Driven Random Dynamic Reservoir with Homeostatic Variance Control

1 Model Description

1.1 Dynamics

$$y_i^{t+1} = \tanh\left(a_i^t I_i^{t+1}\right) \tag{1}$$

$$I_i^{t+1} = \sum_{j=1}^{N} W_{ij} y_j^t + E_i^{t+1} \tag{2}$$

$$b_i^{t+1} = b_i^t + \epsilon_b \left[\langle y_i \rangle_T - \mu_t\right] \tag{3}$$

$$a_i^{t+1} = a_i^t + \epsilon_a \left[\sigma_t^2 - \left(y_i^t - \overline{y}(t)_i\right)^2\right] \tag{4}$$

$$\overline{y}_i^{t+1} = \epsilon_{\text{trail}} \left[y_i^{t+1} - \overline{y}_i^t\right] \tag{5}$$

We denote by $\langle\cdot\rangle_T$ an average over time and by $\langle\cdot\rangle_P$ over the population.

1.2 Parameters / Settings

W_{ij} is a sparse random matrix with connection probability p. Nonzero entries were drawn from a Gaussian distribution $\mathcal{N}(\mu = 0, \sigma = \sigma_W)$. Diagonal entries were always set to zero.

E_i^t are random vectors of size N with independent entries drawn from a Gaussian distribution $\mathcal{N}(\mu = 0, \sigma = \sigma_{\text{ext}})$. External input is turned off after t_0.

By changing individual gain and bias values a_i and b_i, the homeostatic control tries to drive the activity standard deviation and mean of every cell to the value given by σ_t and μ_t. However, these mechanisms are also switched off after t_0. This is done because we can assume that homeostatic processes would biologically act on much slower timescales than changes in input. Before t_0, we can set ϵ_a and ϵ_b to relatively high values to let homeostasis converge under external drive.

All parameters that were unchanged throughout all simulations are shown in Table 1.

Table 1: Model Parameters

N	p	σ_W	ϵ_b	ϵ_a	ϵ_{trail}	μ_t	t_0
1000	0.1	$w_0/\sqrt{Np}$	$2 \cdot 10^{-4}$	10^{-3}	10^{-4}	0	10^5

2 Derivation of Gain Dynamics from Objective Function

We would like to justify our choice of (4) by deriving an update rule from a suitable objective function F_σ:

$$F_\sigma = \left[\left\langle (y - \langle y \rangle_T)^2 \right\rangle_T - \sigma_t^2\right]^2 \tag{6}$$

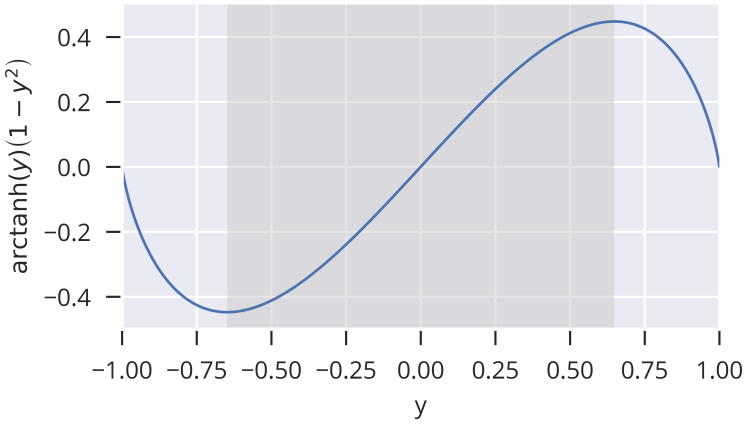

Figure 1: Covariance term of (8). If y is bound within the gray area, positive covariance is ensured.

Using $\Delta a \propto -\frac{\partial F_\sigma}{\partial a}$:

$$\Delta a \propto a^{-1}\left[\sigma_t^2 - \left\langle (y - \langle y\rangle_T)^2 \right\rangle_T\right] \left\langle (y - \langle y\rangle_T) \operatorname{arctanh}(y)\left(1-y^2\right)\right\rangle_T \tag{7}$$

$$= a^{-1}\left[\sigma_t^2 - \left\langle (y - \langle y\rangle_T)^2 \right\rangle_T\right] \operatorname{Cov}\left[y, \operatorname{arctanh}(y)\left(1-y^2\right)\right]_T \tag{8}$$

The sign in the covariance term in (8) determines whether the dynamics given in (4) are recovered, or if the gain adaptation is reversed. Intuitively, it indicates if increasing a broadens or narrows the output distribution. Fig. 1 illustrates that a positive covariance is always ensured if the activity remains within the region highlighted in gray, which is approximately given by $(-0.648, 0.648)$. For symmetry reasons, $\operatorname{Cov}\left[y, \operatorname{arctanh}(y)\left(1-y^2\right)\right]_T \geq 0$ also holds for distributions exceeding this bound if they are symmetric around $y = 0$. In our numerical simulations, this symmetry condition was fulfilled due to the Gaussian distribution of the neural input, which was centered around zero. This allowed us to reduce dynamics to the essential term $\left[\sigma_t^2 - \left\langle (y - \langle y\rangle_T)^2 \right\rangle_T\right]$.

3 Results

Exemplary results are shown in Fig. 2.

4 Mean Field Approximation

We would like to find an approximate relation between the gain resulting from homeostasis and the input and target variance. If we linearize the neural activation function and take an average over time, we get

$$\left\langle y_i^2 \right\rangle_T = a_i^2 \left\langle \left(\sum_{j=1}^{N} W_{ij} y_j + E_i \right)^2 \right\rangle_T \tag{9}$$

$$= a_i^2 \left\langle \left(\sum_{j=1}^{N} W_{ij} y_j \right)^2 \right\rangle_T + a_i^2 E_i^2 \tag{10}$$

$$= a_i^2 \sum_{j,k=1}^{N} W_{ij} W_{ik} \left\langle y_j x_k \right\rangle_T + a_i^2 E_i^2 \ . \tag{11}$$

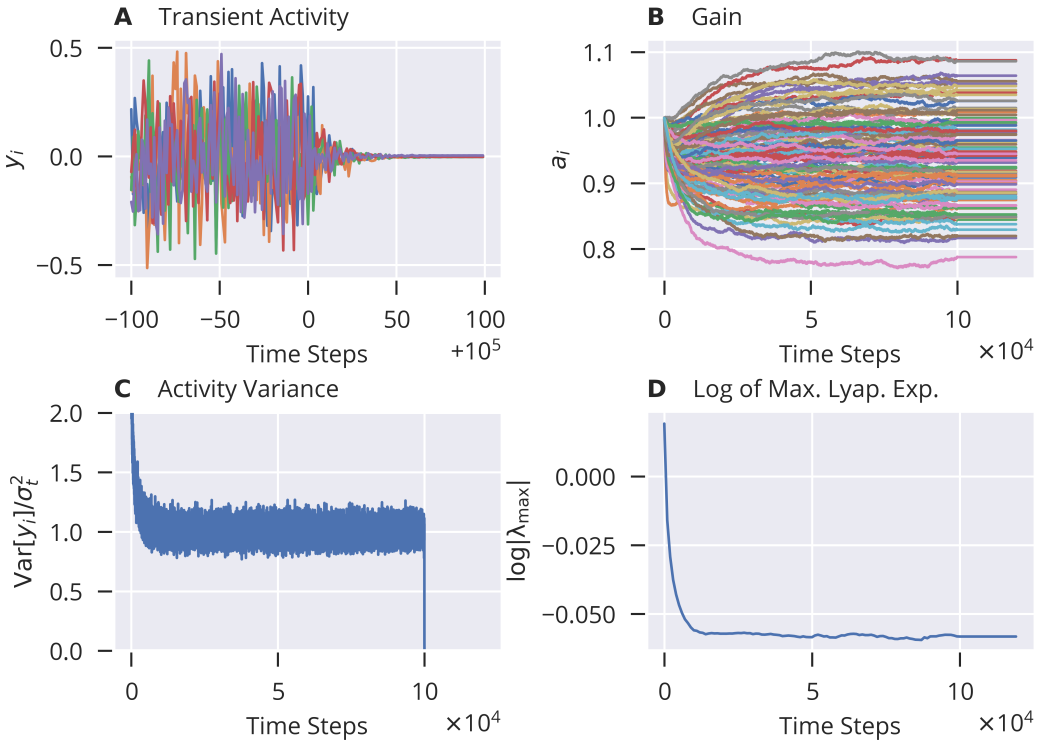

Figure 2: **A**: Sample of activity within $[t_0 - 100, t_0 + 100]$. **B**: Gain dynamics of $N/10$ exemplary neurons. **C**: Population mean of squared activity, normalized by target variance. **D**: Log. of largest absolute value of eigenvalues of $a_i^t W_{ij}$. Standard parameters are as given in Table 1. Furthermore, $\sigma_{\text{ext}} = 0.1$, $\sigma_t = 0.2$, total runtime $n_t = 1.2 \cdot 10^5$.

If we assume that the system is in a chaotic state we can set $\langle y_j x_k \rangle_T = 0$ for $j \neq k$. This leads to

$$\langle y_i^2 \rangle_T = a_i^2 \left(\sum_{j=1}^{N} W_{ij}^2 \langle y_j^2 \rangle_T + \sigma_{\text{ext}}^2 \right) \tag{12}$$

where we have assumed $\langle E_i \rangle_T = 0$ for all i.

By design, our homeostatic mechanism fixes all $\langle y_i^2 \rangle_T$ to σ_t^2. Thus,

$$\sigma_t^2 = a_i^2 \left(\sigma_t^2 \sum_{j=1}^{N} W_{ij}^2 + \sigma_{\text{ext}}^2 \right) \tag{13}$$

$$a_i = \left(\sum_{j=1}^{N} W_{ij}^2 + \sigma_{\text{ext}}^2 / \sigma_t^2 \right)^{-1/2} . \tag{14}$$

Since W_{ij} is a random Gaussian matrix with variance $\sigma_W^2 = w_0^2/(Np)$, $\sum_{j=1}^{N} W_{ij}^2$ follows a χ^2 - distribution with variance $\frac{2Np\sigma_W^2}{N^2p^2} = \frac{2\sigma_W^2}{Np}$. For $N \to \infty$, its variance vanishes and consequently, all a_i converge to the same value, namely

$$a = \left(\sigma_W^2 + \sigma_{\text{ext}}^2 / \sigma_t^2 \right)^{-1/2} . \tag{15}$$

This equation predicts that a should not change if the ratio between target and input variance remains constant. We ran a parameter sweep over σ_{ext} and σ_t with a network of $N = 1000$ neurons and looked at the resulting distribution of gains and the maximal Lyapunov exponent. Importantly, this approximation suggests that the network should

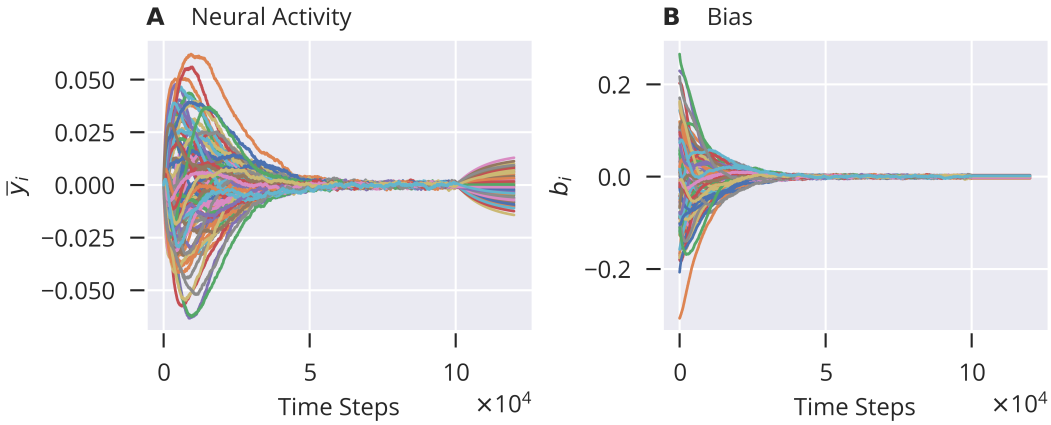

Figure 3: **A**: Dynamics of population mean of neural activity. **B**: Bias dynamics of $N/10$ exemplary neurons.

tune into a subcritical configuration for any non-vanishing external input. Even though this is not strictly verified in the numerical simulation, see Fig. 4A, it holds for the majority of $\sigma_{\text{ext}}/\sigma_t$ combinations.

4.1 Self-Consistency Equation with Full Activation Function

If we do not restrict ourself to a linearized version of the activation function, (9) becomes

$$\left\langle y_i^2 \right\rangle_T = \left\langle \tanh\left(a_i \left(\sum_{j=1}^{N} W_{ij} y_j + E_i \right) \right)^2 \right\rangle_T . \tag{16}$$

Again, assuming that $\langle y_j x_k \rangle_T = 0$ for $j \neq k$ and $N \to \infty$, the sum of weight and inputs follows a Gaussian distribution $\mathcal{N}\left(\mu = 0, \sigma = \sqrt{\sigma_W^2 \sigma_t^2 + \sigma_{\text{ext}}^2}\right)$. Replacing the time average with an integral over the distribution of the input, we arrive at the full self-consistency equation

$$\sigma_t^2 = \int_{-\infty}^{\infty} \tanh^2(ay)\, \mathcal{N}\left(y, \mu = 0, \sigma = \sqrt{\sigma_W^2 \sigma_t^2 + \sigma_{\text{ext}}^2}\right) \mathrm{d}y \tag{17}$$

As a second order approximation with respect to tanh, we can use $\tanh^2(y) \approx y^2 - \frac{2}{3} y^4$ in the integral, which gives

$$\begin{aligned}
\sigma_t^2 &\approx \frac{1}{a\sqrt{2\pi}\sigma_{\text{total}}} \int_{-\infty}^{\infty} \left(y^2 - \frac{2}{3} y^4 \right) \exp\left(\frac{-y^2}{2a^2 \sigma_{\text{total}}^2} \right) \mathrm{d}y && (18)\\
&= a^2 \sigma_{\text{total}}^2 - 2a^4 \sigma_{\text{total}}^4 && (19)\\
&= a^2 \left(\sigma_W^2 \sigma_t^2 + \sigma_{\text{ext}}^2 \right) - 2a^4 \left(\sigma_W^2 \sigma_t^2 + \sigma_{\text{ext}}^2 \right) && (20)
\end{aligned}$$

where we have defined $\sigma_{\text{total}} = \sqrt{\sigma_W^2 \sigma_t^2 + \sigma_{\text{ext}}^2}$ for convenience. Note that neglecting the second term on the right hand side of (20) reduces the equation to (15). Since this first order approximation does not explain the $a = 1$ transition traced in Fig. 4B as a dashed blue line, we were particularly interested in the solution of (20) for $a = 1$. Moreover, we set $\sigma_W = 1$ as in the simulations. We therefore had to solve

$$\sigma_t^4 + \sigma_{\text{ext}}^4 + 2\sigma_t^2 \sigma_{\text{ext}}^2 - \frac{\sigma_{\text{ext}}^2}{2} = 0 . \tag{21}$$

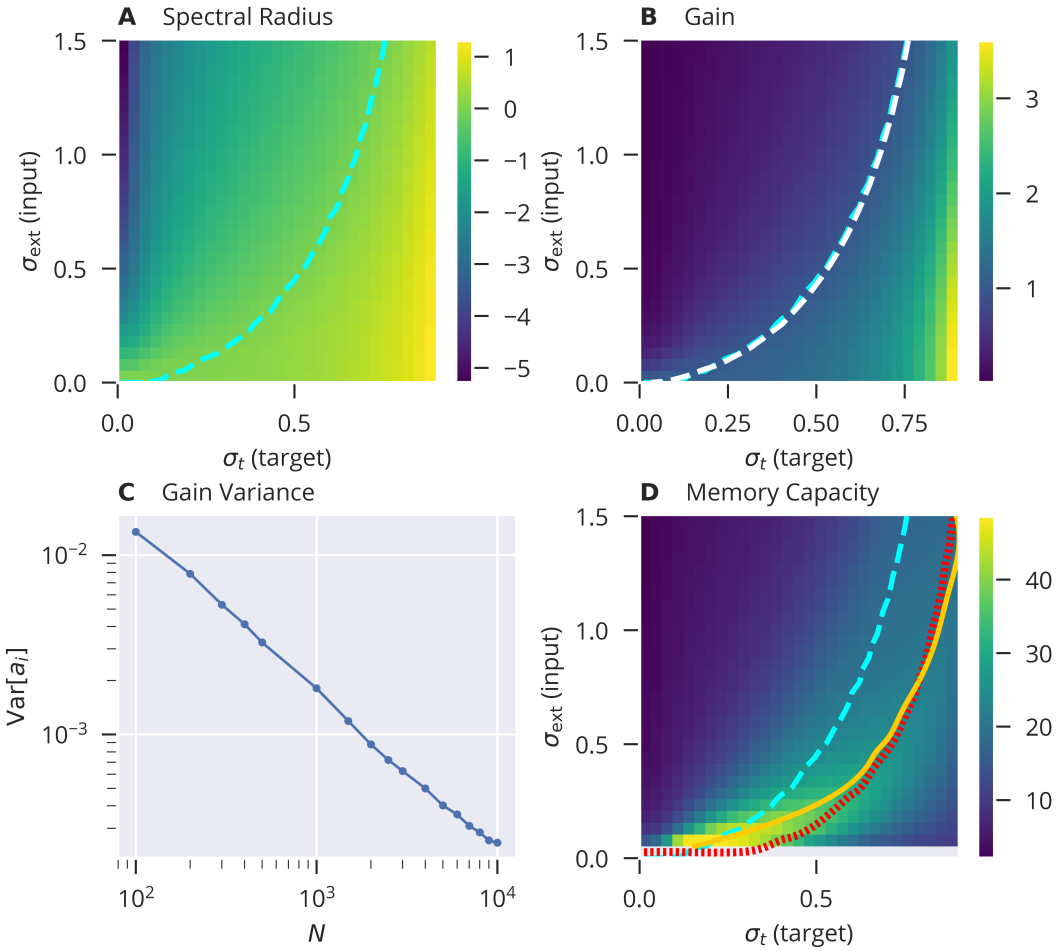

Figure 4: Parameter sweep, run on a network with $N = 1000$ (except **C**). **A**: Log of largest absolute value of eigenvalues of $a_i W_{ij}$. turquois dashed line marks the zero transition. **B**: $\langle a_i \rangle_P$. White dashed line marks the $\langle a_i \rangle_P = 1$ transition. Turquois line as in **A**. **C**: $\left\langle (a_i - \langle a_i \rangle_P)^2 \right\rangle_P$ as a function of N. **D**: Memory Capacity (color coded), zero transition of largest absolute eigenvalue (turquois line), maximal memory capacity for a given standard deviation of external drive (orange). Red line marks the loss of the echo state property.

This equation can be rearranged to

$$\left[\left(\sigma_{\text{ext}} - \frac{1}{\sqrt{8}} \right)^2 + \sigma_t^2 - \frac{1}{8} \right] \left[\left(\sigma_{\text{ext}} + \frac{1}{\sqrt{8}} \right)^2 + \sigma_t^2 - \frac{1}{8} \right] = 0 \,. \tag{22}$$

From this form we can see that the solution set consists of two circles with radius $1/\sqrt{8}$ in the σ_t, σ_{ext} space, centered around $\sigma_t = 0$, $\sigma_{\text{ext}} = \pm 1/\sqrt{8}$. Fig. 6 shows a comparison between this approximation and the simulation.

5 Behavior for Input with Correlations Across Time/Population

Test with e.g. Ornstein-Uhlenbeck Process and/or dominant princ. comp. in input space

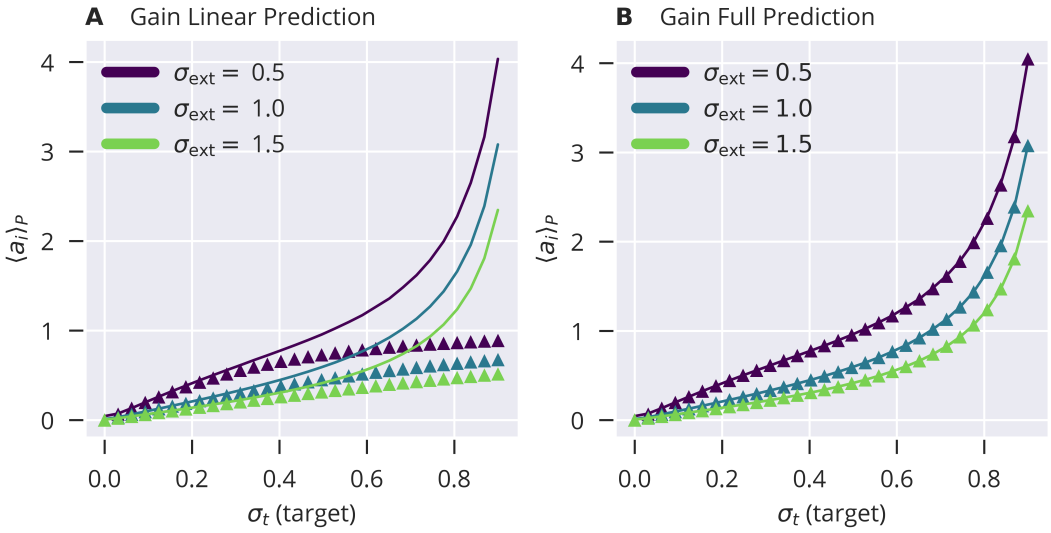

Figure 5: **A**: Prediction of (15) (triangles) and numerical result (lines) of $\langle a_i \rangle_P$. **B**: $\langle a_i \rangle_P$ from the simulation matches the numerically determined solution of (17).

6 Related Literature

The effect of homoestatic mechanisms onto echo state network performance has been investigated in several works. A common approach derives dynamic adaptation rules from the minimization of a functional (usually KL-divergence) of an empirical estimate of each neuron's output distribution and a target PDF [Triesch (2007); Schrauwen et al. (2008); Boedecker et al. (2009)].

On the relation between Echo State Property and Input Signal Properties: Manjunath and Jaeger (2013)

References

Boedecker, J., Obst, O., Mayer, N. M., and Asada, M. (2009). Initialization and self-organized optimization of recurrent neural network connectivity. *HFSP Journal*, 3(5):340–349.

Manjunath, G. and Jaeger, H. (2013). Echo State Property Linked to an Input: Exploring a Fundamental Characteristic of Recurrent Neural Networks. *Neural Computation*, 25(3):671–696. PMID: 23272918.

Schrauwen, B., Wardermann, M., Verstraeten, D., Steil, J. J., and Stroobandt, D. (2008). Improving reservoirs using intrinsic plasticity. *Neurocomputing*, 71(7-9):1159–1171.

Triesch, J. (2007). Synergies Between Intrinsic and Synaptic Plasticity Mechanisms. *Neural Computation*, 19(4):885–909. PMID: 17348766.

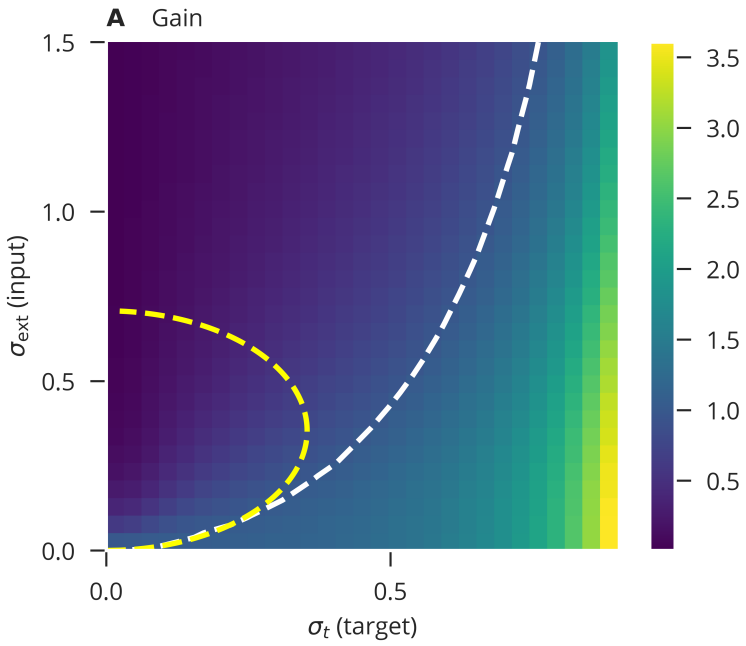

Figure 6: Comparison of the $\langle a_i \rangle_P = 1$ transition (blue) shown in Fig. 4B and the second order approximate solution (yellow) given by (22).