SUPPLEMENTARY INFORMATION

ROBUST TIMING AND MOTOR PATTERNS BY TAMING CHAOS IN RECURRENT NEURAL NETWORKS

Rodrigo Laje and Dean V. Buonomano

Supplementary Figures S1, S2, S3, S4, S5, S6, S7, S8.
Supplementary Modeling.
Supplementary example Matlab programs

SUPPLEMENTARY FIGURES

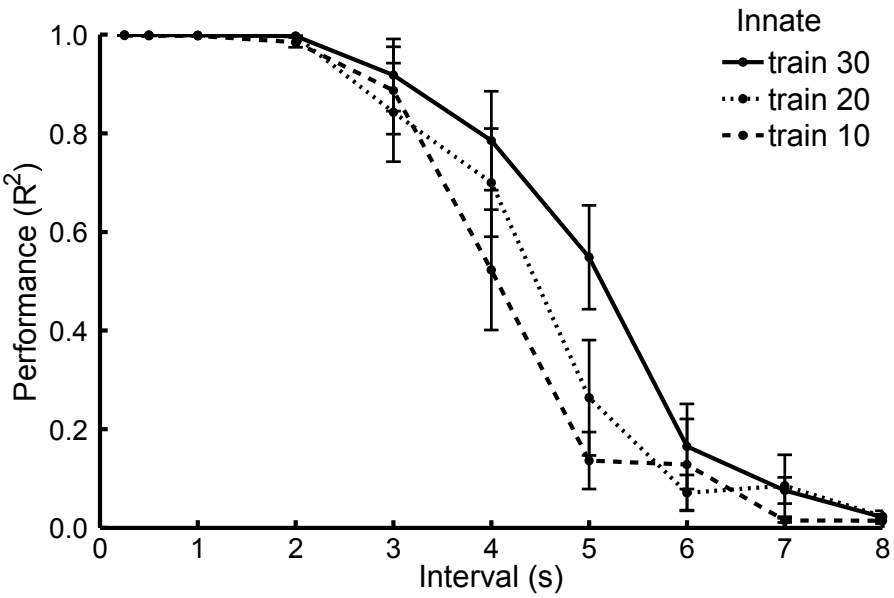

Figure S1. Trade-off between amount of training and target interval in the timing task. Performance of Innate training as in Figure 3, for different amounts of training (10, 20, and 30 training loops). The curve labeled "train 30" is the same as "Innate" in Figure 3b (Architecture 1). Horizontal cross-section: longer target durations need more training to reach a given intermediate performance level. Vertical cross-section: more training leads to better performance at intermediate given duration. Mean ± SEM across networks (the same 10 seed networks as in Fig. 3).

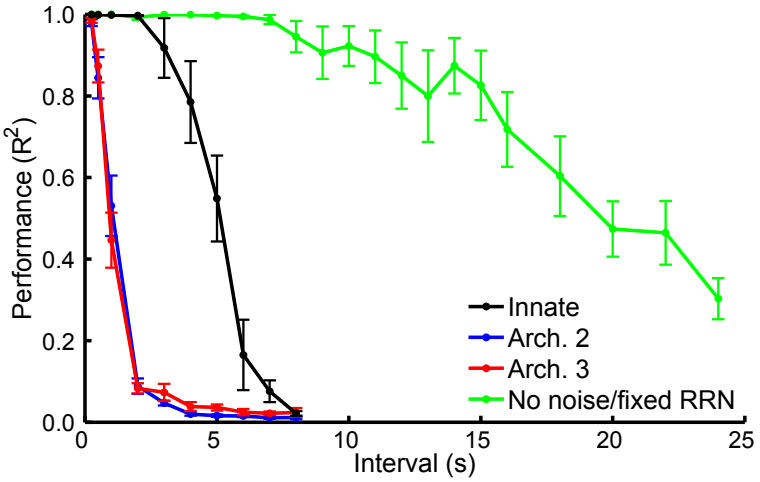

Figure S2. Upper bound on the timing capacity. Performance in the timing task. Black, blue, and red curves are exactly the same as in Figure 3b. The green curve represents the average performance of the same set of ten "seed" networks but with no recurrent training, no noise, no feedback, and always the same initial conditions. The readout unit was trained to match the pulsed target of the corresponding duration. Note that the green performance levels are essentially useless from a computational perspective because the network trajectories cannot be reporduced even with minute levels of noise. Mean ± SEM across networks (the same 10 seed networks as in Fig. 3).

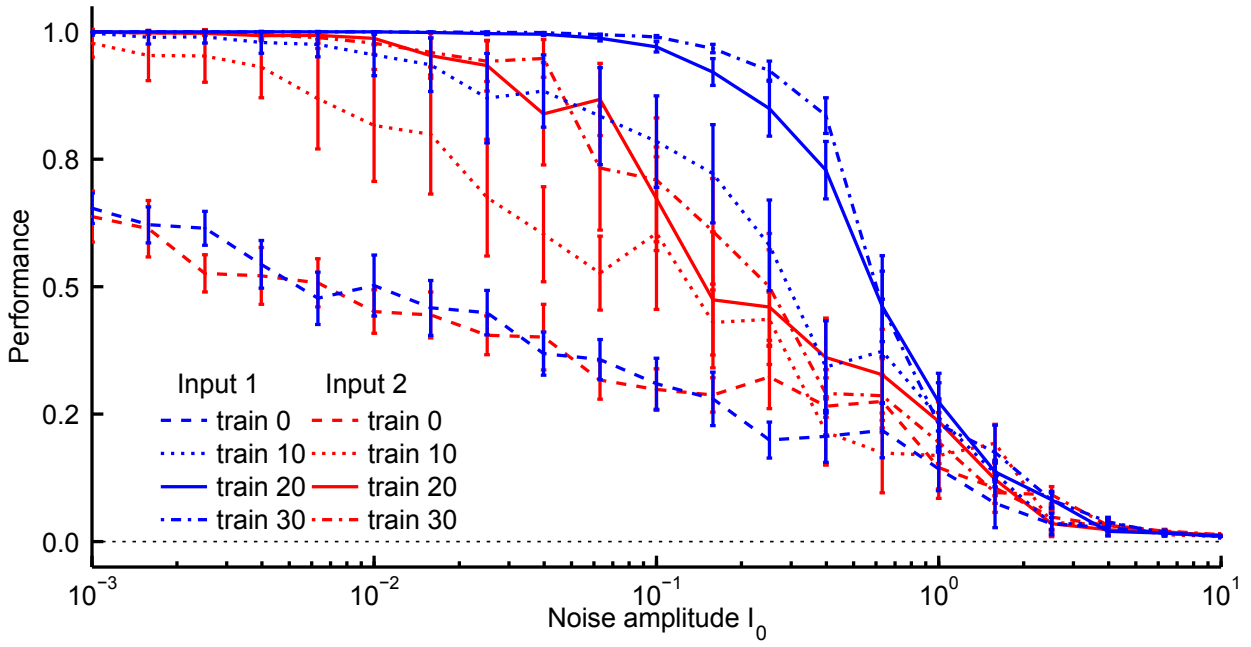

Figure S3. Stability of trained and control trajectories as a function of training. Performance as defined in Figure 5b, for different amounts of training (none, 10, 20, and 30 training loops). Dashed- and solid-line curves ("train 0" and "train 20") correspond to "Pre" and "Post" in Figure 5b, respectively. Training differentially improves performance at all stages of training above "train 0". Performance seems to reach a saturation beyond 20 training loops, both for the trained and the control trajectories. Mean ± SEM across networks (the same 10 seed networks as in Fig. 5).

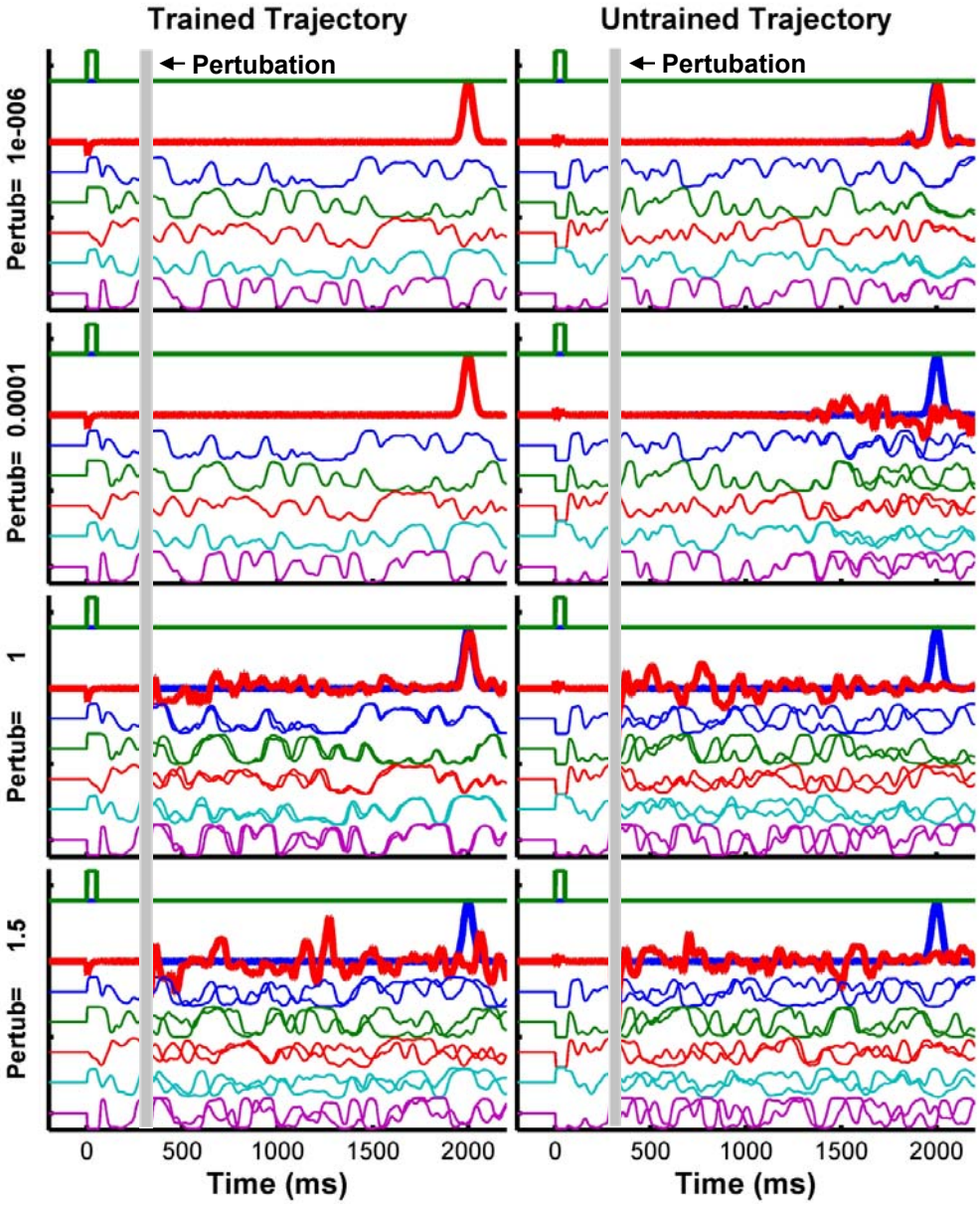

Figure S4. Trained trajectories are dynamic attractors, while the untrained trajectory is unstable. Results from one of the 10 networks shown in Figs. 5 and 6. In each panel the top green line represents the 50-ms input (amplitude = 5); the underlying blue trace (only visible in the lower panels) represents the 25-ms perturbation of the amplitude noted on the y axis. Thick blue and red traces correspond to the output traces in the absence and presence of the perturbation, respectively (noise is off). The output weights were trained to generate the same 2000-ms delayed response to both the trained and untrained trajectories (in the absence of any noise or perturbations). Lower colored traces show the activity of 5 sample recurrent units. In response to tiny perturbations the divergence can initially be "invisible" and only produce detectable changes late in the trial of the untrained input. In response to a perturbation of amplitude=1, both the trained and untrained trajectories are knocked off their path, but only the trained trajectory "finds its way back". The magnitude of the perturbation that still generates the appropriate output will be slightly different in different simulations and vary according to the level of noise used for training the recurrent network, and the amount of training.

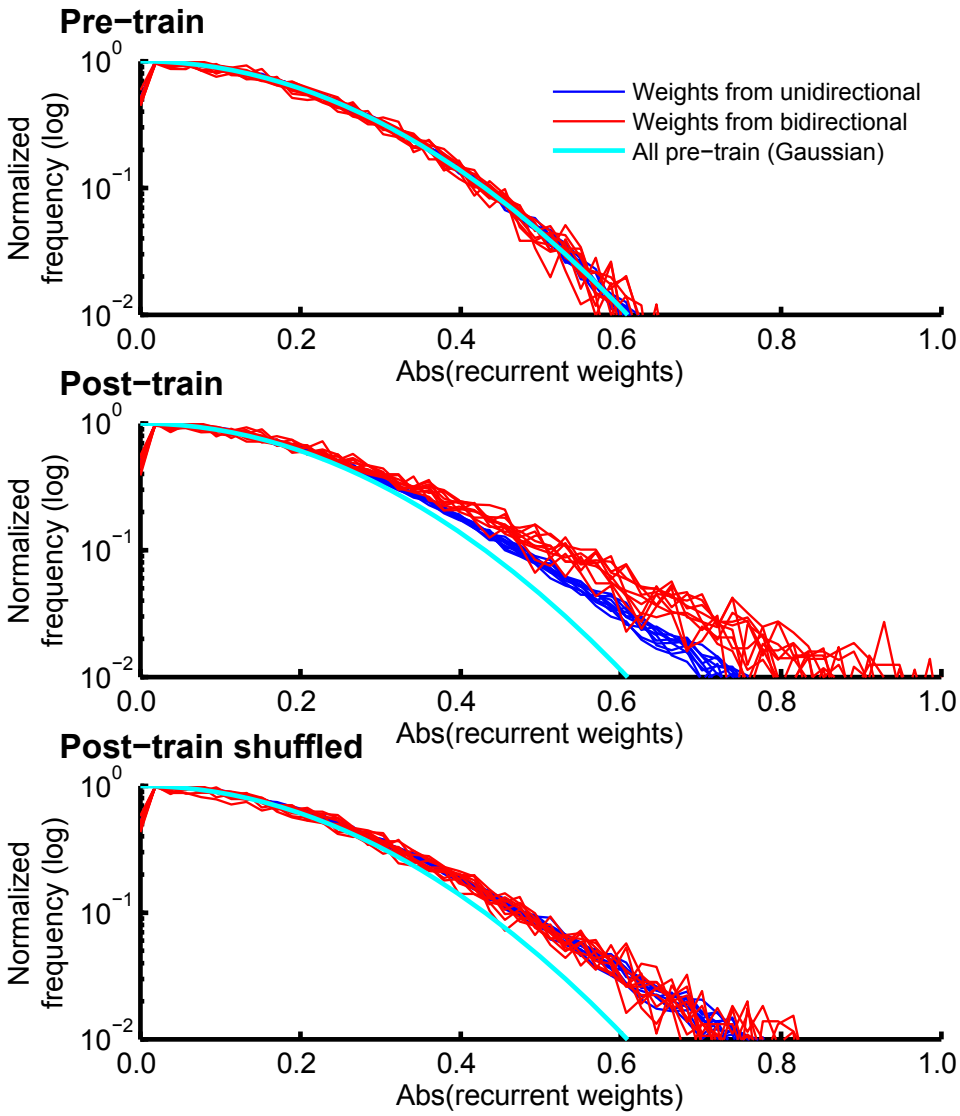

Figure S5. Weights from unidirectional versus bidirectional connections. Pre-training (top panel): weights from uni- and bi-directional connections are indistinguishable. Post-training (middle panel): weights from bidirectional connections became stronger in absolute value than those from unidirectional connections (unidirectional median ± MAD across networks: 0.145 ± 0.001; bidirectional: 0.161 ± 0.003; paired Wilcoxon sign-rank test, p=0.002). Shuffling of the weights while keeping the binary connectivity (bottom panel) eliminates the difference between unidirectional and bidirectional (paired Wilcoxon sign-rank test between uni- and bi-, p=0.6).

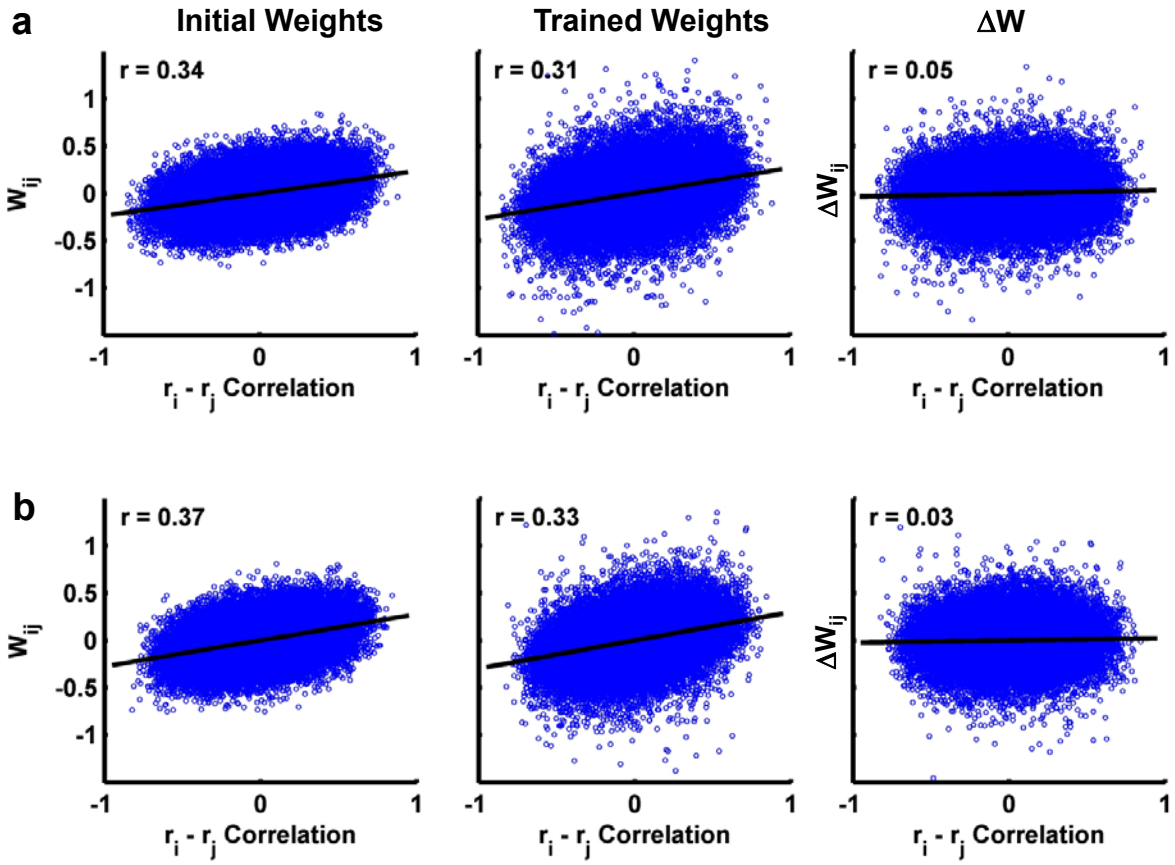

Figure S6. Correlation between synaptic weights and the firing rates of the pre- and postsynaptic units during the target ("innate") pattern. Panels **a** and **b** represent sample correlations from two different networks from Figure 5. (*Left*) Correlation between the initial synaptic weights W_{ij} (only those weights that were plastic were included, ~38400) and the correlation between the firing rates of units *i* and *j*. A small number of "autapses" (that had an activity correlation of 1) were excluded in these analyses. (*Center*) Correlation between the trained weights W_{ij} and correlated activity. (*Right*) Correlation between ΔW_{ij} and correlated activity. Across five networks the average correlation for the Initial, Trained, and Delta weights was $R = 0.355 \pm 0.005$, 0.322 ± 0.005, and 0.038 ± 0.004, respectively. The correlation values for the Initial and Trained weights were highly significant ($p < 10^{-16}$ for all 5 networks analyzed); the same was true for the correlations of the ΔW ($p < 10^{-6}$). Finally, when the *R* values of the 5 networks were grouped there was a significant decrease in the *R* values after training ($t_4 = 10.8$, $p < 0.0005$).

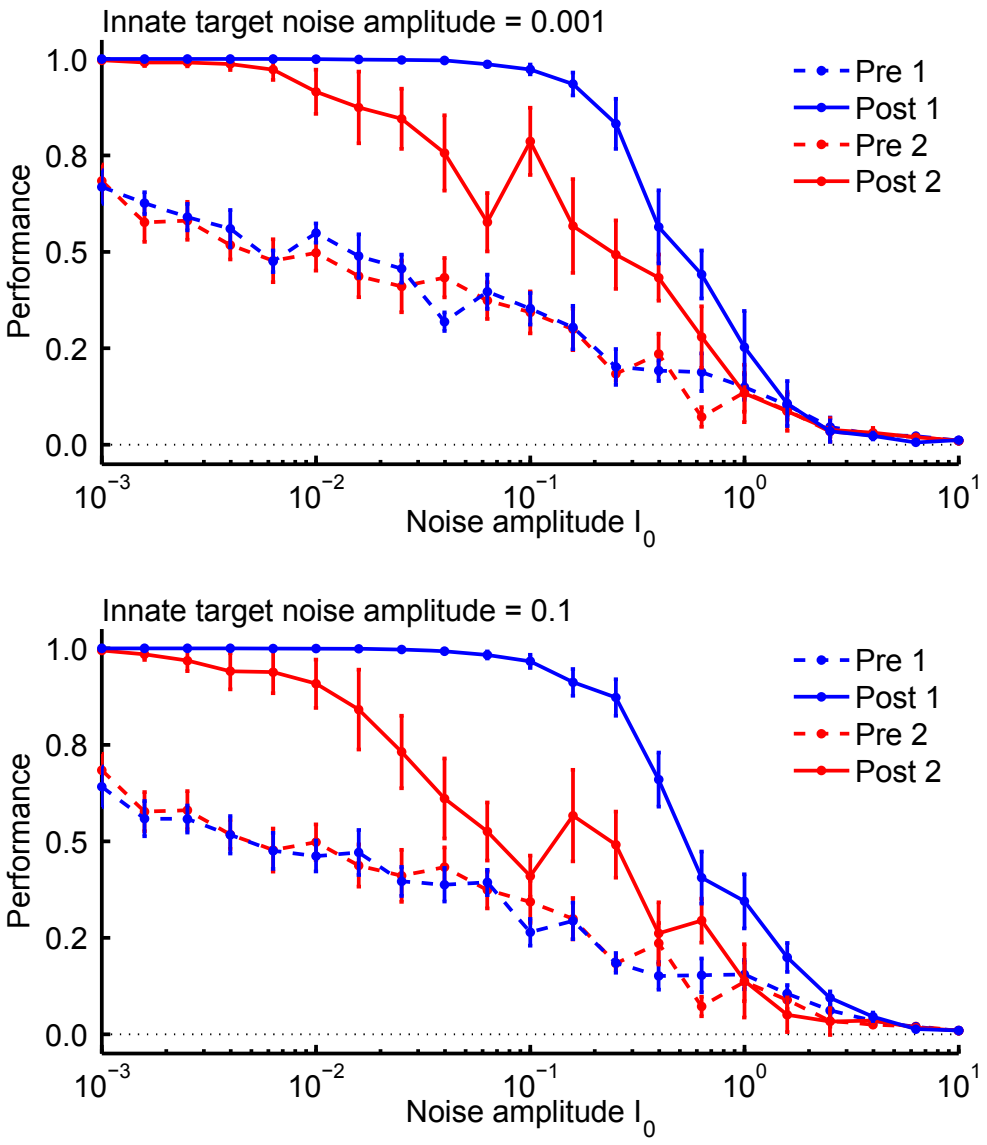

Figure S7. The innate trajectory can be chosen in the presence of noise. Average performance of 10 networks as a function of noise amplitude during testing, as in Fig. 5; the innate trajectory used for training was chosen under noise with two different amplitudes: 0.001 (top panel) and 0.1 (bottom panel). Training noise amplitude is always 0.001, as in Fig. 5; testing noise amplitude as shown on the x-axis. Mean ± SEM across 10 networks (the same 10 seed networks as in Fig. 5).

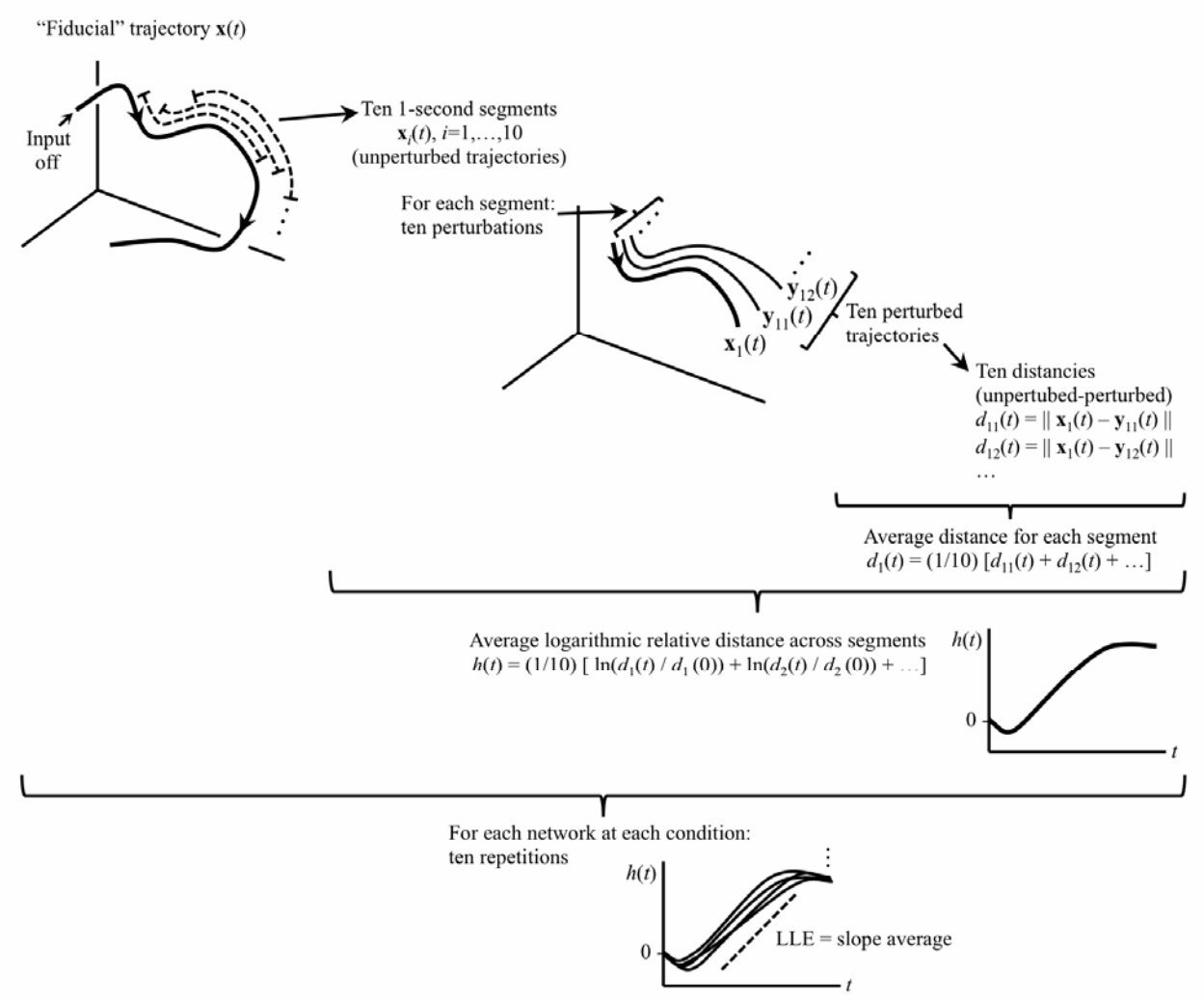

Figure S8. Schematic of calculation of Largest Lyapunov Exponents. Phase space is depicted as 3-dimensional for visualization purposes only.

SUPPLEMENTARY MODELING

Comparison to other chaos-related regimes

Below we briefly summarize other chaos-related regimes in complex systems to provide a contrast between different related phenomena. To the best of our knowledge a critical difference between our results and all previous known regimes is that the stable transient channel demonstrated here is the result of explicit modifications of the system itself—modifications designed to create stable trajectories within a chaotic system. This contrasts with previous work, which describe "natural" regimes of intact systems.

Regular chaos. Although many of our RRNs retained a chaotic attractor after training, the fact that the phase space has some locally stable trajectories makes our RRNs different from a regular chaotic system. In a regular chaotic system, the divergence rate of trajectories naturally fluctuates along the trajectory, and thus finding a particular portion of a trajectory where the flow converges rather than diverges is possible[1]; however, the probability of finding a long stable trajectory is very low. The convergence displayed at very short times (<50 ms) in **Fig. 6a** is only due to the fast relaxation of a random initial condition towards the attractor[2], and can also be found in a regular chaotic system.

Stable Chaos. Among the different types of regimes observed in complex and chaotic systems, *"stable chaos"* deserves special mention because it has been studied in randomly connected recurrent networks composed of integrate-and-fire units[3-4]. In this context *"stable chaos"* has been used to refer to irregular transient behavior despite a negative LLE. In these regimes periodic solutions are generally reached after a very long transient that is very irregular. Although the largest non-zero Lyapunov exponent is negative (i.e., the system is stable under infinitesimal perturbations), finite-size perturbations can lead to diverging trajectories; this is reflected in intricate basins of attraction for the different periodic solutions. There are a couple of differences between this phenomenon and our finding. First, we observe locally stable trajectories in response to large perturbations or noise (e.g., **Figures 1c**, **2**, **4**, **5**, **S3**, **S4**, **S7**). Second, in most cases the system described here does not converge to a limit cycle and thus remains formally chaotic outside the training window or in response to different inputs.

For a second and distinct use of the term *stable chaos* see Milani and colleagues[5-6].

Strange nonchaotic attractors. First described by Grebogi et al[7], certain types of dynamical systems display solutions that are complex but do not show sensitive dependence on initial conditions. In these systems, the attractor is a geometrically complex object (e.g., it has a fractal structure), which leads to complex trajectories in phase space and thus to complex time series. Two initially close trajectories, however, don't diverge exponentially in time and thus the LLE is non-positive. Although some features of our trained trajectories might resemble those of strange nonchaotic attractors (namely, complex time series and a non-positive LLE), there are critical differences: our stable trajectories are transient, whereas a strange nonchaotic attractor is an invariant solution (i.e., once the system is in the attractor it will remain in the attractor) with well-defined stationary statistical properties. A second, important difference is that in our case the LLE after training is non-positive for the trained trajectory only—for many networks, the rest of the phase space retains a positive LLE.

Chaotic transients. Under certain circumstances, the transient evolution of a system before reaching a non-chaotic stable solution can be chaotic itself, what is called "transient chaos"[8]. This phenomenon is in many ways the opposite of the stable transients we report: in transient chaos the LLE is locally positive and in our networks the LLE is locally non-positive.

1 Kantz, H. A Robust Method to Estimate the Maximal Lyapunov Exponent of a Time-Series. *Phys Lett A* **185**, 77-87 (1994).
2 Wolf, A. l., Swift, J. B., Swinney, H. L. & Vastano, J. B. Determining Lyapunov Exponents from a Time Series. *Physica D* **16**, 285-317 (1985).
3 Politi, A., Livi, R., Oppo, G. L. & Kapral, R. Unpredictable Behavior in Stable Systems. *Europhys Lett* **22**, 571-576 (1993).
4 Zillmer, R., Brunel, N. & Hansel, D. Very long transients, irregular firing, and chaotic dynamics in networks of randomly connected inhibitory integrate-and-fire neurons. *Physical Review E* **79**, 031909 (2009).
5 Milani, A., Nobili, A. M. & Knezevic, Z. Stable chaos in the asteroid belt. *Icarus* **125**, 13-31 (1997).
6 Milani, A. & Nobili, A. M. An Example of Stable Chaos in the Solar-System. *Nature* **357**, 569-571 (1992).
7 Grebogi, C., Ott, E., Pelikan, S. & Yorke, J. A. Strange Attractors That Are Not Chaotic. *Physica D* **13**, 261-268 (1984).
8 Grebogi, C., Ott, E. & Yorke, J. A. Crises, Sudden Changes in Chaotic Attractors, and Transient Chaos. *Physica D* **7**, 181-200 (1983).