the computation efficiency, we incorporate RNNs into the multiphysics simulation. In the proposed approach, we first apply the FEM-based multiphysics simulation to obtain the transient behaviors of VCAs on a portion of the required time sequence. The obtained data are used to train and validate the RNNs. Once the training completes, the RNNs can make predictions on the transient behaviors for the rest of the required time sequence. This RNN-based approach requires no substantial domain knowledge and at the same time saves complex and expensive 3-D simulations. Through the loudspeaker example, it demonstrates that RNNs can make accurate predictions on the transient behaviors of VCAs in a highly efficient manner. Three types of RNNs, to be specific, the vanilla RNN, LSTM network, and GRU network are compared based on their performance in modeling VCAs. Due to the limitation imposed by the temporal dependency, the vanilla RNN cannot achieve adequate prediction accuracy in the loudspeaker example, wheres both the LSTM and GRU networks provide pretty accurate predictions. Various tunable parameters including optimization methods and activation functions associated with the LSTM network are investigated through the loudspeaker example. It is critical to select the appropriate activation functions in order for RNNs to make meaningful and accurate predictions on the transient behaviors of VCAs.

References

[1] H.-C. Yu, T.-Y. Lee, S.-J. Wang, M. L. Lai, J.-J. Ju, D. R. Huang, and S.-K. Lin, "Design of a voice coil motor used in the focusing system of a digital video camera," *IEEE Transactions on Magnetics*, vol. 41, no. 10, pp. 3979–3981, Oct 2005.

[2] L. Wang, K. Maslov, J. Yao, B. Rao, and L. V. Wang, "Fast voice-coil scanning optical-resolution photoacoustic microscopy," *Optics letters*, vol. 36, no. 2, pp. 139–141, 2011.

[3] Q. Xu, "Design and development of a compact flexure-based xy precision positioning system with centimeter range," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 2, pp. 893–903, 2014.

[4] E. C. Chubb, J. E. Colgate, and M. A. Peshkin, "Shiverpad: A glass haptic surface that produces shear force on a bare finger," *IEEE Transactions on Haptics*, vol. 3, no. 3, pp. 189–198, 2010.

[5] S. Choi and K. J. Kuchenbecker, "Vibrotactile display: Perception, technology, and applications," *Proceedings of the IEEE*, vol. 101, no. 9, pp. 2093–2104, 2013.

[6] W. Klippel, "Diagnosis and remedy of nonlinearities in electrodynamical transducers," in *Audio Engineering Society Convention 109*. Audio Engineering Society, 2000.

[7] J.-M. Jin, *The finite element method in electromagnetics*. John Wiley & Sons, 2015.

[8] T. Tsuchiya, Y. Kagawa, M. Doi, and T. Tsuji, "Finite element simulation of non-linear acoustic generation in a horn loudspeaker," *Journal of sound and vibration*, vol. 266, no. 5, pp. 993–1008, 2003.

[9] W. Klippel, "Loudspeaker nonlinearities–causes, parameters, symptoms," in *Audio Engineering Society Convention 119*. Audio Engineering Society, 2005.

[10] T. Lu, F. Zhang, and J.-M. Jin, "Multiphysics simulation of 3-D ICs with integrated microchannel cooling," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 6, no. 11, pp. 1620–1629, 2016.

[11] T. Lu and J.-M. Jin, "Electrical-thermal co-simulation for analysis of high-power RF/microwave components," *IEEE Transactions on Electromagnetic Compatibility*, vol. 59, no. 1, pp. 93–102, 2017.

[12] ——, "Coupled electrical–thermal–mechanical simulation for the reliability analysis of large-scale 3-D interconnects," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 7, no. 2, pp. 229–237, 2017.

[13] P. J. Werbos, "Backpropagation through time: what it does and how to do it," *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990.

[14] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.

[15] A. Graves, A.-r. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Acoustics, speech and signal processing (icassp), 2013 ieee international conference on*. IEEE, 2013, pp. 6645–6649.

[16] H. T. Chorsi, M. T. Chorsi, and S. D. Gedney, "A conceptual study of microelectromechanical disk resonators," *IEEE Journal on Multiscale and Multiphysics Computational Techniques*, vol. 2, pp. 29–37, 2017.

[17] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[18] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," 1999.

[19] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.
[20] N. M. Newmark, "A method of computation for structural dynamics," *Journal of the engineering mechanics division*, vol. 85, no. 3, pp. 67–94, 1959.
[21] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 157–166, 1994.
[22] *ANSYS workbench help system*, ANSYS, Inc., Canonsburg, PA, 2018, release 19.1.
[23] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016.

Tianjian Lu (S13-M17) received the B.E. degree in electrical engineering from the National University of Singapore, Singapore, in 2010, the M.S. and Ph.D degrees in electrical engineering from the University of Illinois at Urbana-Champaign, in 2012 and 2016, respectively. Since 2016, he has been a hardware engineer with Google. His research interests include multiphysics modeling and simulation, signal and power integrity, and machine learning. He was a recipient of the Best Student Paper Award (The First Place Winner) at the 31th International Review of Progress in ACES, Williamsburg, VA, USA, in 2015 and the Best Student Paper Award at the IEEE Electrical Design of Advanced Packaging and Systems (EDAPS), Honolulu, HI USA, in 2016. He was also a recipient of the P. D. Coleman Outstanding Research Award by the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, in 2016.

Ken Wu received the B.S. degree from the University of Science and Technology of China in 1989 and the M.S. degree from the Institute of Geophysics, Chinese Academy of Science, in 1992, both in geophysics. He received his Ph.D from the Department of Electrical Engineering at the State University of New York at Binghamton in 1996. He has been working at Google Consumer Hardware Division for over five years. He specializes in signal and power integrity modeling, simulation and lab validation from PCB system to package and die levels. He is one of the early contributors to USB TypeC connector electrical spec. Previously he worked in Apple, Broadcom, Force10 Networks (part of Dell now), 3dfx (part of NVIDIA now), and LSI Logic.