Supplement to “Nearly Extensive Sequential Memory Lifetime Achieved by Coupled Nonlinear Neurons” by Taro Toyoizumi, *Neural Computation,* Vol. 24, No. 10 (October 2012), pp. 2678–2699.

Supplement contains color versions of Figures 1-5 from the article.

URL: http://www.mitpressjournals.org/doi/abs/10.1162/NECO_a_00324

Supplemental materials for "Nearly extensive sequential memory lifetime achieved by coupled nonlinear neurons" by Taro Toyoizumi, *Neural Computation* (2012).

A color version of Figs. 1-5:

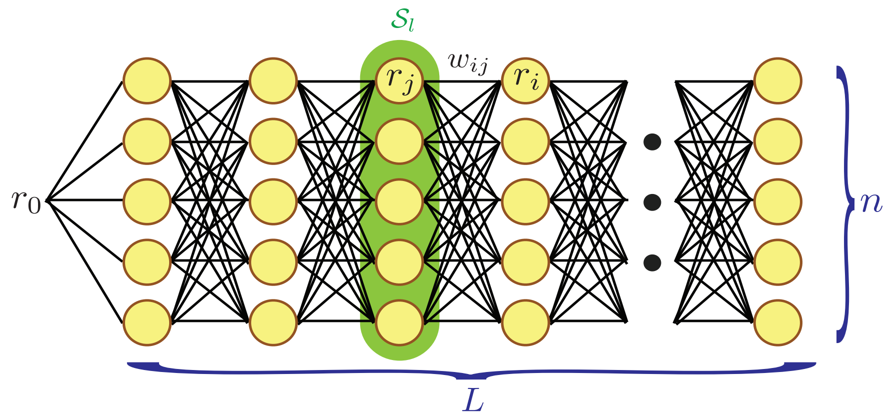

Figure 1: Simple feedforward network model of size N. There are n neurons in each layer, and there are $L = N/n$ successive layers. Each neuron is connected to all the neurons in the previous layer with a uniform synaptic strength. We study how the input of strength, r_0, propagates down the feedforward chain.

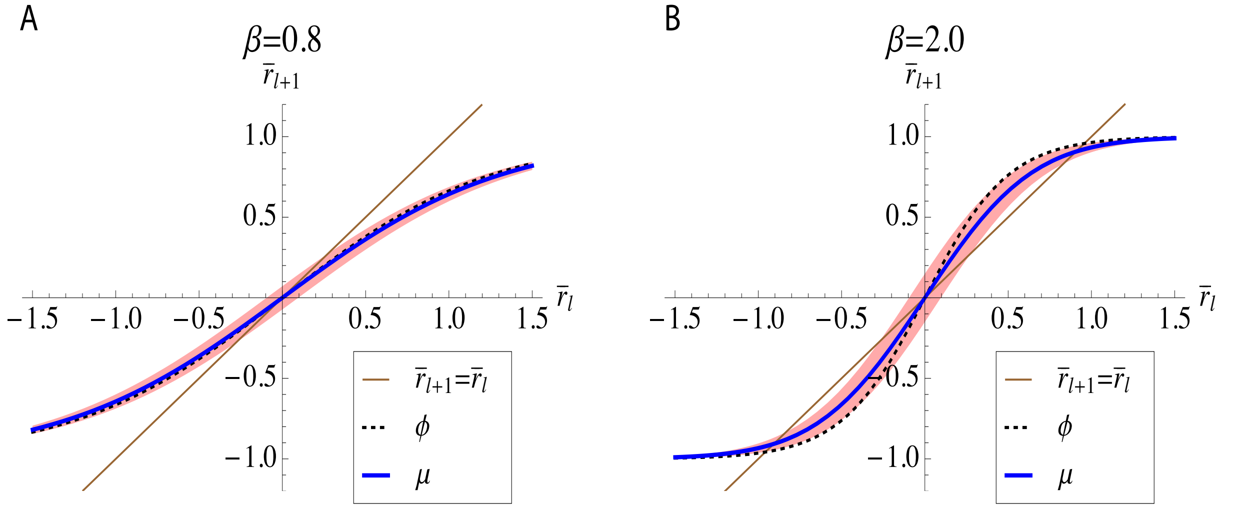

Figure 2: Conditional probability $P(\bar{r}_{l+1}|\bar{r}_l)$ of the average activity. The blue solid line is the conditional mean, $\mu(\bar{r}_l)$, and the pink band is the conditional standard deviation, $\sqrt{v(\bar{r}_l)/n}$. The brown line indicates the condition $\bar{r}_{l+1} = \bar{r}_l$, and the dashed line shows the nonlinear response function $\phi(x) = \tanh(\beta x)$. (A) The slope of ϕ is $\beta = 0.8$. Here, $\bar{r} = \mu(\bar{r})$ has only one attracting solution at $\bar{r} = 0$. Hence, the activity tends to decay toward 0. (B) The slope of ϕ is $\beta = 2$. Here, $\bar{r}_{l+1} = \mu(\bar{r}_l)$ has three fixed points: two ($\bar{r} \approx \pm 0.9$) are attractive and one ($\bar{r} = 0$) is repulsive. When $\bar{r}$ is close to one of the attracting fixed points, noise does not accumulate because it is partially removed at each time step. Other parameters are set to $\sigma = 0.3$ and $n = 10$.

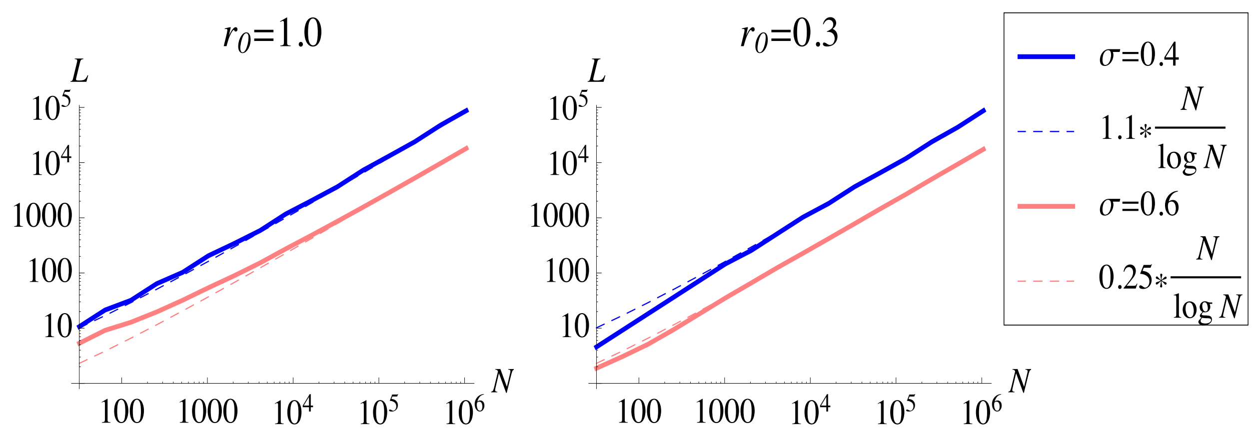

Figure 3: Memory lifetime of binary neurons is scaled close to the network size. The memory lifetime, L, was evaluated at two noise levels: $\sigma = 0.4$ (blue solid) and $\sigma = 0.6$ (red solid), and two inputs: $r_0 = 1.0$ (*Left*) and $r_0 = 0.3$ (*Right*). The offset of two curves at different noise levels reflects the different number of neurons in each layer, n, chosen to achieve the 90% decoding criterion. The scaling behavior was well fitted by $\sim N/\log N$ in all cases as suggested by the theoretical result.

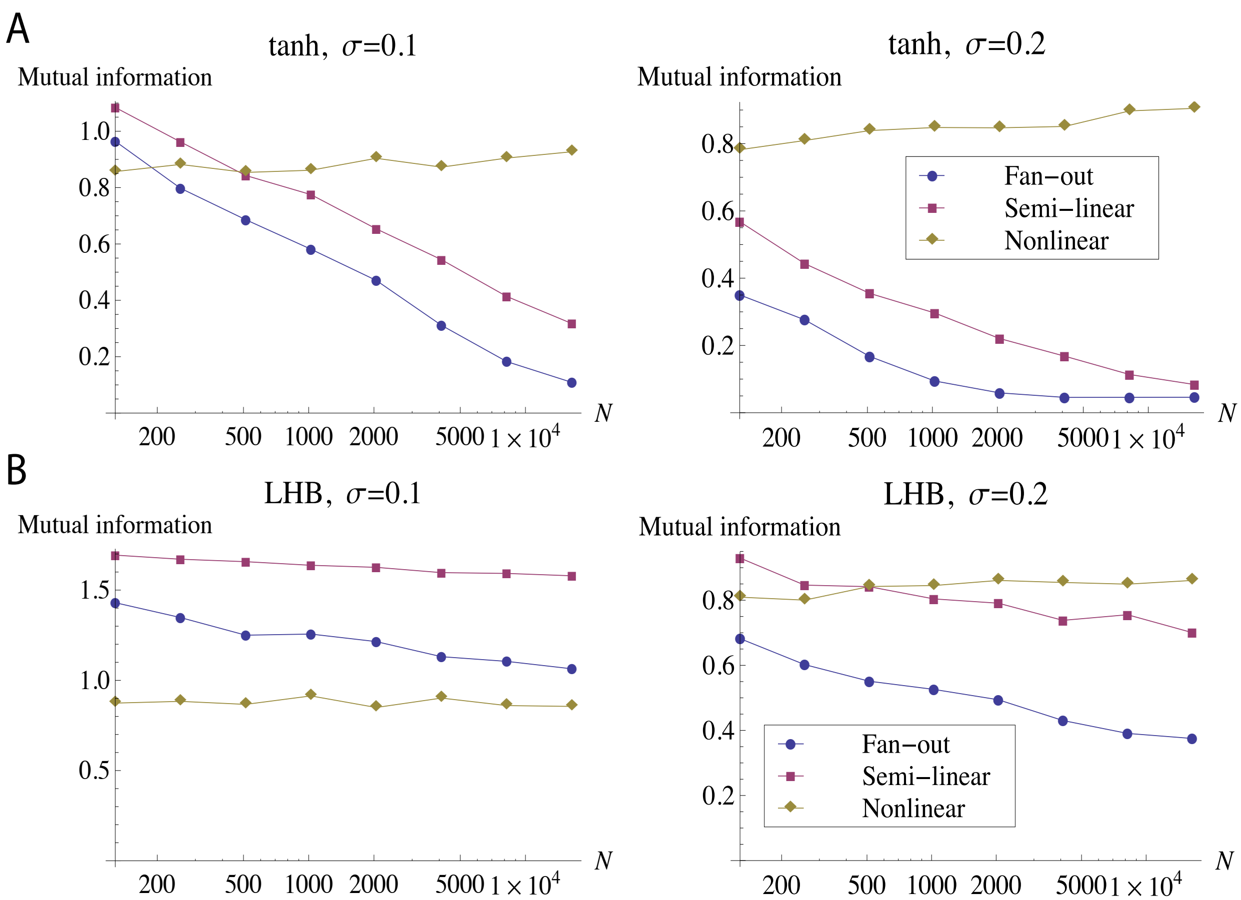

Figure 4: The nonlinear network outperformed the semi-linear network for large network sizes or with large noise. A Gaussian random input with zero mean and standard deviation $\sigma_0 = 0.5$ was provided to the first layer, and the performance was measured by the mutual information of the input and the activity in the final layer. The sequential memory performance was numerically examined at noise level σ, shown in each panel and with two nonlinear functions: (A) the hyperbolic tangent nonlinearity $\phi(x) = \tanh(\beta x)$ and (B) a piecewise-linear function $\phi(x) = \beta x$ for $|x| < 1/\beta$ with hard saturating bounds (LHB). Three types of networks were compared with approximately the same size, N, and the same number of layers $L \approx \sqrt{2N}$ for a fair comparison: the fan-out network (Ganguli et al., 2008) with $\beta = 1$ and a linearly increasing number of neurons along deeper layers ($n_l = l$; order 1 memory lifetime); the semi-linear network with a β solution that yielded a gain equal to 1 and a fixed number of neurons in each layer ($n_l = N/L$; order $\sqrt{N}$ memory lifetime); and the nonlinear network with the same network architecture as the semi-linear network but with $\beta = 2$ ($n_l = N/L$, order $N/\log N$ memory lifetime). The semi-linear network always showed better performance than the fan-out network and the nonlinear network was superior to the other two except at a small network size and with a small amount of noise.

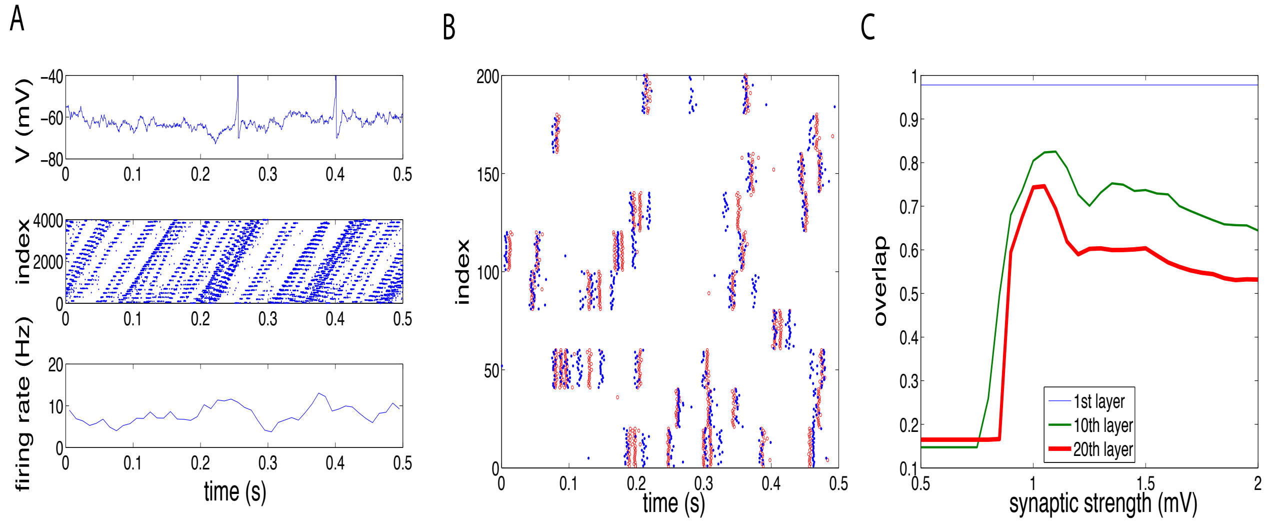

Figure 5: A feedforward network of leaky integrate-and-fire neurons reliably buffered spike patterns. (A) *Top*: Membrane potential of a single neuron. *Middle*: Spike-timing of all neurons in the network. The network consisted of $k = 10$ independent synfire chains, where each synfire chain had $L = 20$ layers and $n = 20$ neurons in each layer. The oblique patterns describe feedforward propagation of synfire activity. *Bottom*: The population firing rate of all the neurons averaged in 10-ms bins. (B) The spiking pattern of the first layers (blue dots) was well preserved even until the final 20th layers (red circles). The spike pattern of the final layers was shifted so that the spike overlap with the first layers was maximized. The feedforward synaptic strengths were set to 1 mV. Note that input pulses were somewhat degraded by background noise even in the first layer. (C) The spike overlap with the input pulses in the 1st, 10th, and 20th layers plotted for different feedforward synaptic strengths.