# Locally coupled oscillatory recurrent networks learn traveling waves and topographic organization

**Abstract:** Complex spatio-temporal neural population dynamics such as traveling waves are known to exist across multiple brain regions (Lubenov & Siapas, 2009; Muller et al., 2014), and have been hypothesized to play diverse roles from information transfer (Sato et al., 2012) to long-term memory consolidation (Muller et al., 2016). To-date, however, the empirical validation of these computational hypotheses has been hindered by the lack of a flexible and efficiently trainable model of such behavior. In this work, we introduce the Locally Coupled Oscillatory Recurrent Neural Network (LocoRNN), and show that it indeed learns to leverage traveling waves, and other well known coordinated dynamics of coupled oscillators (Kuramoto, 1975), in the service of structured sequence modeling. However, unlike previous models of such dynamics, we show that our model remains a flexible, trainable, sequence model competitive with state of the art on benchmarks such as the Hamiltonian Dynamics Forecasting Suite (Botev et al., 2021). Furthermore, when trained to model simple image sequences such as simulated retinal waves, we see that the orientation selectivity of hidden neurons becomes topographically organized, while such organization is absent when trained on unstructured noise. The resulting organization is reminiscent of orientation columns observed in the visual cortex and in line with prior work on activity-dependent organization in the visual system during development (Ackman et al, 2012). Due to local connectivity, our model is both more biologically plausible and parameter efficient than its globally coupled counterpart, the coRNN (Rusch & Mishra, 2021), while also being substantially more amenable to gradient-based training than recent spiking neural network counterparts (Davis et al, 2021) due to provably bounded gradients. Overall, we believe our results highlight the value of the LocoRNN as a novel tool for investigating the diversity of hypothesized roles of synchronous neural dynamics and their impact on computation.

**Additional Detail:** Neural oscillations and traveling waves have long been a subject of study in neuroscience and neurophysiology (Hughes, 1995; Muller et al., 2018). Recent advances in measurement technology have brought renewed attention to the subject by allowing for much more precise observation of traveling waves in awake subjects (Davis et al., 2020). Along with the increase in experimental results, a diversity of hypotheses have been developed for the computational roles of these dynamics, including: the contextualization and integration of distant information (Sato et al., 2012), the sequencing of motor cortex activation (Takahashi et al., 2015), and the consolidation of memories during sleep (Muller et al., 2016). However, to-date, the computational models capable of investigating these hypotheses have been limited to those which either are built for the primary purpose of analysis (Davis et al, 2021), or those which perform very simple binary operations (Gong & van Leeuwen, 2009), with neither set leveraging the flexible computational capabilities of modern deep neural networks.

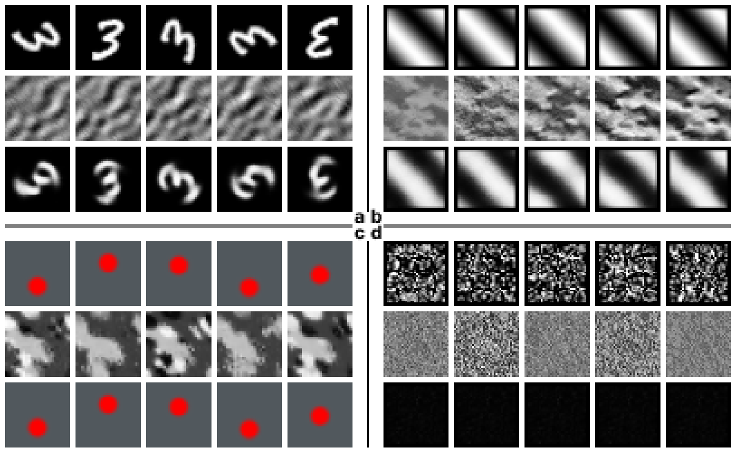


**Figure 1:** Visualization of input sequences **x** (top), hidden state **z** (middle), and reconstructions $\hat{\mathbf{x}}$ (bottom), on four datasets (quadrants **a**, **b**, **c**, **d**). Pixel brighteness of **z** depicts the instantaneous firing rate of spatially located neurons. The LocoRNN learns to exhibit different dynamics per dataset: compact traveling waves on rotating MNIST (**a**), diffuse interacting waves on simple periodic inputs (**b**), standing waves when modeling spring dynamics (**c**), and unstructured independent oscillatory dynamics for unstructured noise (**d**).

A well known model of traveling waves from the neuroscience literature is based upon a dynamical system of locally coupled oscillators (Diamant & Bortoff, 1969; Ermentrout & Kleinfeld, 2001). In this work, we leverage the recently popularized connection between recurrent neural networks (RNNs) and discretized ordinary differential equations to integrate the structured and biologically relevant dynamics of locally coupled oscillator networks into a powerful RNN model. Specifically, we consider the system:

$$\frac{\partial^2 \mathbf{z}}{\partial t^2} = \sigma\left(\mathbf{w}_z \star \mathbf{z} + \mathbf{w}_{\dot{z}} \star \frac{\partial \mathbf{z}}{\partial t} + \mathbf{w}_x \star f_\theta(\mathbf{x}) + \mathbf{b}\right) - \gamma \mathbf{z} - \alpha \frac{\partial \mathbf{z}}{\partial t} \tag{1}$$

The function $f_\theta$ denotes the encoder, $\sigma(\cdot) = \tanh(\cdot)$ the transfer/activation function, $\mathbf{x}(t)$ the time-varying input, and $\mathbf{z}(t) \in \mathbb{R}^{D \times D}$ the dynamic hidden state organized as a fixed 2-dimensional lattice. Locality of the recurrent connections is enforced by the convolution operation ($\star$) performed spatially over the lattice. In practice, this equation is discretized and integrated numerically yielding the following RNN equations:

$$\mathbf{z}_{t+1} = \mathbf{z}_t + \Delta t(\mathbf{v}_{t+1}), \qquad \mathbf{v}_{t+1} = \mathbf{v}_t + \Delta t\Big(\sigma\big(\mathbf{w}_z \star \mathbf{z}_t + \mathbf{w}_{\dot{z}} \star \mathbf{v}_t + \mathbf{w}_x \star f_\theta(\mathbf{x}_{t+1}) + \mathbf{b}\big) - \gamma \mathbf{z}_t - \alpha \mathbf{v}_t\Big) \tag{2}$$

We denote this model the Locally Coupled Oscillatory Recurrent Neural Network (LocoRNN), owing to the fact that it can be seen as a Coupled Oscillatory Recurrent Neural Network (coRNN) (Rusch & Mishra, 2021) modified to restrict recurrent connectivity to be topographically local.

To validate the computational capabilities of our model, we introduce a decoder $g_\theta(\mathbf{z}) = \hat{\mathbf{x}}$, and train the parameters of this model ($\theta$, $\mathbf{w}_\cdot$, & $\mathbf{b}$) with backpropagation through time (BPTT) to minimize a forward predictive reconstruction loss $\mathcal{L}$ over sequences: $\mathcal{L} = \sum_t^T ||g_\theta(\mathbf{z}_{t+1}) - \mathbf{x}_{t+2}||_2^2$. In Table 1, we see that indeed, as desired, our model achieves performance competitive with state of the art on a complex physical system forecasting benchmark, placing it in stark contrast to existing traveling wave models. To validate the structured synchronous dynamics of our model, in Figure 1 we plot the hidden state over time with corresponding input sequences (above) and reconstructions (below). We see the model exhibits a diversity of well known complex synchronous dynamics such as traveling and standing waves for different datasets. Interestingly, the model also exhibits no spatio-temporal structure of hidden dynamics when trained on random noise (bottom right), or before training (not shown), suggesting the model has learned to leverage synchronous dynamics in the service of modeling spatio-temporally structured observations. Qualitatively, we further observe that the training loss of the model dramatically decreases precisely when synchronous dynamics appear to emerge. In future work, we intend to quantify this phenomenon more accurately.

| | HGN++ | ODE | LSTM | LocoRNN |
|---|---|---|---|---|
| Spring | 447 (0) | 430 (26) | 302 (63) | 311.8 (27) |
| Pendulum | 105 (21) | 212 (65) | 3 (4) | 155.1 (24) |
| Dbl. Pend. | 11 (5) | 22 (7) | 0 (0) | 9 (9) |
| Two Body | 444 (3) | 439 (11) | 263 (92) | 413 (53) |
| RPS | 141 (23) | 124 (23) | 77 (15) | 133 (18) |
| Pennies | 79 (6) | 164 (14) | 118 (25) | 141 (37) |

**Table 1:** Number of future prediction steps with reconstruction error below a small value $\epsilon$ (0.025), known as Valid Prediction Time 'VPT' (± std. 3 runs), on the Hamiltonian Dynamics Benchmark. We see the LocoRNN combined with deep convolutional encoders is able to forecast diverse physical dynamics accurately up to hundreds of steps, similar to state of the art models such as Hamiltonian Generative Networks (HGN++), Neural ODEs (ODE), and LSTMs.

Compared to existing RNN and spiking models of synchronous dynamics, our model inherits the provably bounded hidden state and gradient magnitudes of the coRNN system, thereby ameliorating the exploding & vanishing gradient problem and allowing BPTT to scale to much longer sequences. Further, distinct from the original coRNN, the LocoRNN only contains local recurrent connections, reducing parameter complexity while permitting the study of the spatio-temporally distributed processing. In summary, we believe our model is the first to both exhibit complex wave phenomena and simultaneously remain a powerful trainable sequence model, thereby presenting value as an investigative tool of such synchronous dynamics.

To exemplify this potential value, we perform a preliminary experiment to test the hypothesis that spontaneously generated retinal waves contribute to visual system topographic organization prior to visual experience (Wiesel & Hubel, 1974; Ackman et al, 2012). Specifically, we train the model to reconstruct simple periodic sine waves (depicted in Figure 1, top right), and then measure the orientation selectivity of each hidden neurons time-average response to static sequences of oriented gratings using Cohen's $d$ metric (Cohen, 1988). In Figure 2 we plot the resulting color/angle of maximal $d$ value for each of the $72 \times 72$ neurons (or a black $\mathbf{x}$ if all $d < 0.65$). We see that the simulated retinal waves do appear to induce topographic organization of orientation selectivity reminiscent of the orientation columns of primary visual cortex. In relation to prior models of orientation columns (Swindale, 1982), our work does not presuppose the existence of orientation selectivity, but rather it is absent at initialization and it is instead learned in conjunction with topographic organization. On MNIST, we further see topographic organization of class-selectivity (not shown), suggesting a more general method for topographic organization. Anonymized code and videos at: github.com/q2w4/LocoRNN

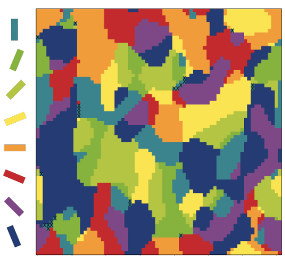

**Figure 2:** Orientation selectivity ($d' > 0.65$) of LocoRNN hidden neurons $\mathbf{z}$ trained on simulated retinal waves.