

Transient attractor strength in a hippocampal model

Xiangbao Wu and William B Levy

University of Virginia Health System

P.O. Box 800420 Neurosurgery

Charlottesville, VA 22908-0420, USA

Abstract

Using computer simulations of a hippocampal model, this report investigates how stutter length (timesteps of repeating each external input pattern) affects transient attractor strength. When using parameters that reproduce hippocampally dependent cognitive learning tasks, permanent attractors do not form but transient attractors do. Such transient attractors are not simply localized, but certain characteristics of these attractors can be implied from the relative trajectories of state space transitions. The simple sequence completion problem with orthogonal inputs is employed and a measurement is suggested for this purpose. We show that stutter length profoundly affects transient attractor strength.

Keywords: CA3; neural network; dynamics; computation.

Introduction

In neural networks asymptotically stable attractors seem unrealistic. That is, brain states seem to be in continuous flux. Therefore the idea of a “transient attractor”

seems more applicable than permanent attractor. In particular, a transient attractor pulls a state space sequence towards itself but only briefly, and then it is replaced by another such transient attractor.

Such a model seems consistent with sequence learning models (e.g. [1, 2]). However, there is little to guide us in terms of establishing the existence of and quantifying these transient attractors. Here we present one possible quantification.

The model

Our hippocampal model is essentially a model of region CA3 (see e.g. [2, 4, 5, 6]). The input layer corresponds to a combination of the entorhinal cortex and dentate gyrus. The CA3 model is a sparsely (10%) interconnected feedback network of 2048 neurons where all direct connections are excitatory and the network elements are McCulloch-Pitts neurons. There is an interneuron mediating feedforward inhibition, and one mediating feedback inhibition. Inhibition is of the divisive form, but the system is not purely competitive because of a slight delay. Synaptic modification develops over training. The process controlling synaptic modification is a local, self-adaptive postsynaptic rule that includes both potentiation and depression aspects [1, 3]. The network computations are all local and are contained in three equations: spatial summation adjusted by inhibition; threshold to fire or not; and local Hebbian synaptic modification (see [2, 4, 6, 8] for details).

The simulations described here used suitable inhibitory parameters to approximately achieve a 5% activity. At each timestep, 32 external inputs accounted for

about one-third of the total activity ($2048 \times 0.05 \approx 102$). Thus, feedback activation accounts for more cell firings and, by the end of training, does so via long local context neuronal firings (see e.g. [5, 8, 9]).

Detecting the transient attractor

A sequence completion task is employed in this report. The inputs are 42 orthogonal input patterns (e.g. p1, p2, p3, ..., p42). Each pattern is repeated 3 timesteps in the stutter 3 case and repeated 2 timesteps in the stutter 2 case [9].

Each sequence is presented to the network for 300 trials of learning. Before each training and test trial, CA3 neurons are randomly excited to produce a predefined number of active neurons (here 102; for more details, see e.g. [6]), and this initialization state vector is called $Z(0)$. To test how strong the transient attractor is at the end of training, the synaptic modification is disabled, and we twice give each network a sequence, e.g. $Z(0)(p1)(p2)...(p20)(p21 \text{ plus noise})$ and then let it free run. We then compare the two sequences of cell firing, where differences in the two sequences are only due to the differences of random excitation at timestep 21.

Hamming distance is used to measure the distance between two trajectories, time point matched to time point. Pairs of simulations are compared. Each of the pair has had identical training so they are identical networks. Two trajectories are produced because input pattern 21 is perturbed. If 4 random neurons are turned on at timestep 21, the noise/signal ratio is defined as $4/32$, i.e. 0.125.

To measure how strong the transient attractor around p21 is, we calculate Hamming

distance between all pairs of the two sequences of firing states. We especially pay attention to the Hamming distance between the pair at time step 22. That is, the states following by p21 plus noise of the two testing sequences. The smaller the Hamming distance is, the stronger the transient attractor is.

Results

Figure 1 shows Hamming distance between corresponding pairs of two trajectories by giving $Z(0)(p1)(p2) \dots (p20)(p21 \text{ plus noise})$ twice, where differences in the two sequences are only due to the differences of random excitation at timestep 21. Note how the two trajectories converge initially and diverge later on. That is, the two trajectories converge (the Hamming distance decreases) for one timestep and they stay close for two more then diverge.

Figure 2 shows the average Hamming distance between the pairs of trajectories at timestep 22 as a linear function of stutter length. From this figure, one can see clearly that the transient attractor is stronger for stutter 3 than for stutter 2. And the transient attractor is stronger for stutter 2 than for stutter 1.

Figure 3 plots the average Hamming distance as a function of noise/signal ratio. All three curves have a similar shape, but as the noise/signal ratio increases, the Hamming distance tends to increase. Just as in the previous figure, the Hamming distance is larger for stutter 1 than for stutter 2 or stutter 3. For stutter 3, the Hamming distance is quite robust against increases of the noise/signal ratio and changes very little.

Conclusion

We have shown that stutter length can profoundly affect transient attractor strength. This is consistent with, and extends, our previous findings that stuttering the orthogonal inputs helps learning [9]. Moreover, these results suggest a way to understand how strong a transient attractor is.

Acknowledgments

This work was supported by NIH MH48161 to WBL and MH57358 to N. Goddard (with subcontract #163194-54321 to WBL).

References

- [1] W.B Levy, A computational approach to hippocampal function, in: R.D. Hawkins and G.H. Bower, eds., Computational Models of Learning in Simple Neural Systems, (Academic, New York, 1989) 243-305.
- [2] W.B Levy, A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks, *Hippocampus* 6 (1996) 579-590.
- [3] W.B Levy and O. Steward, Synapses as associative memory elements in the hippocampal formation, *Brain. Res.* 175 (1979) 233-245.
- [4] W.B Levy and X.B. Wu, The relationship of local context codes to sequence length memory capacity, *Network:Computation in Neural Systems* 7 (1996) 371-384.

- [5] W.B Levy and X.B. Wu, A simple, biologically motivated neural network solves the transitive inference problem, IEEE International Conference on Neural Networks, Vol. I (IEEE, Piscataway, 1997) 368-371.
- [6] A. Shon, X.B. Wu, D. Sullivan and W. B Levy, Initial state randomness improves sequence learning in a model hippocampal network, Phys. Rev. E, in press.
- [7] A. A. Tsonis, Chaos: From Theory to Applications, (New York: Plenum Press, 1992).
- [8] X.B. Wu, R.A. Baxter and W.B Levy, Context codes and the effect of noisy learning on a simplified hippocampal CA3 model, Biol. Cybern. 74 (1996) 159-165.
- [9] X. B. Wu, J. Tyrcha and W. B Levy, A neural network solution to the transverse patterning problem depends on repetition in the input code, Biol. Cybern., 79 (1998), 203-213.

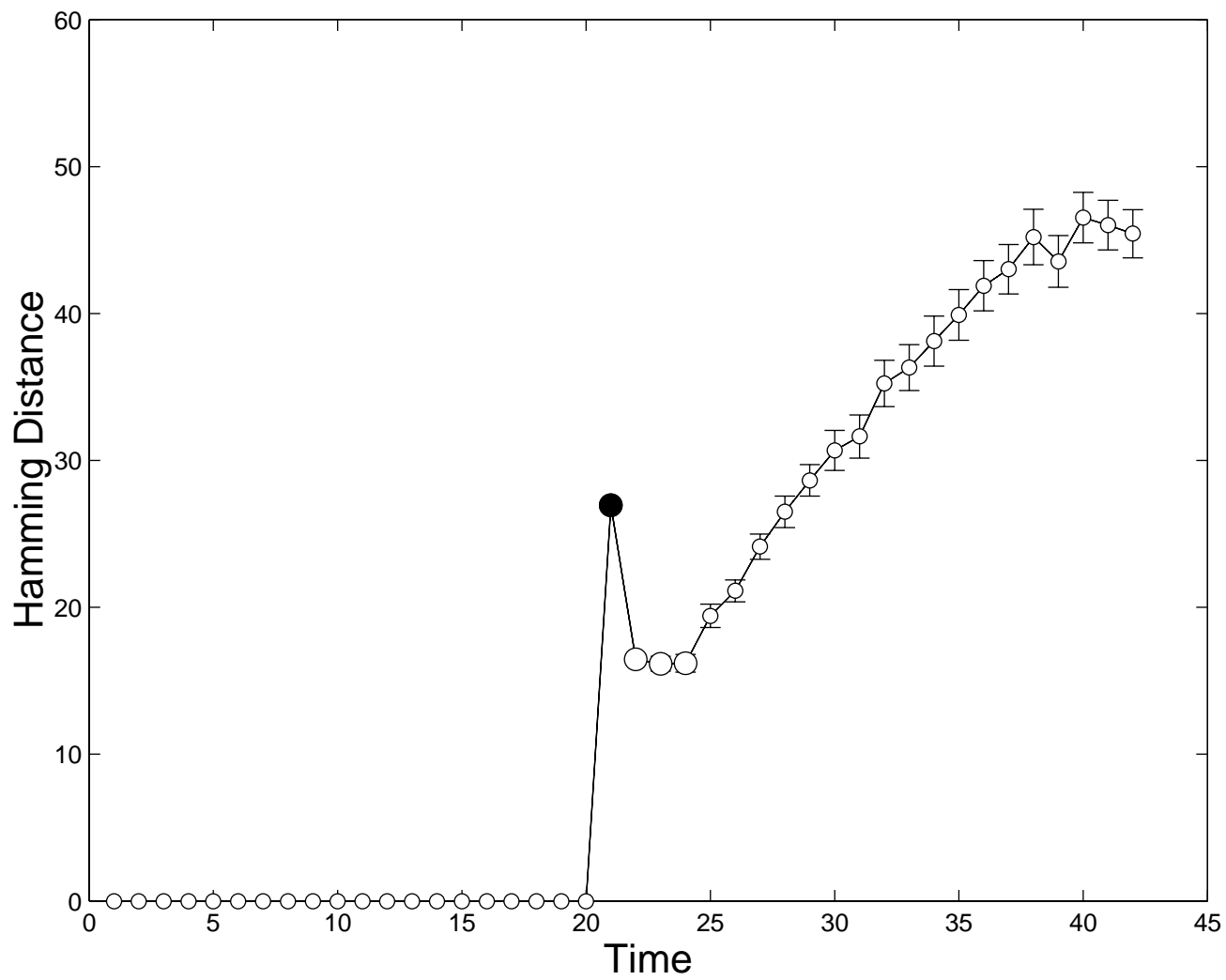


Figure 1.

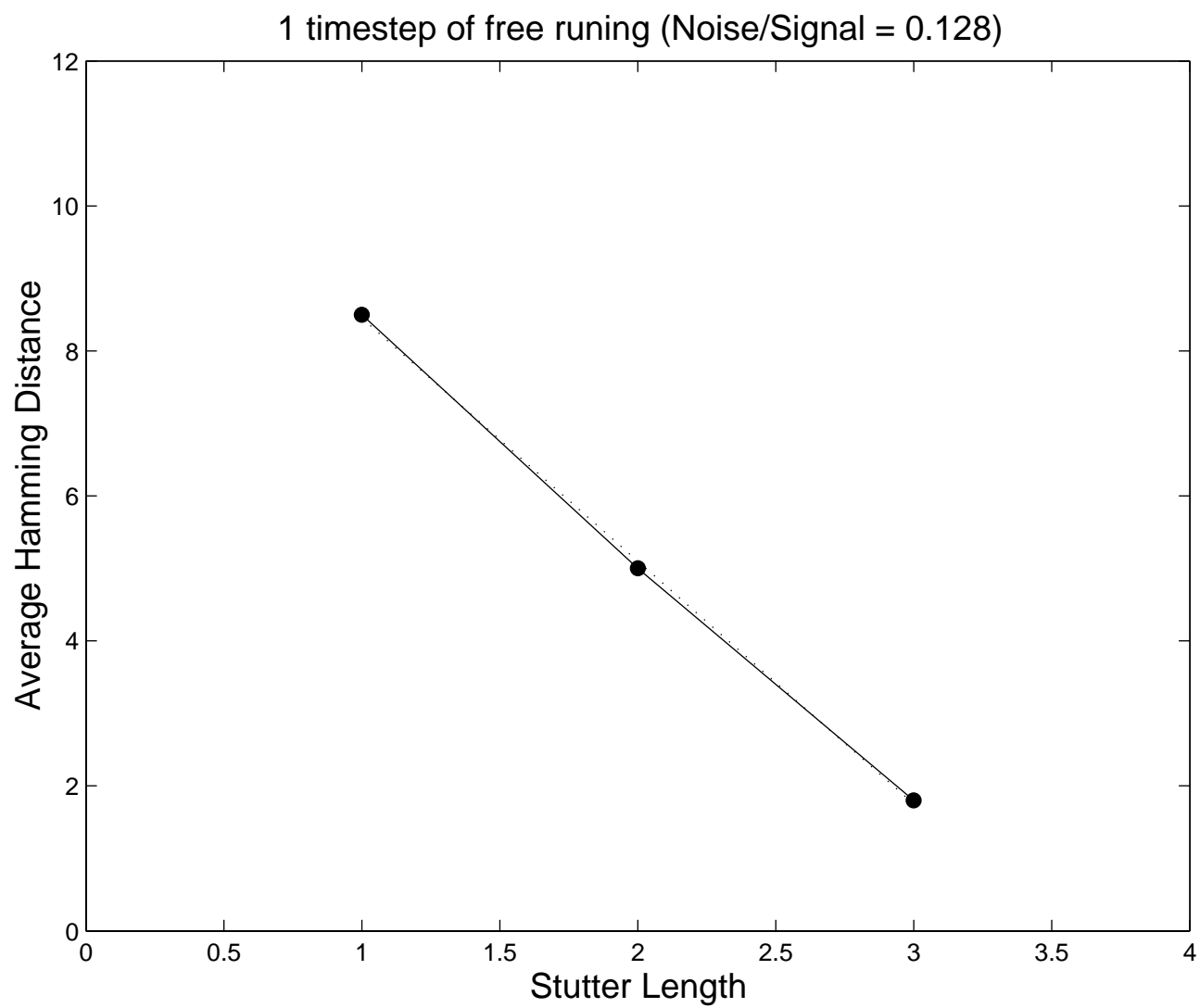


Figure 2.

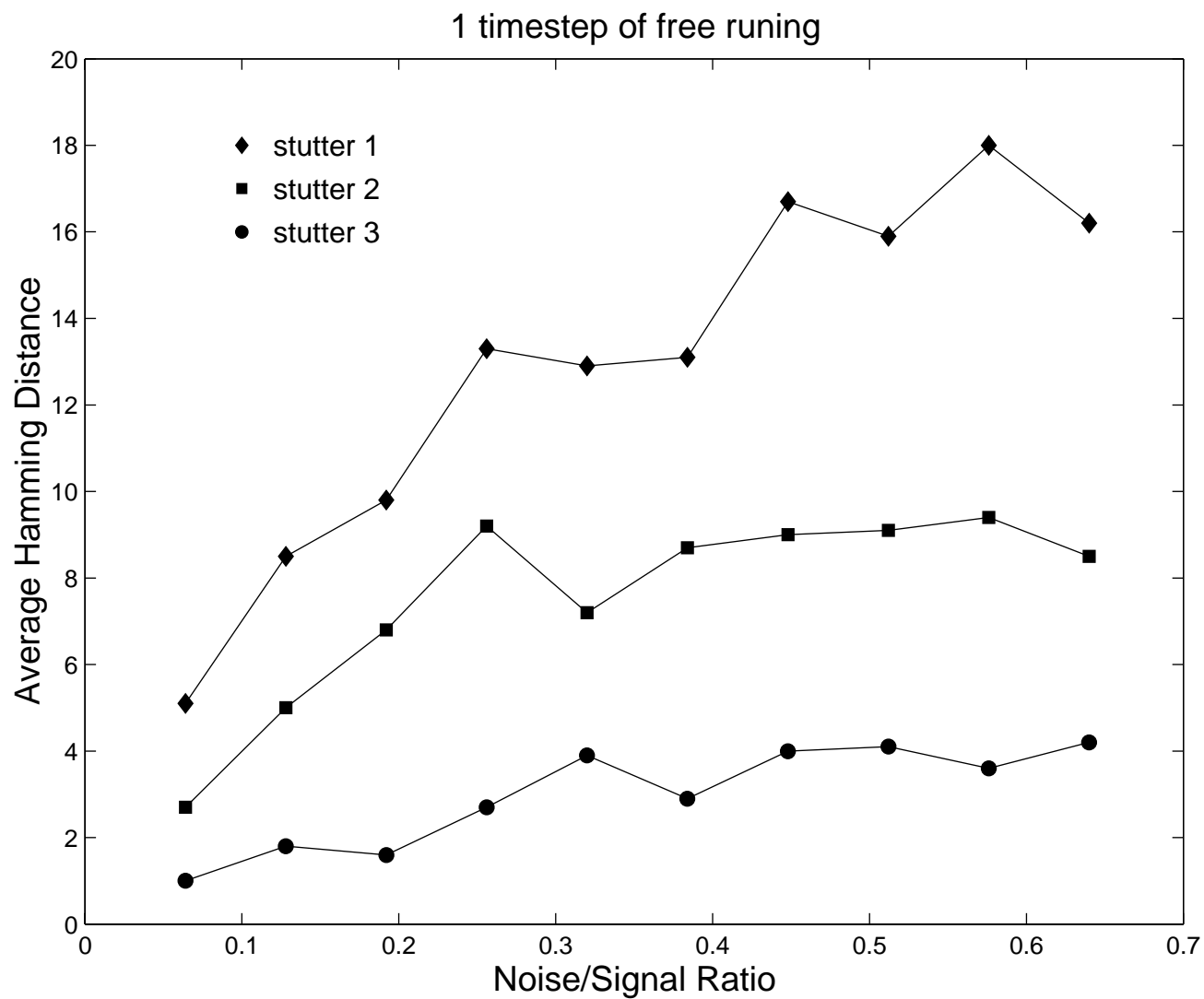


Figure 3.

Figure Legends

Figure 1: An example to show Hamming distance between two sequences near a transient attractor. With the synaptic modification disabled, we twice give each network a sequence, e.g. $Z(0)(p1)(p2) \dots (p20)(p21)$ plus noise. We then compare the two sequences of cell firing, where differences in the two sequences are only due to the differences of random input excitation at timestep 21. Note how the two trajectories converge for one timestep and they stay close for two more (see the three larger open circles); subsequently they diverge. These three special timesteps could be defined as the transient attractor lifetime. The filled circle represents the Hamming distance due to the random neurons turned on at timestep 21. Error bars represent the standard error of the mean for 100 pairs of simulations.

Figure 2: The network gives stronger transient attractors for stutter 3 than for stutter 2 and the transient attractor for stutter 2 is stronger for stutter 1. Average Hamming distance between the two trajectories at timestep 22 is a linear function of stutter length. Each point is an average of ten simulations and the correlation coefficient of the linear fit is 0.99.

Figure 3: Hamming distance as a function of noise/signal ratio at timestep 22. Hamming distance is much higher for stutter 1 than for stutter 2 or stutter 3. Note the tendency that Hamming distances increase as the noise/signal ratio increases but reach a relative plateau.