Local homeostatic regulation of the spectral radius of echo-state networks

Fabian Schubert[1*], Claudius Gros[1]

1 Institute for Theoretical Physics, Goethe University Frankfurt, Frankfurt am Main, Hesse, Germany

* fschubert@itp.uni-frankfurt.de

Abstract

Recurrent cortical network dynamics plays a crucial role for sequential information processing in the brain. While the theoretical framework of reservoir computing provides a conceptual basis for the understanding of recurrent neural computation, it often requires manual adjustments of global network parameters, in particular of the spectral radius of the recurrent synaptic weight matrix. Being a mathematical and relatively complex quantity, the spectral radius is not readily accessible to biological neural networks, which generally adhere to the principle that information about the network state should either be encoded in local intrinsic dynamical quantities (e.g. membrane potentials), or transmitted via synaptic connectivity. We present two synaptic scaling rules for echo state networks that solely rely on locally accessible variables. Both rules work online, in the presence of a continuous stream of input signals. The first rule, termed flow control, is based on a local comparison between the mean squared recurrent membrane potential and the mean squared activity of the neuron itself. It is derived from a global scaling condition on the dynamic flow of neural activities and requires the separability of external and recurrent input currents. We gained further insight into the adaptation dynamics of flow control by using a mean field approximation on the variances of neural activities that allowed us to describe the interplay between network activity and adaptation as a two-dimensional dynamical system. The second rule that we considered, variance control, directly regulates the variance of neural activities by locally scaling the recurrent synaptic weights. The target set point of this homeostatic mechanism is dynamically determined as a function of the variance of the locally measured external input. This functional relation was derived from the same mean-field approach that was used to describe the approximate dynamics of flow control.

The effectiveness of the presented mechanisms was tested numerically using different external input protocols. The network performance after adaptation was evaluated by training the network to perform a time delayed XOR operation on binary sequences. As our main result, we found that flow control can reliably regulate the spectral radius under different input statistics, but precise tuning is negatively affected by interneural correlations. Furthermore, flow control showed a consistent task performance over a wide range of input strengths/variances. Variance control, on the other side, did not yield the desired spectral radii with the same precision. Moreover, task performance was less consistent across different input strengths.

Given the better performance and simpler mathematical form of flow control, we concluded that a local control of the spectral radius via an *implicit* adaptation scheme is a realistic alternative to approaches using classical "set point" homeostatic feedback controls of neural firing.

Author summary

How can a neural network control its recurrent synaptic strengths such that network dynamics are optimal for sequential information processing? An important quantity in this respect, the spectral radius of the recurrent synaptic weight matrix, is a non-local quantity. Therefore, a direct calculation of the spectral radius is not feasible for biological networks. However, we show that there exist a local and biologically plausible adaptation mechanism, flow control, which allows to control the recurrent weight spectral radius while the network is operating under the influence of external inputs. Flow control is based on a theorem of random matrix theory, which is applicable if inter-synaptic correlations are weak. We apply the new adaption rule to echo-state networks having the task to perform a time-delayed XOR operation on random binary input sequences. We find that flow-controlled networks can adapt to a wide range of input strengths while retaining essentially constant task performance.

Introduction

Echo state networks are a class of recurrent neural networks that can be easily trained on time dependent signal processing tasks. Once externally induced, the recurrent activity provides a reservoir of non-linear transformations [1], both in time and space, that may be utilized by a linear readout unit. Training the linear output units makes an echo state network a highly effective prediction machine [1, 2]. However, achieving optimal performance needs fine tuning of the network properties, in particular the spectral radius $R = |\Lambda_{max}|$. It is closely related to the so-called echo state property [3]. While different equivalent formulations exist, a common definition for the echo-state property posits that specific input sequences lead to uniquely determined sequences of neural activity. This presumption implies that perturbations of the network's initial conditions decay over time. Theoretical approaches linking the echo-state property to network parameters have led to results that can be divided into necessary and sufficient conditions for obtaining the echo state property. In particular, early works on echo state networks already established coarse bounds for the choice of the recurrent weight matrix, namely $R < 1$ as a necessary condition and $\sigma_{\max} < 1$ as a sufficient condition, where $\sigma_{\max}$ is the largest singular value [3]. In practice, setting $R \approx 1$ has proven to be sufficient for achieving the echo state property and good network performance [1, 3, 4]. This observation was supported by a theoretical result stating that for $R < 1$, the probability for the network dynamics being contracting tends to 1 for a large number of nodes [5]. If, however, R is significantly smaller than 1, information is lost too fast over time, which is detrimental for tasks involving sequential memory. A spectral radius of about one is hence best [6], in the sense that it provides a maximal memory capacity if the network operates in a linear regime [7, 8]. Similarly, Boedecker et al. found the best memory capacity to be close to the largest Lyapunov exponent being equal to zero [9], which can be understood as a generalization of the aforementioned results for nonlinear dynamics.

Aside their applications as efficient machine learning algorithms, echo state networks are potentially relevant as models of information processing in the brain [10–12].

An extension to layered ESN architectures was presented by Gallicchio and Micheli [13], which bears greater resemblance to the hierarchical structure of cortical networks than the usual shallow ESN architecture. This line of research illustrates the importance to examine whether local and biological plausible principles exist that would allow to tune the properties of the neural reservoir to the "edge of chaos" [14], in particular when a continuous stream of inputs is present. The rule has to be independent of the network topology, which is not a locally accessible information, and

of the distribution of synaptic weights.

Generally speaking, strategies for optimizing ESN hyperparameters can be divided in two categories: supervised and unsupervised methods, where the first one utilizes an error signal, while the latter uses only information contained within the network dynamics. Our work is concerned with unsupervised optimization, independent of the actual task at hand. Several studies investigated local unsupervised intrinsic plasticity rules for improving neural reservoirs [15–17], usually by defining a target output distribution that each neuron attempts to reproduce by changing neural gains and biases, or a target distribution for the local optimization of information transmission [18, 19]. In general, however, it is difficult to formulate an explicit analytic theory that quantifies the relation between intrinsic neural parameters and global network dynamics, or corresponding parameters, as, for example, the spectral radius. One issue with choosing a fixed target distribution for optimizing network performance is that changing the external input, e.g. its mean or variance, affects the configuration of gains and biases resulting from the intrinsic plasticity mechanism. It can happen that the induced changes decrease performance by driving the network away from the critical point, which is known to be a beneficial dynamical state for sequence learning tasks [20–25].

Here, we propose and compare two unsupervised homeostatic mechanisms, which we termed flow control and variance control. Both are supposed to regulate the mean and variance of neuronal firing such that the network works in an optimal regime with regard to sequence learning tasks. The mechanisms act on two sets of node-specific parameters, the biases b_i and the neural gain factors a_i. This approach can be considered a realization of dual homeostasis, which has been investigated previously with respect to a stable control of the mean and the variance of neural activity [26]. In this framework, the adaptation of the bias acts an intrinsic plasticity for the control of the internal excitability of a neuron [27–29], while the gain factors functionally correspond to a synaptic scaling of the recurrent weights [30–32].

We restricted ourselves to 'biologically plausible' adaption mechanisms, viz mechanisms for which the dynamics of all variables is local, i.e., bound to a specific neuron. Additional variables enter only when locally accessible. In a strict sense, this implies that local dynamics is determined exclusively by the dynamical variables of the neuron and by information about the activity of afferent neurons. Being less restrictive, one could claim that it should also be possible to access aggregate or 'mean-field' quantities that average a property of interest with respect to the population. For example, nitric oxide is a diffusive neurotransmitter that can act as a measure for the population average of neural firing rates [33].

In this study, we used standard discrete-time dynamics, which is markovian, with information about past states being stored dynamically. Following a general description of the model and the network, we present the flow control: it involves a normalization constraint in phase space (Sect. *Autonomous spectral radius regulation*). We evaluate the performance of networks that were subject to adaptation in Sect. *XOR-memory recall*, using a nonlinear recall task based on a sequential binary XOR-operation. Finally, we discuss the influence of node-to-node cross-correlations within the population in Sect. *Input induced correlations*.

Model

A full description of the network model and parameters can be found in the methods section. We briefly introduce the network dynamics as

$$x_i(t) = a_i \sum_{j=1}^{N} W_{ij} y_j(t-1) + I_i(t), \qquad y_i(t) = \tanh\left(x_i(t) - b_i\right) \ . \tag{1}$$

Each neuron's membrane potential x_i consists of a recurrent contribution $a_i \sum_{j=1}^{N} W_{ij} y_j(t-1)$ and some external input $I_i(t)$. The biases b_i were subject to the following homeostatic adaptation:

$$b_i(t) = b_i(t-1) + \epsilon_{\mathrm{b}} \left[y_i(t) - \mu_{\mathrm{t}} \right] \ . \tag{2}$$

Here, μ_{t} defines a target for the average activity.

For a_i, we considered two different forms of update rules. The first, flow control, is given by

$$a_i(t) = a_i(t-1)\Big[1 + \epsilon_{\mathrm{a}} \Delta R_i(t)\Big], \qquad \Delta R_i(t) = R_{\mathrm{t}}^2 |y_i(t-1)|^2 - |x_{\mathrm{r},i}(t)|^2 \ . \tag{3}$$

The parameter R_{t} defines the desired target spectral radius. We also considered an alternative global update rule where $\Delta R_i(t)$ was given by

$$\Delta R_i(t) = \frac{1}{N}\Big[R_{\mathrm{t}}^2 \, ||\mathbf{y}(t-1)||^2 - ||\mathbf{x}_{\mathrm{r}}(t)||^2\Big] \ . \tag{4}$$

However, since this is a non-local rule, it only served as a comparative model to Eq. (3) when we investigated the effectiveness of the adaptation mechanism.

The second rule, variance control, has the form

$$a_i(t) = a_i(t-1) + \epsilon_{\mathrm{a}} \left[\sigma_{\mathrm{t},i}^2(t) - \left(y_i(t) - \mu_i^{\mathrm{y}}(t)\right)^2 \right] \tag{5}$$

$$\sigma_{\mathrm{t},i}^2(t) = 1 - \sqrt{1 + 2R_{\mathrm{t}}^2 y_i(t)^2 + 2\sigma_{\mathrm{ext},i}^2(t)} \tag{6}$$

$$\mu_i^{\mathrm{y}}(t) = \mu_i^{\mathrm{y}}(t-1) + \epsilon_\mu \left[y_i(t) - \mu_i^{\mathrm{y}}(t-1) \right] \tag{7}$$

$$\sigma_{\mathrm{ext},i}^2(t) = \sigma_{\mathrm{ext},i}^2(t-1) + \epsilon_\sigma \left[\left(I_i(t) - \mu_{\mathrm{ext},i}(t)\right)^2 - \sigma_{\mathrm{ext},i}^2(t-1) \right] \tag{8}$$

$$\mu_{\mathrm{ext},i}(t) = \mu_{\mathrm{ext},i}(t-1) + \epsilon_\mu \left[I_i(t) - \mu_{\mathrm{ext},i}(t-1) \right] \ . \tag{9}$$

Eq. (5) drives the average variance of each neuron towards a desired target variance. While Eq. (7)–(9) are simply local trailing averages of the mean activity and the mean and variance of the external input, Eq. (6) is an analytic expression that was derived from a mean field approximation which is explained in S2 Appendix. Similar to flow control, we also considered a non-local version for comparative reasons where (6) was replaced with

$$\sigma_{\mathrm{t},i}^2(t) = 1 - \sqrt{1 + 2R_{\mathrm{t}}^2 ||\mathbf{y}(t)||^2/N + 2\sigma_{\mathrm{ext},i}^2(t)} \ . \tag{10}$$

Before proceeding to the results, we explain the mathematical background of the proposed adaptation rules in more detail.

Autonomous spectral radius regulation

We would like to explain the theoretical framework that led us to propose Eq. (3) and (6) as a regulatory mechanism for the spectral radius of the recurrent weight matrix.

The circular law of random matrix theory states that the eigenvalues λ_j are distributed uniformly on the complex unit circle if the elements of a real $N \times N$ matrix are drawn from distributions having zero mean and standard deviation $1/\sqrt{N}$ [34]. Noting that the internal weight matrix W_{ij} has $p_{\mathrm{r}} N$ non-zero elements per row, the circular law implies that the spectral radius of $a_i W_{ij}$, the maximum of $|\lambda_j|$, is unity when the synaptic scaling factors a_i are set uniformly to $1/\sigma_{\mathrm{w}}$. Our goal is to investigate adaption rules for the synaptic rescaling factors that are based on dynamic quantities, which includes the membrane potential x_i, the neural activity y_i and the input I_i.

A $N \times N$ matrix with i.i.d. entries with zero mean and $1/N$ variance will have a spectral radius of one as $N \to \infty$. Rajan and Abbott [35] investigated the case where

the statistics of the columns of the matrix differ in their means and variances: given row-wise E-I balance for the recurrent weights, the square of the spectral radius of a random $N \times N$ matrix whose columns have variances σ_i^2 is $N \left\langle \sigma_i^2 \right\rangle_i$ for $N \to \infty$. Since the eigenvalues are invariant under transposition, this result can also be applied to our case, where node-specific gain factors a_i were applied to each *row* of the recurrent weights. Thus, the spectral radius R_a of random matrices $a_i W_{ij}$ is approximately given by

$$R_\mathrm{a}^2 \approxeq \frac{1}{N} \sum_i R_{\mathrm{a},i}^2, \qquad R_{\mathrm{a},i}^2 := a_i^2 \sum_j (W_{ij})^2 \tag{11}$$

for large N, assuming that the distribution underlying $\widehat{W}$ has zero mean. Note that R_a^2 can also be expressed in terms of the Frobenius norm $\left\| \widehat{W}_\mathrm{a} \right\|_\mathrm{F}$ via

$$R_\mathrm{a}^2 \approxeq \left\| \widehat{W}_\mathrm{a} \right\|_\mathrm{F}^2 / N \,. \tag{12}$$

Note that, before proceeding to further investigations, we numerically tested this approximation for $N = 500$ and heterogeneous random sets of a_i drawn from a uniform $[0, 1]$-distribution and found a very close match to the actual spectral radii (1-2% relative error).

Considering the $R_{\mathrm{a},i}$ as per site estimates for the spectral radius, one can use the generalized circular law (11) to regulate R_a on the basis of local adaption rules, one for every a_i.

In the particular case of flow control, this was done using a comparison between the variance of neural activity that is present in the network and the resulting recurrent contribution to the membrane potential. A more detailed explanation is given in Sect. *Spectral radius, singular values and global Lyapunov exponents* and S1 Appendix. In short, we propose that,

$$\left\langle \left|\left|\mathbf{x}_\mathrm{r}(t)\right|\right|^2 \right\rangle_\mathrm{t} \approxeq R_\mathrm{a}^2 \left\langle \left|\left|\mathbf{y}(t-1)\right|\right|^2 \right\rangle_\mathrm{t}, \tag{13}$$

where $x_{\mathrm{r},i}$ is the recurrent contribution to the membrane potential x_i. This stationarity condition led to the adaptation rule given in Eq. (3).

Instead of directly imposing Eq. (13) via an appropriate adaptation mechanism, we also considered the possibility of transferring this condition into a set point for the variance of neural activities as a function the external driving. To do so, we used a mean-field approach to describe the effect of the recurrent input onto the resulting neural activity variance. A detailed description is given in S2 Appendix. This led to the update rule given by Eq. (5)–(9).

Results

Flow control

Flow control, see Eq. (3), robustly led to the desired spectral radius R_t when correlations were weak. It is of particular interest that an internal network parameter, the spectral radius, can be regulated in the continuing present of external inputs $\mathbf{I}(t)$. For correlated input streams, as listed in Sect. *Input protocols*, substantial inter-neuron correlation may be induced. In this case a stationary state was still attained, but with a spectral radius deviating to a certain extent from the parameter R_t entering (3).

In Fig. 1, we present a simulation for a network with $N = 500$ sites, a connection probability $p_\mathrm{r} = 0.1$ and an adaption rate $\epsilon_\mathrm{a} = 10^{-3}$. The standard deviation of the external driving is $\sigma_\mathrm{ext} = 0.5$. Numerically, we found that the time needed to converge to the stationary states depends substantially on R_t, slowing down when the spectral

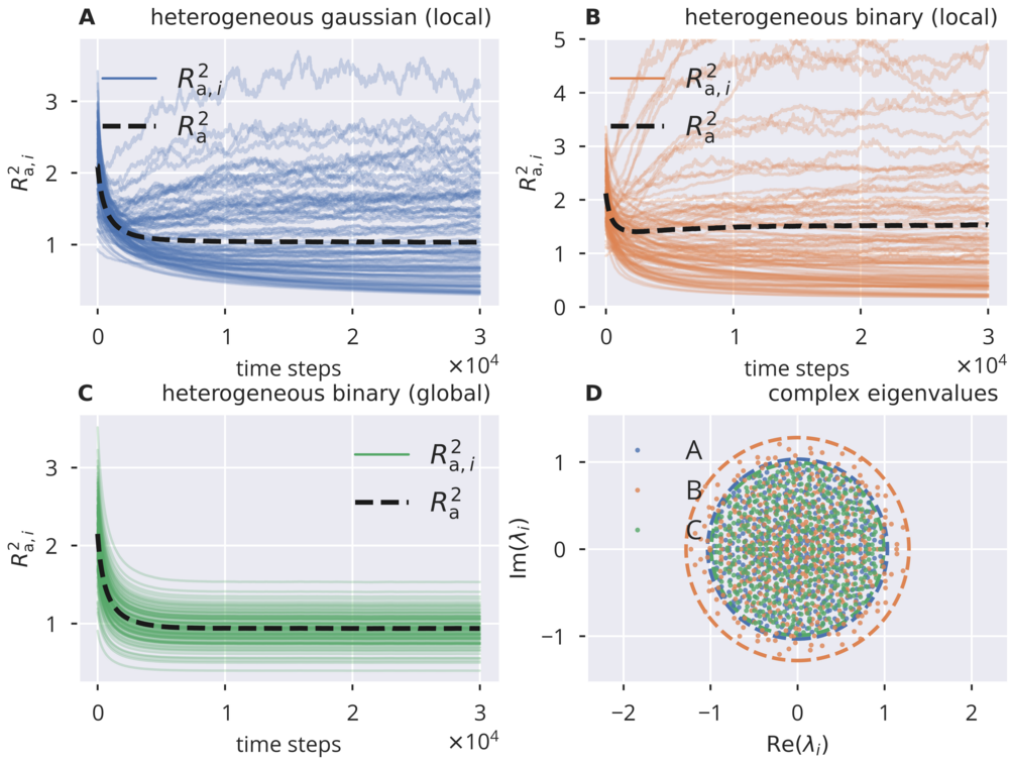


Fig 1. Online spectral radius regulation using flow control. The spectral radius R_a and the respective local estimates $R_{a,i}$ as defined by (11). For the input protocols see Sect. *Input protocols*, for the local and global adaption rules (3) and (4). **A**: Dynamics of $R^2_{a,i}$ and R^2_a, in the presence of heterogeneous independent Gaussian inputs. Local adaptation. **B**: Heterogeneous identical binary input, local adaptation. **C**: Heterogeneous identical binary input, global adaptation. **D**: Distribution of eigenvalues of the corresponding affective synaptic matrices $\widehat{W}_a$. Circles denote the respective spectral radii. Local adaption converges to a spectral radius that is too large (the target is unity), when the driving input is strongly correlated.

radius becomes small. It is then advantageous, as we have done, to scale the adaption rate ϵ_a inversely with the trailing average $\bar{x}_r^2$ of $||x_r||^2$, viz as $\epsilon_a \to \epsilon_a/\bar{x}_r^2$.

For all input protocols and adaption rules, the spectral radius was found to converge to a finite value. The desired value, in this case $R_t = 1$, was however not attained when a strongly correlated input interferes with local adaption, see Eq. (3). The local estimates $R_{a,i}$ for the spectral radius show a substantial heterogeneity, which was reduced when adaption is global, using (16).

Variance control

In comparison, variance control, shown in Fig. 2, resulted in a spectral radius that was consistently off by a factor of approx. 20%. Note that we had to introduce a lower bound of 0 for a_i, since some gain factors would otherwise have become negative. Overall, in comparison to flow control, variance control did not exhibit the same level of precision in tuning the system towards a desired spectral spectral radius.

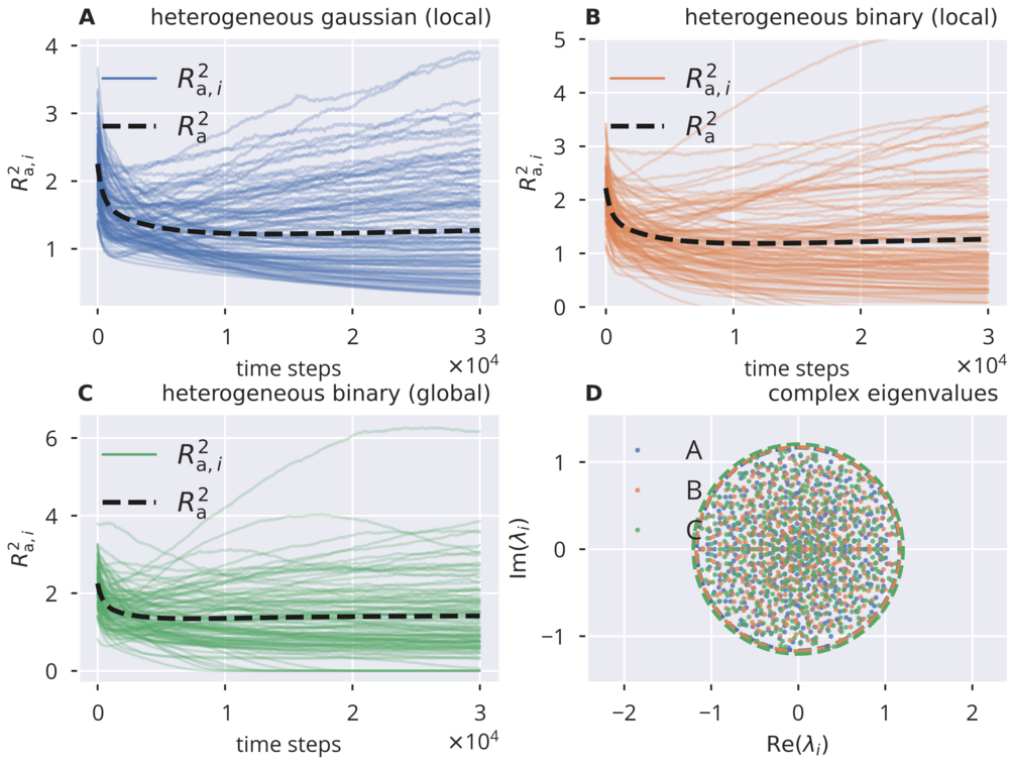


Fig 2. Online spectral radius regulation using variance control. The spectral radius R_a and the respective local estimates $R_{\mathrm{a},i}$ as defined by (11). For the input protocols see Sect. *Input protocols*, for the local and global adaption rules (5) and (10). **A**: Dynamics of $R^2_{\mathrm{a},i}$ and R^2_a, in the presence of heterogeneous independent Gaussian inputs. Local adaptation. **B**: Heterogeneous identical binary input, local adaptation. **C**: Heterogeneous identical binary input, global adaptation. **D**: Distribution of eigenvalues of the corresponding affective synaptic matrices $\widehat{W}_\mathrm{a}$. Circles denote the respective spectral radii. Local adaption converges to a spectral radius that is too large (the target is $R_\mathrm{t} = 1$) under all input protocols.

Spectral radius, singular values and global Lyapunov exponents

Apart from the spectral radius R_a of the matrix $\widehat{W}_\mathrm{a}$, one may consider the relation between the adaptation dynamics and the respective singular values σ_i of $\widehat{W}_\mathrm{a}$. We recall that the spectrum of $\hat{U}_\mathrm{a} = \widehat{W}_\mathrm{a}^\dagger \widehat{W}_\mathrm{a}$ is given by the squared singular values, σ_i^2, and that the relation $||\mathbf{x}_\mathrm{r}||^2 = \mathbf{y}^\dagger \widehat{W}_\mathrm{a}^\dagger \widehat{W}_\mathrm{a} \mathbf{y}$ holds. Now, assume that the time-averaged projection of neural activity $\mathbf{y} = \mathbf{y}(t)$ onto all eigenvectors of $U(t)$ is approximately the same, that is, there is no preferred direction of neural activity in phase space. From this idealized case, it follows that the time average of the recurrent contribution to the membrane potential can be expressed with

$$\left\langle \, ||\mathbf{x}_\mathrm{r}||^2 \, \right\rangle_\mathrm{t} \approx \frac{\left\langle \, ||\mathbf{y}||^2 \, \right\rangle_\mathrm{t}}{N} \sum_i \sigma_i^2 = \frac{\left\langle \, ||\mathbf{y}||^2, \right\rangle_\mathrm{t}}{N} \sum_{i,j} \left(a_i W_{ij} \right)^2 \tag{14}$$

as the rescaled average of the σ_i^2. For the second equation, we used the fact that the $\sum_i \sigma_i^2$ equals the sum of all matrix elements squared [36, 37]. With (11), one finds that (14) is equivalent to $\left\langle \, ||\mathbf{x}_\mathrm{r}||^2, \right\rangle_\mathrm{t} = R_a^2 \left\langle \, ||\mathbf{y}||^2, \right\rangle_\mathrm{t}$ and hence to the flow condition (13). This result can be generalized, as done in S1 Appendix, to the case that the neural

activities have inhomogeneous variances, while still being uncorrelated with zero mean. We then show that the stationarity condition leads to a spectral radius of (approximately) unity.

It is worthwhile to note that the singular values of $\hat{U}_{\mathrm{a}} = \widehat{W}_{\mathrm{a}}^{\dagger}\widehat{W}_{\mathrm{a}}$ do exceed unity when $R_{\mathrm{a}} = 1$. More precisely, for a random matrix with i.i.d. entries, one finds in the limit of large N that the largest singular value is given by $\sigma_{\max} = 2R_{\mathrm{a}}$, in accordance with the Marchenko-Pastur law for large random matrices [38]. Consequently, directions in phase space exist, in which the norm of the phase space vector is elongated by factors greater than one. Still, this does not contradict the fact that a unit spectral radius coincides with a transition to chaos for the non-driven case. The reason is that the global Lyapunov exponents are given by

$$\lim_{n\to\infty} \frac{1}{2n} \ln\left(\left(\widehat{W}_{\mathrm{a}}^{n}\right)^{\dagger} \widehat{W}_{\mathrm{a}}^{n}\right) \tag{15}$$

which eventually converge to $\ln\|\lambda_i\|$, see S1 Fig and [39], where λ_i is the ith eigenvalue of $\widehat{W}_{\mathrm{a}}$. The largest singular value of the nth power of a random matrix with a spectral radius R_{a} scales like R_{a}^{n} in the limit of large powers n. The global Lyapunov exponent goes to zero as a consequence when $R_{\mathrm{a}} \to 1$.

Spectral radius adaption dynamics

For an understanding of the spectral radius adaption dynamics of flow control, it is of interest to examine the effect of using the global adaption constraint

$$\Delta R_i(t) = \frac{1}{N}\left[R_{\mathrm{t}}^{2}\,||\mathbf{y}(t-1)||^{2} - ||\mathbf{x}_{\mathrm{r}}(t)||^{2}\right] \tag{16}$$

in (3). The spectral radius condition (13) is then enforced directly, with the consequence that (16) is stable and precise even in the presence of correlated neural activities (see Fig. 1C). This rule, while not biologically plausible, provides an opportunity to examine the dynamical flow, besides the resulting state. There are two dynamical variables, $a \equiv a_i = a_j$, which, for the sake of simplicity, are assumed to be homogeneous, and the activity variance $\sigma_{\mathrm{y}}^{2} = ||\mathbf{y}||^{2}/N$. The evolution of $(a, \sigma_{\mathrm{y}}^{2})$ resulting from the global rule (4) is shown in Fig. 3.

For the flow, $\Delta a = a(t+1) - a(t)$ and $\Delta\sigma_{\mathrm{y}}^{2} = \sigma_{\mathrm{y}}^{2}(t) - \sigma_{\mathrm{y}}^{2}(t-1)$, the approximation

$$\Delta a = \epsilon_{\mathrm{a}} a \left(1 - a^{2}\right)\sigma_{\mathrm{y}}^{2} \tag{17}$$

$$\Delta\sigma_{\mathrm{y}}^{2} = 1 - \sigma_{\mathrm{y}}^{2} - \frac{1}{\sqrt{1 + 2a^{2}\sigma_{\mathrm{y}}^{2} + 2\sigma_{\mathrm{ext}}}} \tag{18}$$

is obtained when using $R_{\mathrm{t}} = 1$ in (16), and $\sigma_{\mathrm{w}} = 1$ for the normalized variance of the synaptic weights, as defined by (26). We used the mean-field approximation for neural variances that is derived in in S2 Appendix. The analytic flow compares well with numerics, as shown in Fig. 3. For a subcritical rescaling factor a and $\sigma_{\mathrm{ext}} = 0$, the system flows towards a line of fixpoints defined by a vanishing σ_{y}^{2} and a finite $a \in [0, 1]$. When starting with $a > 0$, the fixpoint is instead $(a, \sigma_{\mathrm{y}}^{2}) = (1, 0)$. The situation changes qualitatively for finite external inputs, viz when $\sigma_{\mathrm{ext}} > 0$. The nullcline $\Delta\sigma_{\mathrm{y}}^{2} = 0$ is now continuous and the system flows, as shown in Fig. 3 to $a = 1$, with the value of σ_{y}^{2} being determined by the intersection of the two nullclines.

This analysis shows that external input is necessary for a robust flow towards the desired spectral weight, the reason being that the dynamics dies out before the spectral weight can be adapted when the isolated systems starts in the subcritical regime.

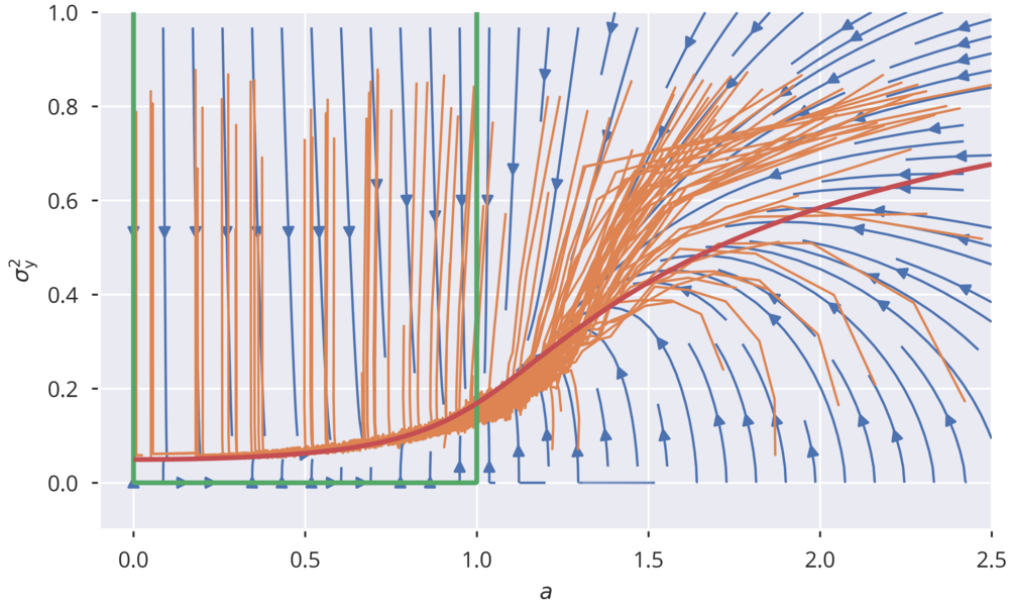


Fig 3. Spectral radius adaption dynamics. The dynamics of the synaptic rescaling factor a and the squared activity σ_{y}^2 (orange), as given by (4), for $R_{\mathrm{t}} = 1$. Also shown is the analytic approximation to the flow (blue), see (17) and (18), and the respective nullclines $\Delta a = 0$ (green) and $\Delta\sigma_{\mathrm{y}}^2 = 0$ (red). For the input, the heterogeneous binary protocol is used, with $\sigma_{\mathrm{ext}} = 0.25$ and $\epsilon_{\mathrm{a}} = 0.1$.

XOR-memory recall

To this point, we only presented results regarding the effectiveness of the introduced adaption rules. However, we did not account for their effects onto a given learning task. Therefore, we tested the performance of locally adapted networks under the delayed XOR task, which evaluates the memory capacity of the echo state network in combination with a non-linear operation. For the task, the XOR operation is to be taken with respect to a delayed pair of two consecutive binary inputs signals, $u(t-\tau)$ and $u(t-\tau-1)$, where τ is a fixed time delay. The readout layer is given by a single unit, which has the task to reproduce

$$f_\tau(t) = \mathrm{XOR}\left[u(t-\tau), u(t-\tau-1)\right], \qquad t, \tau = 1,\, 2,\, \ldots\,, \tag{19}$$

where $\mathrm{XOR}[u, u']$ is 0/1 if u and u' are identical/not identical.

The readout vector $\mathbf{w}_{\mathrm{out}}$ is trained with respect to the mean squared output error,

$$||\widehat{Y}\mathbf{w}_{\mathrm{out}} - \mathbf{f}_\tau||^2 + \alpha||\mathbf{w}_{\mathrm{out}}||^2\,, \tag{20}$$

using ridge regression on a sample batch of $T_{\mathrm{batch}} = 10N$ time steps, here for $N = 500$, and a regularization factor of $\alpha = 0.01$. The batch matrix $\widehat{Y}$, of size $T_{\mathrm{batch}} \times (N+1)$, holds the neural activities as well as one node with constant activity serving as a potential bias. The readout (column) vector $\mathbf{w}_{\mathrm{out}}$ is similarly of size $(N+1)$. The T_{batch} entries of $\mathbf{f}_\tau$ are the $f_\tau(t)$, viz the target values of the XOR problem. Minimizing (20) leads to

$$\mathbf{w}_{\mathrm{out}} = \left(\widehat{Y}^\dagger\widehat{Y} + \alpha^2\right)^{-1}\widehat{Y}^\dagger\,\mathbf{f}_\tau\,. \tag{21}$$

The learning procedure is repeated independently for each time delay τ. We quantified

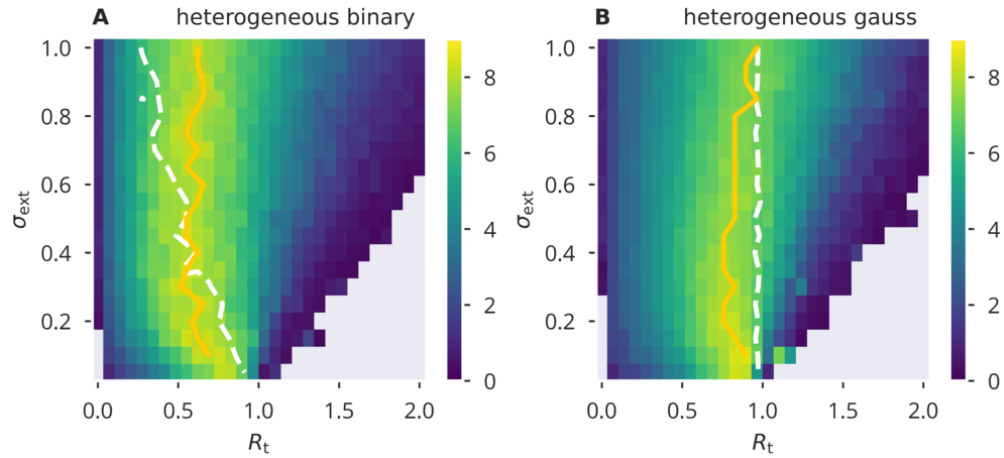


Fig 4. XOR performance for flow control. Color-coded performance sweeps for the XOR-performance (22) after adaptation using flow control. Averaged over five trials. The input has variance σ^2_{ext} and the target for the spectral radius R_{t}. A/B panels are for heterogeneous binary/Gaussian input protocols. Optimal performance (yellow solid line) is in general close to criticality, $R_{\text{a}} = 1$, as measured (white dashed lines).

the performance by the total memory capacity, MC_{XOR}, as

$$\text{MC}_{\text{XOR}} = \sum_{k=1}^{\infty} \text{MC}_{\text{XOR},k} \tag{22}$$

$$\text{MC}_{\text{XOR},k} = \frac{\text{Cov}^2\left[f_k(t), y_{\text{out}}(t)\right]_t}{\text{Var}\left[f_k(t)\right]_t \text{Var}\left[y_{\text{out}}(t)\right]_t}\,. \tag{23}$$

The activity $y_{\text{out}} = \sum_{i=1}^{N+1} w_{\text{out},i}\, y_i$ of the readout unit is compared in (23) with the XOR prediction task, with the additional neuron, $y_{N+1} = 1$, corresponding to the bias of the readout unit. The variance and covariance are calculated with respect to the batch size T_{batch}.

The results for flow control presented in Fig. 4 correspond to two input protocols, heterogeneous Gaussian and binary inputs. Shown are sweeps over a range of σ_{ext} and R_{t}. The update rule (3) was applied to the network for each pair of parameters until the a_i values converged to a stable configuration. We then measured the task performance as described above. Note that in the case of Gaussian input, this protocol was only used during the adaptation phases. Due to the nature of the XOR task, we chose to use binary inputs with the corresponding variances during the performance testing. See S2 Fig in the appendix for a performance sweep using the homogeneous binary and Gaussian input protocol. Optimal performance is generally attained around the $R_{\text{a}} \approx 1$ line. A spectral radius R_{a} slightly smaller than unity was optimal when using Gaussian input, but not for binary input signals. In this case the measured spectral radius R_{a} deviates linearly from the target R_{t}, with increasing strength of the input, as parameterized by the standard deviation σ_{ext}. Optimal performance is, however, essentially independent of the input strength, apart from $\sigma_{\text{ext}} \to 0$, with the peak located roughly at $R_{\text{t}} \approx 0.55$.

Comparing these results to variance control, as shown in Fig. 5, we found that variance control resulted in an overall lower performance. To our surprise, for external input with a large variance, Gaussian input caused stronger deviations from the desired spectral radius as compared to binary input. Therefore, in a sense, it appeared to behave opposite to what we found for flow control. However, similar to flow control, the

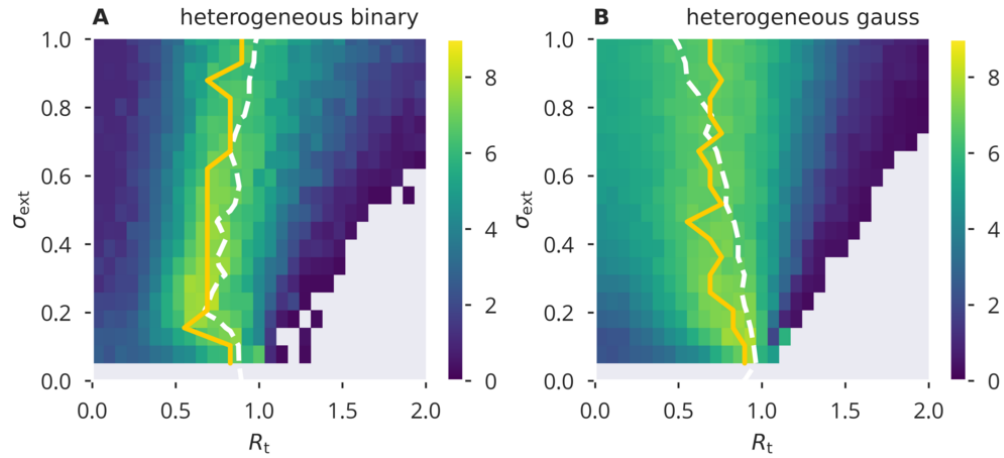


Fig 5. XOR performance for variance control. Color-coded performance sweeps for the XOR-performance (22) after adaptation using variance control. Averaged over five trials. The input has variance $\sigma^2_{\rm ext}$ and the target for the spectral radius $R_{\rm t}$. A/B panels are for heterogeneous binary/Gaussian input protocols. Optimal performance (yellow solid line) is in general close to criticality, $R_{\rm a} = 1$, as measured (white dashed lines).

value of $R_{\rm t}$ giving optimal performance under a given $\sigma_{\rm ext}$ remained relatively stable over the range of external input strength measured. On the other hand, using homogeneous input, see S3 Fig, did cause substantial deviations from the target spectral radius when using binary input.

The two input protocols used for the data shown in Fig. 4 and 5 relate to two different (biological) scenarios: First, an idealized case where a local recurrent neural ensemble only receives input from one particular source that has very specific statistics, such as, in our model case, a binary input sequence. In this scenario, it is plausible to assume that the input statistics under which long term adaptation is taking place coincide with the type of input patterns that are considered for the evaluation of the actual task performance. A second, presumably more plausible situation is the case where different input streams are integrated in the neural ensemble. Each neuron receives a weighted superposition of those input signals. Unless those inputs were strongly correlated, this would mean that random Gaussian input—as used in Fig. 4 B and 5 B—is a reasonable approximation to this scenario.

Input induced correlations

The theory for the echo-state layer, Eq. (49), which was used in Eq. (18) and (6), assumes that the variance $\sigma^2_{\rm bare}$ of the bare recurrent contribution to the membrane potential, $x_{\rm bare} = \sum_j W_{ij} y_j$, is given by $\sigma^2_{\rm w}\sigma^2_{\rm y}$. In Fig. 6, a comparison is presented for the four input protocols introduced in Sect. *Input protocols*. For the Gaussian protocols, for which neurons receive statistically uncorrelated external signals, one observes that $\sigma^2_{\rm bare} \to \sigma^2_{\rm w}\sigma^2_{\rm y}$ in the thermodynamic limit $N \to \infty$ via a power law, which is to be expected when the presynaptic neural activities are decorrelated. On the other side, binary 0/1 inputs act synchronous on all sites, either with site-dependent or site-independent strengths (heterogeneous/homogeneous). Corresponding activity correlations are induced and a finite and only weakly size-dependent difference between $\sigma^2_{\rm bare}$ and $\sigma^2_{\rm w}\sigma^2_{\rm y}$ shows up. Substantial corrections to the analytic theory are to be expected in this case. To this extend we measured the cross-correlation $C(y_i, y_j)$,

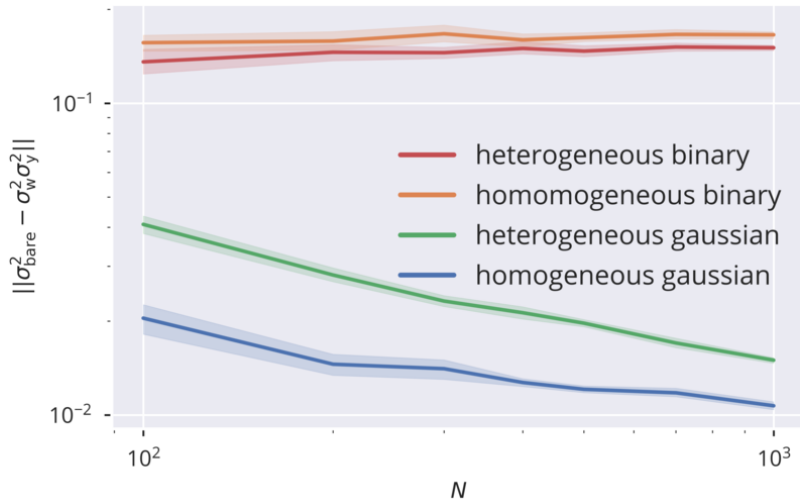


Fig 6. Size dependence of correlation. Comparison between the variance σ^2_{bare} of the bare recurrent input $x_{\text{bare}} = \sum_j W_{ij} y_j$ with $\sigma^2_{\text{w}}\sigma^2_{\text{y}}$. Equality is given when the presynaptic activities are statistically independent. This can be observed in the limit of large network sizes N for uncorrelated input data streams (homogeneous and heterogeneous Gaussian input protocols), but not for correlated inputs (homogeneous and heterogeneous binary input protocols). Compare Sect. *Input protocols* for the input protocols. Parameters are $\sigma_{\text{ext}} = 0.5$, $R_{\text{a}} = 1$ and $\mu_t = 0.05$.

defined as

$$\bar{C} = \frac{1}{N(N-1)} \sum_{i \neq j} |C(y_i, y_j)|, \qquad C(y_i, y_j) = \frac{\text{Cov}(y_i, y_j)}{\sqrt{\text{Cov}(y_i, y_i)\text{Cov}(y_j, y_j)}}, \tag{24}$$

with the covariance given by $\text{Cov}(y_i, y_j) = \langle (y_i - \langle y_i \rangle_t)(y_j - \langle y_j \rangle_t) \rangle_t$. For a system of $N = 500$ neurons the results for the averaged absolute correlation $\bar{C}$ are presented in Fig. 7 (see S4 Fig in the supplementary material for homogeneous input protocols). Autonomous echo-state layers are in chaotic states when supporting a finite activity level, which implies that correlations vanish in the thermodynamic limit $N \to \infty$. The case $\sigma_{\text{ext}} = 0$, as included in Fig. 7, serves consequently as a yardstick for the magnitude of correlations that are due to the finite number of neurons.

Input correlations are substantially above the autonomous case for correlated binary inputs, with the magnitude of $\bar{C}$ decreasing when the relative contribution of the recurrent activity increases. This is the case for increasing R_{t}. The effect is opposite for the Gaussian protocol, for which the input does not induce correlations, but contributes to decorrelating neural activity. The mean absolute correlation $\bar{C}$ is in this case suppressed when the internal activity becomes small in the limit $R_{\text{t}} \to 0$. Finite-size effects are relevant only when the internal activity is of the order of the input-induced activity, viz for increasing R_{t}.

Discussion

The mechanisms for tuning the spectral radius via a local homeostatic adaption rule introduced in the present study require neurons to have the ability to distinguish and measure locally both external and recurrent input contributions. For flow control, neurons need to be able to compare the recurrent membrane potential with their own activity, as assumed in Sect. *Autonomous spectral radius regulation*. On the other hand,

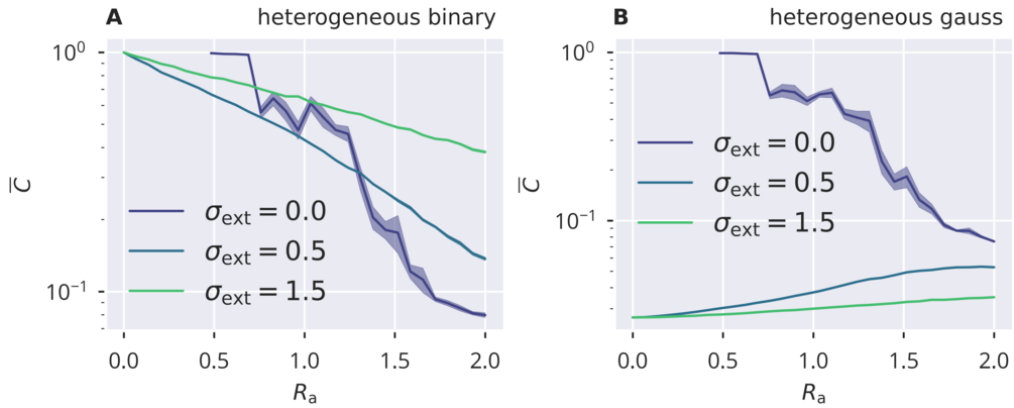


Fig 7. Input induced activity correlations. For heterogeneous binary and Gaussian inputs (A/B), the dependency of mean activity cross correlations $\bar{C}$, see Eq. (24). $\bar{C}$ is shown as a function of the target spectral radius R_a. Results are obtained for $N = 500$ sites by averaging over five trials, with shadows indicating the accuracy. Correlations are due to finite-size effect for the autonomous case $\sigma_{\text{ext}} = 0$.

variance control directly measures the variance of the external input and derives the activity target variance accordingly. The limiting factor to a successful spectral radius control is the amount of cross-correlation induced by external driving statistics. As such, the functionality and validity of the proposed mechanisms depended on the ratio between external input, i.e. feed-forward or feedback connections, with respect to recurrent, or lateral connections. In general, it is not straightforward to directly connect experimental evidence regarding the ratio between recurrent and feed-forward contributions to the effects observed in the model. It is, however, worthwhile to note that the fraction of synapses associated with interlaminar loops and intralaminar lateral connections are estimated to make up roughly 50% [40]. Relating this to our model, it implies that the significant interneural correlations that we observed when external input strengths were of the same order of magnitude as the recurrent inputs, can not generally be considered an artifact of biologically implausible parameter choices. In fact, synchronization is a widely observed phenomenon in the brain and is even considered important for information processing [41, 42].

Overall, we found flow control to be generally more robust than variance control in the sense that, while still being affected by the amount of correlations within the neural reservoir, the task performance was less so prone to changes in the external input strength. Comparatively stable network performance could be observed, in spite of significant deviations from the desired spectral radius (see Fig. 4). A possible explanation may be that flow control uses a distribution of samples from only a restricted part of phase space, that is, from the phase space regions that are actually visited or "used" for a given input. Therefore, while a spectral radius of unity ensures –statistically speaking– the desired scaling properties in all phase-space directions, it seem to be enough to control the correct scaling for the subspace of activities that is actually used for a given set of input patters. Variance control, on the other hand, relies more strictly on the assumptions regarding the statistical independence of neural activities. In consequence, the desired results could only be achieved under a rather narrow set of input statistics (independent Gaussian input with small variance). In addition, the approximate expression derived for the nonlinear transformation appearing in the mean field approximation adds another layer of potential source of systematic error to the control mechanism. This aspect also speaks in favor of flow control, since its rules are mathematically more simple. In contrast to variance control, the

stationarity condition stated in Eq. (13) is independent of the actual nonlinear activation function used and could easily be adopted in a modified neuron model. It should be noted, however, that the actual target R_t giving optimal performance might then also be affected.

Interestingly, flow control distinguishes itself from a conventional local activity-target perspective of synaptic homeostasis: There is no predefined set point in Eq. (3). This allows heterogeneities of variances of neural activity to develop across the network, while retaining the average neural activity at a fixed predefined level.

Conclusion

Apart from being relevant from a theoretical perspetive, we propose that the separability of recurrent and external contributions to the membrane potential is an aspect that is potentially relevant for the understanding of local homeostasis in biological networks. While homeostasis in neural compartments has been a subject of experimental research [43, 44], to our knowledge, it has not yet been further investigated on a theoretical basis. Given that the neural network model used in this study lacks some features characterizing biological neural networks (e.g. strict positivity of the neural firing rate, Dale's law), future research should therefore investigate whether the here presented framework of local flow control can be implemented within more realistic biological neural network models. A particular concern regarding our findings is that biological neurons are spiking. The concept of an underlying instantaneous firing rate is, strictly speaking, a theoretical construct, let alone the definition of higher moments, such as the "variance of neural activity". However, it is important to note that real-world biological control mechanims, e.g. of the activity, rely on physical quantities that serve as measurable correlates. A well-known example is the intracellular calcium concentration, which is essentially a linearly filtered version of the neural spike train [32]. On a theoretical level, Cannon and Miller showed that dual homeostasis can successfully control the mean and variance of this type of spike-averaging physical quantities [26]. An extension of the flow control to filtered spike trains of spiking neurons could be an interesting subject of further investigations. However, using spiking neuron models would have shifted the focus of our research towards the theory of liquid state machines [45, 46], exceeding the scope of this publication. We therefore leave the extension to more realistic network/neuron models to future work.

Materials and methods

Model

We implemented an echo state network with N neurons, receiving D_in inputs. The neural activity is $y_i \in [-1, 1]$, x_i the membrane potential, u_i the input activities, W_{ij} the internal synaptic weights and I_i the external input received. The output layer will be specified later. The dynamics

$$x_i(t) = a_i \sum_{j=1}^{N} W_{ij} y_j(t-1) + I_i(t), \qquad y_i(t) = \tanh\left(x_i(t) - b_i\right) \tag{25}$$

is discrete in time, where the input I_i is treated instantaneously. A tanh-sigmoidal has been used as a nonlinear activation function.

The synaptic renormalization factor a_i in (25) can be thought of as a synaptic scaling parameter that neurons use to regulate the overall strength of the recurrent

Table 1. Standard values for model parameters

N	p_r	σ_w	μ_t	ϵ_b	ϵ_a	ϵ_μ	ϵ_σ
500	0.1	1	0.05	10^{-3}	10^{-3}	10^{-4}	10^{-3}

inputs. The strength of the inputs I_i is unaffected, which is biologically plausible if external and recurrent signals arrive at separate branches of the dendritic tree [44].

The W_{ij} are the bare synaptic weights, with $a_i W_{ij}$ being the components of the effective weight matrix $\widehat{W}_\mathrm{a}$. Key to our approach is that the propagation of activity is determined by $\widehat{W}_\mathrm{a}$, which implies that the spectral radius of the effective, and not of the bare weight matrix needs to be regulated.

The bare synaptic matrix W_{ij} is sparse, with a connection probability $p_\mathrm{r} = 0.1$. The non-zero elements are drawn from a Gaussian with standard deviation

$$\sigma = \frac{\sigma_\mathrm{w}}{\sqrt{N p_\mathrm{r}}}\,, \tag{26}$$

and vanishing mean μ. Here $N p_\mathrm{r}$ corresponds to the mean number of afferent internal synapses, with the scaling $\sim 1/\sqrt{N p_\mathrm{r}}$ enforcing size-consistent synaptic-weight variances.

As introduced in the introduction, we applied the following adaptation mechanisms:

$$b_i(t) = b_i(t-1) + \epsilon_\mathrm{b}\left[y_i(t) - \mu_\mathrm{t}\right]\ . \tag{27}$$

For the gains, using flow control:

$$a_i(t) = a_i(t-1)\Big[1 + \epsilon_\mathrm{a}\Delta R_i(t)\Big], \qquad \Delta R_i(t) = R_\mathrm{t}^2 |y_i(t-1)|^2 - |x_{\mathrm{r},i}(t)|^2\ . \tag{28}$$

For variance control:

$$a_i(t) = a_i(t-1) + \epsilon_\mathrm{a}\left[\sigma_{\mathrm{t},i}^2(t) - \left(y_i(t) - \mu_i^\mathrm{y}(t)\right)^2\right] \tag{29}$$

$$\sigma_{\mathrm{t},i}^2(t) = 1 - \sqrt{1 + 2R_\mathrm{t}^2 y_i(t)^2 + 2\sigma_{\mathrm{ext},i}^2(t)} \tag{30}$$

$$\mu_i^\mathrm{y}(t) = \mu_i^\mathrm{y}(t-1) + \epsilon_\mu\left[y_i(t) - \mu_i^\mathrm{y}(t-1)\right] \tag{31}$$

$$\sigma_{\mathrm{ext},i}^2(t) = \sigma_{\mathrm{ext},i}^2(t-1) + \epsilon_\sigma\left[\left(I_i(t) - \mu_{\mathrm{ext},i}(t)\right)^2 - \sigma_{\mathrm{ext},i}^2(t-1)\right] \tag{32}$$

$$\mu_{\mathrm{ext},i}(t) = \mu_{\mathrm{ext},i}(t-1) + \epsilon_\mu\left[I_i(t) - \mu_{\mathrm{ext},i}(t-1)\right]\ . \tag{33}$$

For a summary of model parameters, see Table 1.

Input protocols

Overall, we examined four distinct input protocols.

- *Homogeneous Gaussian.* Nodes receive inputs $I_i(t)$ that are drawn individually from a Gaussian with vanishing mean and standard deviation σ_ext.
- *Heterogeneous Gaussian.* Nodes receive stochastically independent inputs $I_i(t)$ that are drawn from Gaussian distributions with vanishing mean and node specific standard deviations $\sigma_{i,\mathrm{ext}}$. The individual $\sigma_{i,\mathrm{ext}}$ are normal distributed, as drawn from the positive part of a Gaussian with mean zero and variance σ_ext^2.
- *Homogeneous binary.* Sites receive identical inputs $I_i(t) = \sigma_\mathrm{ext} u(t)$, where $u(t) = \pm 1$ is a binary input sequence.

- *Heterogeneous binary.* We define with

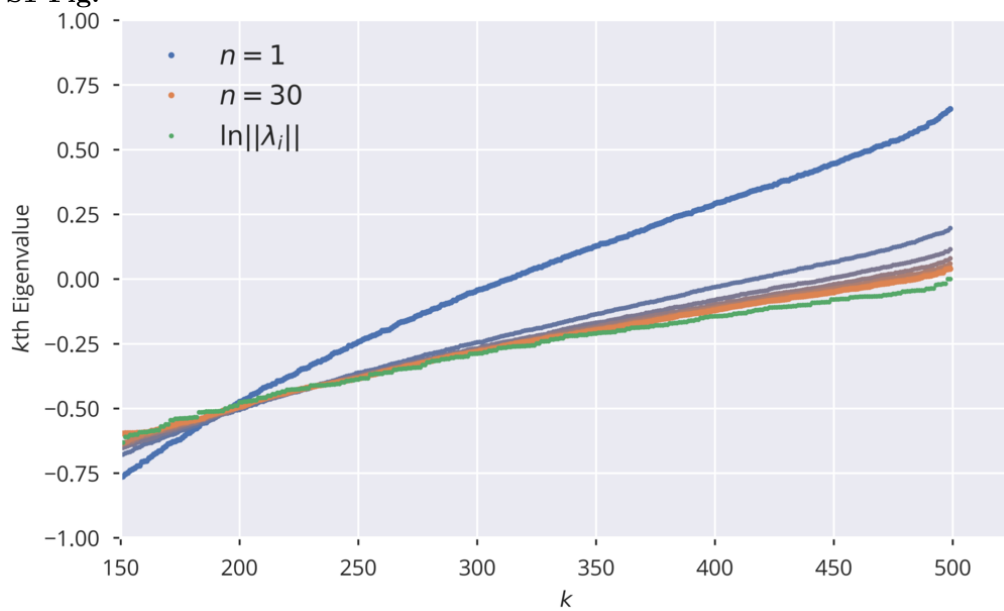
$$I_i = W_i^{\mathrm{u}} u(t), \qquad u_j(t) = \pm 1 \tag{34}$$

  the afferent synaptic weight vector W_i^{u}, which connects the binary input sequence $u(t)$ to the network. All W_i^{u} are drawn independently from a Gaussian with mean zero and standard deviation σ_{ext}.

The Gaussian input variant simulates external noise. We used it in particular to test predictions of the theory developed in Sect. *S2 Appendix*. In order to test the performance of the echo state network with respect to the delayed XOR task, the binary input protocols are employed. A generalization of the here defined protocols to the case of vectorial input signals would be straightforward.

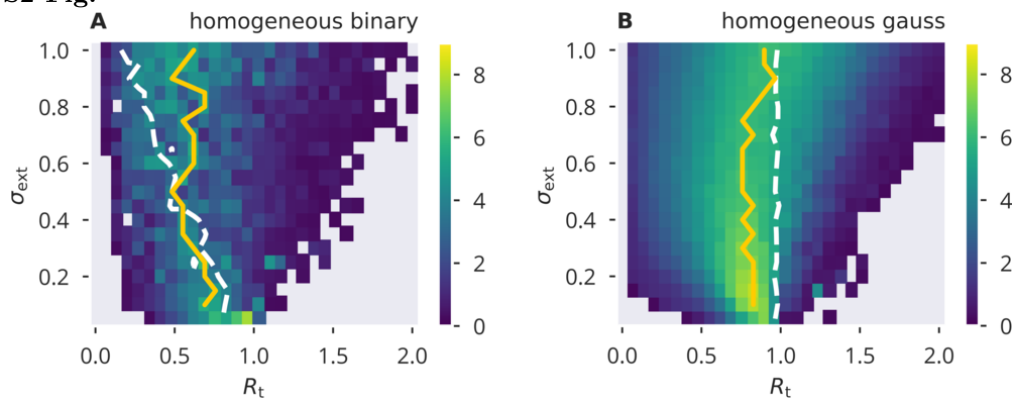
Supporting information

S1 Fig.



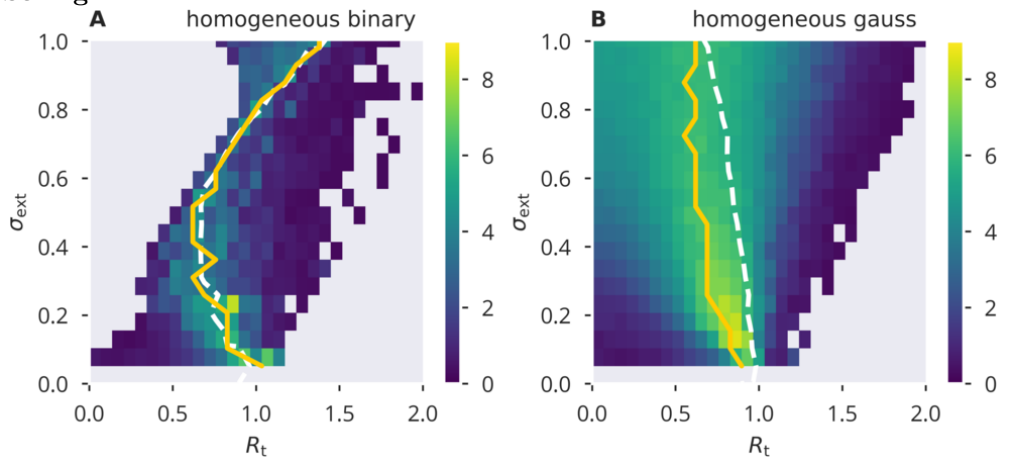
Convergence of Lyapunov Spectrum. Convergence of eigenvalues of $\ln\left((\widehat{W}^n)^\dagger \widehat{W}^n\right)/(2n)$ for different n, as discussed in Sect. . $\widehat{W}$ is a random Gaussian matrix which was rescaled to a spectral radius of one. Colors from blue to orange encode the exponent n ranging between 1 and 30. Green dots show the theoretical limit of $\ln\|\lambda_i\|$, where λ_i is the ith eigenvalue of $\widehat{W}$.

S2 Fig.



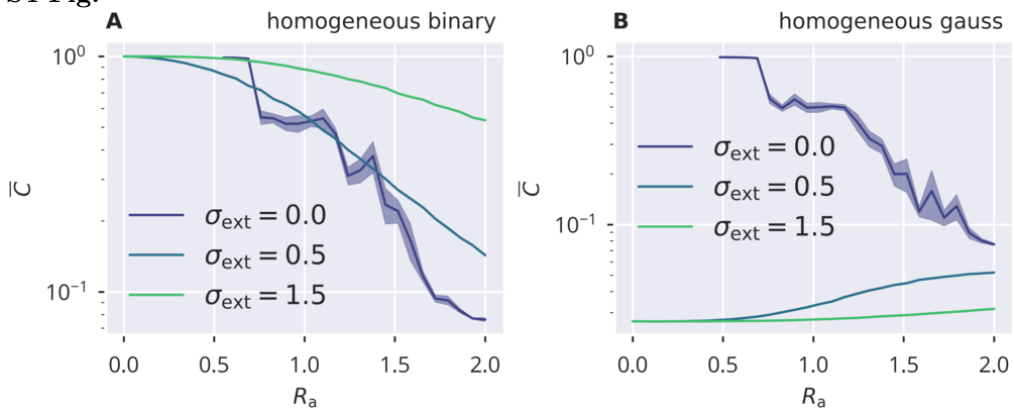
XOR performance for flow control, homogeneous input. Numerical results for the network performance under a time-delayed XOR task, as defined in Sect. *XOR-memory recall*, using homogeneous binary/Gaussian input. Shown are color-coded performance sweeps for the XOR-performance (22), averaged over five trials. The input has variance σ^2_{ext} and the target for the spectral radius R_{t}. A/B panels are for binary/Gaussian input protocols. Optimal performance for a given σ_{ext} is given by yellow solid lines, measured value of $R_{\text{a}} = 1$ by white dashed lines.

S3 Fig.



XOR performance for variance control, homogeneous input. Numerical results for the network performance under a time-delayed XOR task, as defined in Sect. *XOR-memory recall*, using homogeneous binary/Gaussian input. Shown are color-coded performance sweeps for the XOR-performance (22), averaged over five trials. The input has variance σ^2_{ext} and the target for the spectral radius R_{t}. A/B panels are for binary/Gaussian input protocols. Optimal performance for a given σ_{ext} is given by yellow solid lines, measured value of $R_{\text{a}} = 1$ by white dashed lines.

S4 Fig.



Input induced activity correlations. For homogeneous binary and Gaussian inputs (A/B), the dependency of mean activity cross correlations $\bar{C}$, see Eq. (24). $\bar{C}$ is shown as a function of the target spectral radius R_{a}. Results are obtained for $N = 500$ sites by averaging over five trials, with shadows indicating the accuracy. Correlations are due to finite-size effect for the autonomous case $\sigma_{\text{ext}} = 0$.

S1 Appendix.

Extended theory of flow control for independent neural activity

We would like to show that the stationarity condition in Eq. (13) results in the correct spectral radius, under the special case of independently identically distributed neural activities with zero mean.

We start with Eq. (13) as a stationarity condition for a given R_{t}:

$$\left\langle ||\mathbf{x}_{\text{r}}(t)||^2 \right\rangle_{\text{t}} \stackrel{!}{=} R_{\text{t}}^2 \left\langle ||\mathbf{y}(t-1)||^2 \right\rangle_{\text{t}} . \tag{35}$$

We can express the left side of the equation as

$$\text{E}\left[\mathbf{y}^\dagger(t)\widehat{W}_{\text{a}}^\dagger \widehat{W}_{\text{a}}\mathbf{y}(t)\right]_t . \tag{36}$$

We define $\widehat{U}_{\text{a}} \equiv= \widehat{W}_{\text{a}}^\dagger \widehat{W}_{\text{a}}$ with $\{\sigma_k^2\}$ being the set of eigenvalues, which are also the squared singular values of $\widehat{W}_{\text{a}}$, and $\{\mathbf{u}_k\}$ the respective set of orthonormal (column) eigenvectors. We insert the identity $\sum_{k=1}^N \mathbf{u}_k \mathbf{u}_k^\dagger$ and find

$$\text{E}\left[\mathbf{y}^\dagger(t)\widehat{U}_{\text{a}} \sum_{k=1}^N \mathbf{u}_k \mathbf{u}_k^\dagger \mathbf{y}(t)\right]_t \tag{37}$$

$$= \text{E}\left[\sum_{k=1}^N \sigma_k^2 \mathbf{y}^\dagger(t)\mathbf{u}_k \mathbf{u}_k^\dagger \mathbf{y}(t)\right]_t \tag{38}$$

$$= \sum_{k=1}^N \sigma_k^2 \mathbf{u}_k^\dagger \text{E}\left[\mathbf{y}(t)\mathbf{y}^\dagger(t)\right]_t \mathbf{u}_k \tag{39}$$

$$= \sum_{k=1}^N \sigma_k^2 \mathbf{u}_k^\dagger \widehat{C}_{\text{yy}} \mathbf{u}_k \tag{40}$$

$$= \text{Tr}\left(\widehat{D}_{\sigma^2} \widehat{S}_{\text{u}}^\dagger \widehat{C}_{\text{yy}} \widehat{S}_{\text{u}}\right) . \tag{41}$$

Given zero mean neural activity, $\widehat{C}_{\mathrm{yy}} = \mathrm{E}[\mathbf{y}(t)\mathbf{y}^{\dagger}(t)]_t$ is the covariance matrix of neural activities. $\widehat{D}_{\sigma^2}$ is a diagonal matrix holding the $\{\sigma_k^2\}$ and $\widehat{S}_{\mathrm{u}}$ is a unitary matrix whose columns are $\{\mathbf{u}_k\}$. $\widehat{S}_{\mathrm{u}}^{\dagger}\widehat{C}_{\mathrm{yy}}\widehat{S}_{\mathrm{u}}$ is expressing $\widehat{C}_{\mathrm{yy}}$ in the diagonal basis of $\widehat{U}_{\mathrm{a}}$.

Including the right hand side of the equation, we get

$$\mathrm{Tr}\left(\widehat{D}_{\sigma^2}\widehat{S}_{\mathrm{u}}^{\dagger}\widehat{C}_{\mathrm{yy}}\widehat{S}_{\mathrm{u}}\right) = R_{\mathrm{t}}^2\mathrm{Tr}\left(\widehat{C}_{\mathrm{yy}}\right) . \tag{42}$$

However, since the trace is invariant under a change of basis, we find

$$\mathrm{Tr}\left(\widehat{D}_{\sigma^2}\widehat{S}_{\mathrm{u}}^{\dagger}\widehat{C}_{\mathrm{yy}}\widehat{S}_{\mathrm{u}}\right) = R_{\mathrm{t}}^2\mathrm{Tr}\left(\widehat{S}_{\mathrm{u}}^{\dagger}\widehat{C}_{\mathrm{yy}}\widehat{S}_{\mathrm{u}}\right) . \tag{43}$$

Defining $\widehat{C}^{\mathrm{u}} \equiv= \widehat{S}_{\mathrm{u}}^{\dagger}\widehat{C}_{\mathrm{yy}}\widehat{S}_{\mathrm{u}}$, we get

$$\sum_{k=1}^{N}\sigma_k^2 C_{kk}^{\mathrm{u}} = R_{\mathrm{t}}^2\sum_{k=1}^{N}C_{kk}^{\mathrm{u}} . \tag{44}$$

If we assume that the node activities are independently identically distributed with zero mean, we get $(\widehat{C}_{\mathrm{yy}})_{ij} = (\widehat{C}^{\mathrm{u}})_{ij} = \left\langle y^2\right\rangle_t \delta_{ij}$. In this case, which was also laid out in Sect. , the equation reduces to

$$\sum_{k=1}^{N}\sigma_k^2 = R_{\mathrm{t}}^2 N . \tag{45}$$

The Frobenius norm of a square Matrix $\widehat{A}$ is given by $\left\|\widehat{A}\right\|_{\mathrm{F}}^2 \equiv= \sum_{i,j}\widehat{A}_{ij}^2$. Furthermore, the Frobenius norm is linked to the singular values via $\left\|\widehat{A}\right\|_{\mathrm{F}}^2 = \sum_k \sigma_k^2(\widehat{A})$ [36, 37]. This allows us to state

$$\sum_{i,j}\left(\widehat{W}_{\mathrm{a}}\right)_{ij}^2 = R_{\mathrm{t}}^2 N \tag{46}$$

which, by using (11), gives

$$R_{\mathrm{a}}^2 = R_{\mathrm{t}}^2 . \tag{47}$$

A slightly less restrictive case is that of uncorrelated but inhomogeneous activity, that is $(\widehat{C}_{\mathrm{yy}})_{ij} = \left\langle y_i^2\right\rangle_t \delta_{ij}$. The statistical properties of the diagonal elements C_{kk}^{u} then determine to which degree one can still claim that Eq. (44) leads to Eq. (45). Fig. 8 shows an example of a randomly generated realization of $(\widehat{C}_{\mathrm{yy}})_{ij} = \left\langle y_i^2\right\rangle_t$ and the resulting diagonal elements of $\widehat{C}^{\mathrm{u}}$, where the corresponding orthonormal basis $\widehat{S}_{\mathrm{u}}$ was generated from the SVD of a random Gaussian matrix.

As one can see, the basis transformation has a strong smoothing effect on the diagonal entries, while the mean over the diagonal elements is preserved. Note that this effect was not disturbed by introducing random row-wise multiplications to the random matrix from which the orthonormal basis was derived. The smoothing of the diagonal entries allows us to state that $C_{kk}^{\mathrm{u}} \approxeq \left\langle y^2\right\rangle$ is a very good approximation in the case considered, which therefore reduces (44) to the homogeneous case previously described. We can conclude that the adaptation mechanism also gives the desired spectral radius under uncorrelated inhomogeneous activity.

In the most general case, we can still state that if C_{kk}^{u} and σ_k^2 are uncorrelated, for large N, Eq. (44) will tend towards

$$N\left\langle\sigma^2\right\rangle\left\langle C^{\mathrm{u}}\right\rangle = NR_{\mathrm{t}}^2\left\langle C^{\mathrm{u}}\right\rangle \tag{48}$$

which would also lead to Eq. (45). However, we can not generally guarantee statistical independence since the recurrent contribution on neural activities and the resulting entries of $\widehat{C}_{\mathrm{yy}}$ and thus also C_{kk}^{u} are linked to $\widehat{S}$ and σ_k^2, being the SVD of the recurrent weight matrix.

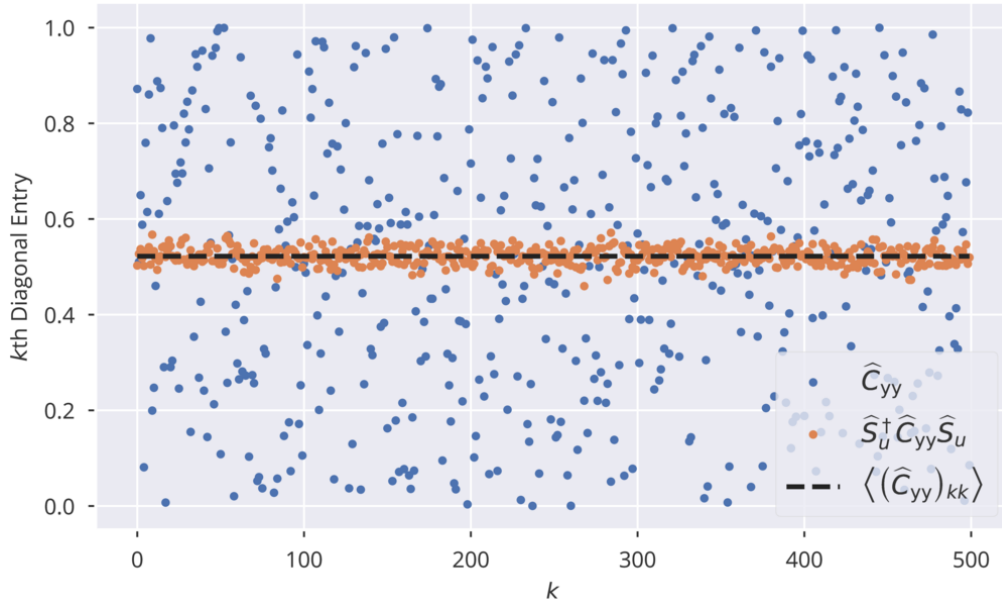


Fig 8. Diagonal Elements of a randomly generated covariance matrix and its representation in the $\widehat{\mathbf{S}}_{\mathrm{u}}$ basis. $\widehat{C}_{\mathrm{yy}}$ is a diagonal matrix with diagonal entries randomly drawn from $[0, 1]$, $\widehat{S}_{\mathrm{u}}$ is the orthonormal eigenbasis of $\widehat{W}_{\mathrm{a}}^{\dagger}\widehat{W}_{\mathrm{a}}$, where $\widehat{W}_{\mathrm{a}}$ is a random Gaussian matrix. The black dashed line denotes the average over the diagonal entries of $\widehat{C}_{\mathrm{yy}}$.

S2 Appendix

Mean field theory for echo state layers

In the following, we deduce analytic expressions allowing to examine the state of echo-state layers subject to a continuous timeline of inputs. Our approach is similar to the one presented by Massar [47].

The recurrent part of the input x_i received by a neuron is a superposition of Np_{r} terms, which are assumed here to be uncorrelated. Given this assumption, the self-consistency equations

$$\sigma_{\mathrm{y},i}^2 = \int_{-\infty}^{\infty} \mathrm{dx} \tanh^2(x) N_{\mu_i,\sigma_i}(x) - \mu_{\mathrm{y},i}^2 \tag{49}$$

$$\mu_{\mathrm{y},i} = \int_{-\infty}^{\infty} \mathrm{dx} \tanh(x) N_{\mu_i,\sigma_i}(x) \tag{50}$$

$$\sigma_i^2 = a_i^2 \sigma_{\mathrm{w}}^2 \left\langle \sigma_{\mathrm{y},j}^2 \right\rangle_j + \sigma_{\mathrm{ext},i}^2, \qquad \mu_i = \mu_{\mathrm{ext},i} - b_i \tag{51}$$

determine the properties of the stationary state. We recall that σ_{w} parameterizes the distribution of bare synaptic weights via (26). The general expressions (49) and (50) hold for all neurons, with the site-dependency entering exclusively via a_i, b_i, $\sigma_{\mathrm{ext},i}$ and $\mu_{\mathrm{ext},i}$, as in (51), with the latter characterizing the standard deviation and the mean of the input. Here, $a_i^2 \sigma_{\mathrm{w}}^2 \sigma_{\mathrm{y}}^2$ is the variance of the recurrent contribution to the membrane potential, x, and σ^2 the respective total variance. The membrane potential is Gaussian distributed, as $N_{\mu,\sigma}(x)$, with mean μ and variance σ^2, which are both to be determined self-consistently. Variances are additive for stochastically independent processes, which has been assumed in (51) to be the case for recurrent activities and the external inputs. The average value for the mean neural activity is μ_i.

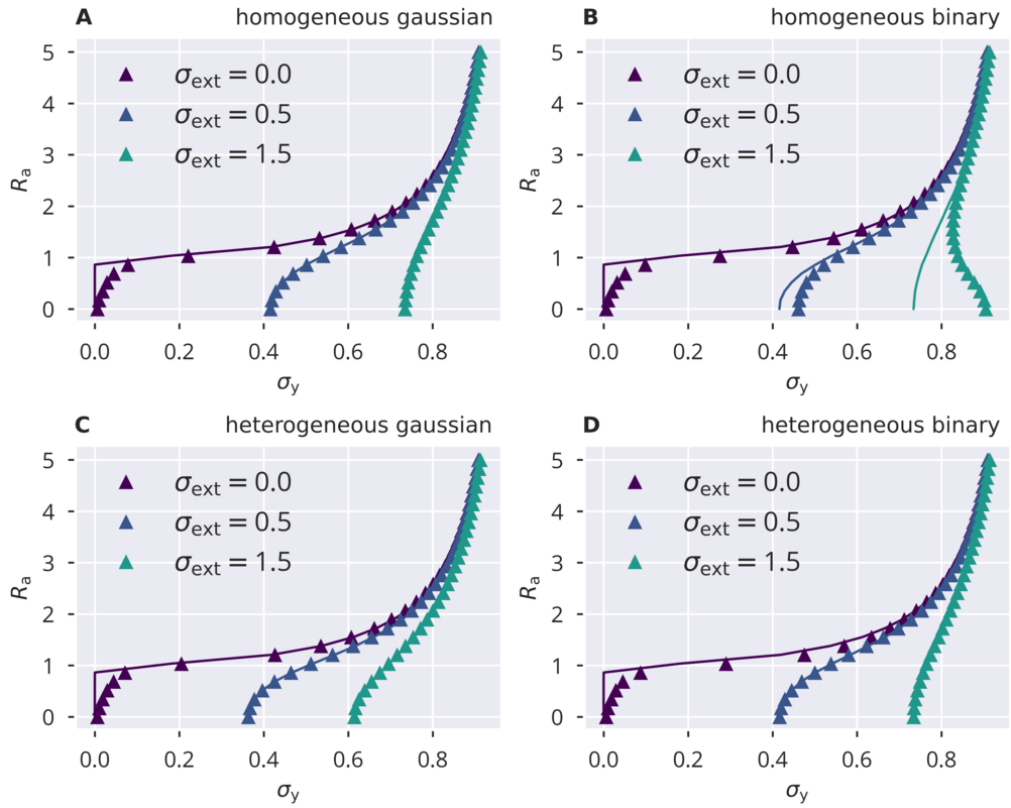


Fig 9. Variance control for the spectral radius. The spectral radius R_{a}, given by the approximation $R_{\text{a}}^2 = \sum_i a_i^2/N$, for the four input protocols defined in Sect. *Input protocols*. Lines show the numerical self-consistency solution of (49), symbols the full network simulations. Note the instability for small σ_{y} and σ_{ext}. **A**: Homogeneous independent Gaussian input. **B**: Homogeneous identical binary input. **C**: Heterogeneous independent Gaussian input. **D**: Heterogeneous identical binary input.

For a given set of a_i, $\sigma_{\text{ext},i}$ and b_i, the means and variances of neural activities, $\sigma_{\text{y},i}^2$ and $\mu_{\text{y},i}$, follow implicitly.

We compared the numerically determined solutions of (49) and (50) against full network simulations using, as throughout this study, $N = 500$, $p_{\text{r}} = 0.1$, $\sigma_{\text{w}} = 1$, $\mu_{\text{t}} = 0.05$. In Fig. 9, the spectral radius R_{a} is given for the four input protocols defined in Sect. *Input protocols*. The identical ensemble of input standard deviations $\sigma_{\text{ext},i}$ enters both theory and simulations.

Theory and simulations are in good accordance for vanishing input. Here, the reason is that finite activity levels are sustained in an autonomous random neural network when the ongoing dynamics is chaotic and hence decorrelated. For reduced activity levels, viz for small variances σ_y^2, the convergence of the network dynamics is comparatively slow, which leads to a certain discrepancy with the analytic prediction (see Fig. 9).

Gaussian approximation The integral occurring in the self-consistency condition (49) can be evaluated explicitly when a tractable approximation to the squared transfer function $\tanh^2()$ is available. A polynomial approximation would capture the leading behavior close to the origin, however without accounting for the fact that $\tanh^2()$ converges to unity for large absolute values of the membrane potential. Alternatively, an approximation incorporating both conditions, the correct second-order scaling for small,

and the correct convergence for large arguments, is given by the Gaussian approximation

$$\tanh^2(x) \approx 1 - \exp\left(-x^2\right) \ . \tag{52}$$

With this approximation the integral in (49) can be evaluated explicitly. The result is

$$\begin{aligned} \frac{1}{1-\sigma_{\mathrm{y}}^2-\mu_{\mathrm{y}}^2} &= \sqrt{1+2\sigma^2}/\exp\left(-\mu^2/\left(1+2\sigma^2\right)\right) \\ &= \sqrt{1+2a^2\sigma_{\mathrm{w}}^2\sigma_{\mathrm{y}}^2+2\sigma_{\mathrm{ext}}^2}/\exp\left(-\mu^2/\left(1+2a^2\sigma_{\mathrm{w}}^2\sigma_{\mathrm{y}}^2+2\sigma_{\mathrm{ext}}^2\right)\right) \ . \end{aligned} \tag{53}$$

Assuming that $\mu \approx 0$ and $\mu_{\mathrm{y}} \approx 0$, inverting the first equation yields a relatively simple analytic approximation for the variance self-consistency equation:

$$\sigma_{\mathrm{y}}^2 = 1 - \frac{1}{\sqrt{1+2a^2\sigma_{\mathrm{w}}^2\sigma_{\mathrm{y}}^2+2\sigma_{\mathrm{ext}}^2}} \ . \tag{54}$$

This equation was then used for the approximate update rule in (6) and (18).

Acknowledgments

References

1. Lukoševičius M, Jaeger H. Reservoir computing approaches to recurrent neural network training. Computer Science Review. 2009;3(3):127–149. doi:https://doi.org/10.1016/j.cosrev.2009.03.005.

2. Ozturk MC, Xu D, Pr\'\incipe JC. Analysis and design of echo state networks. Neural computation. 2007;19(1):111–138.

3. Jaeger H. The "echo state" approach to analysing and training recurrent neural networks. GMD - German National Research Institute for Computer Science; 2001. 148. Available from: `http://www.faculty.jacobs-university.de/hjaeger/pubs/EchoStatesTechRep.pdf`.

4. Caluwaerts K, Wyffels F, Dieleman S, Schrauwen B. The spectral radius remains a valid indicator of the echo state property for large reservoirs. In: Neural Networks (IJCNN), The 2013 International Joint Conference on. IEEE; 2013. p. 1–6.

5. Zhang B, Miller DJ, Wang Y. Nonlinear system modeling with random matrices: Echo state networks revisited. IEEE Transactions on Neural Networks and Learning Systems. 2012;23(1):175–182. doi:10.1109/TNNLS.2011.2178562.

6. Jaeger H. A tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the "echo state network" approach. German National Research Center for Information Technology; 2002. 48.

7. Jaeger H. Short Term Memory in Echo State Networks. Fraunhofer Institute for Autonomous Intelligent Systems; 2002. 152.

8. Farkaš I, Bosák R, Gergel' P. Computational analysis of memory capacity in echo state networks. Neural Networks. 2016;83:109–120.

9. Boedecker J, Obst O, Lizier JT, Mayer NM, Asada M. Information processing in echo state networks at the edge of chaos. Theory in Biosciences. 2012;131(3):205–213. doi:10.1007/s12064-011-0146-8.

10. Nikolić D, Häusler S, Singer W, Maass W. Distributed fading memory for stimulus properties in the primary visual cortex. PLoS biology. 2009;7(12):e1000260.

11. Hinaut X, Lance F, Droin C, Petit M, Pointeau G, Dominey PF. Corticostriatal response selection in sentence production: Insights from neural network simulation with reservoir computing. Brain and Language. 2015;150:54–68. doi:10.1016/j.bandl.2015.08.002.

12. Enel P, Procyk E, Quilodran R, Dominey PF. Reservoir Computing Properties of Neural Dynamics in Prefrontal Cortex. PLOS Computational Biology. 2016;12(6):e1004967. doi:10.1371/journal.pcbi.1004967.

13. Gallicchio C, Micheli A. Echo state property of deep reservoir computing networks. Cognitive Computation. 2017;9(3):337–350.

14. Livi L, Bianchi FM, Alippi C. Determination of the edge of criticality in echo state networks through Fisher information maximization. IEEE Transactions on Neural Networks and Learning Systems. 2018;29(3):706–717.

15. Schrauwen B, Wardermann M, Verstraeten D, Steil JJ, Stroobandt D. Improving reservoirs using intrinsic plasticity. Neurocomputing. 2008;71(7-9):1159–1171. doi:10.1016/j.neucom.2007.12.020.

16. Boedecker J, Obst O, Mayer NM, Asada M. Initialization and self-organized optimization of recurrent neural network connectivity. HFSP Journal. 2009;3(5):340–349. doi:10.2976/1.3240502.

17. Steil JJ. Online reservoir adaptation by intrinsic plasticity for backpropagation-decorrelation and echo state learning. Neural Networks. 2007;20(3):353–364. doi:10.1016/j.neunet.2007.04.011.

18. Bell AJ, Sejnowski TJ. An Information-maximisation approach to blind separation and blind deconvolution. Neural Computation. 1995;7:1129–1159.

19. Triesch J. Synergies Between Intrinsic and Synaptic Plasticity Mechanisms. Neural Computation. 2007;19(4):885–909. doi:10.1162/neco.2007.19.4.885.

20. Bertschinger N, Natschläger T. Real-time computation at the edge of chaos in recurrent neural networks. Neural Computation. 2004;16(7):1413–1436. doi:10.1162/089976604323057443.

21. Legenstein R, Maass W. Edge of chaos and prediction of computational performance for neural circuit models. Neural Networks. 2007;20:323–334. doi:10.1016/j.neunet.2007.04.017.

22. Toyoizumi T, Abbott LF. Beyond the edge of chaos: Amplification and temporal integration by recurrent networks in the chaotic regime. Physical Review E - Statistical, Nonlinear, and Soft Matter Physics. 2011;84(5):051908. doi:10.1103/PhysRevE.84.051908.

23. Aljadeff J, Stern M, Sharpee T. Transition to chaos in random networks with cell-type-specific connectivity. Physical Review Letters. 2015;114(8):088101. doi:10.1103/PhysRevLett.114.088101.

24. Schuecker J, Goedeke S, Helias M. Optimal Sequence Memory in Driven Random Networks. Physical Review X. 2018;8(4):041029. doi:10.1103/PhysRevX.8.041029.

25. Wainrib G, Galtier MN. A local Echo State Property through the largest Lyapunov exponent. Neural Networks. 2016;76:39–45. doi:https://doi.org/10.1016/j.neunet.2015.12.013.

26. Cannon J, Miller P. Stable Control of Firing Rate Mean and Variance by Dual Homeostatic Mechanisms. Journal of Mathematical Neuroscience. 2017;7(1). doi:10.1186/s13408-017-0043-7.

27. Franklin JL, Fickbohm DJ, Willard AL. Long-term regulation of neuronal calcium currents by prolonged changes of membrane potential. Journal of Neuroscience. 1992;12(5):1726–1735. doi:10.1523/jneurosci.12-05-01726.1992.

28. Abbott LF, LeMasson G. Analysis of Neuron Models with Dynamically Regulated Conductances. Neural Computation. 1993;5(6):823–842. doi:10.1162/neco.1993.5.6.823.

29. Borde M, Cazalets JR, Buno W. Activity-dependent response depression in rat hippocampal CA1 pyramidal neurons in vitro. Journal of Neurophysiology. 1995;74(4):1714–1729. doi:10.1152/jn.1995.74.4.1714.

30. Turrigiano GG, Leslie KR, Desai NS, Rutherford LC, Nelson SB. Activity-dependent scaling of quantal amplitude in neocortical neurons. Nature. 1998;391(6670):892–896. doi:10.1038/36103.

31. Turrigiano GG, Nelson SB. Hebb and homeostasis in neuronal plasticity; 2000.

32. Turrigiano GG. The Self-Tuning Neuron: Synaptic Scaling of Excitatory Synapses; 2008.

33. Sweeney Y, Hellgren Kotaleski J, Hennig MH. A Diffusive Homeostatic Signal Maintains Neural Heterogeneity and Responsiveness in Cortical Networks. PLOS Computational Biology. 2015;11(7):e1004389. doi:10.1371/journal.pcbi.1004389.

34. Tao T, Vu V. Random matrices: the circular law. Communications in Contemporary Mathematics. 2008;10(02):261–307.

35. Rajan K, Abbott LF. Eigenvalue Spectra of Random Matrices for Neural Networks. Physical Review Letters. 2006;97(18). doi:10.1103/physrevlett.97.188104.

36. Sengupta AM, Mitra PP. Distributions of singular values for some random matrices. Physical Review E. 1999;60(3):3389.

37. Shen J. On the singular values of Gaussian random matrices. Linear Algebra and its Applications. 2001;326(1-3):1–14.

38. Marčenko VA, Pastur LA. Distribution of Eigenvalues for some Sets of Random Matrices. Mathematics of the USSR-Sbornik. 1967;1(4):457–483. doi:10.1070/sm1967v001n04abeh001994.

39. Wernecke H, Sándor B, Gros C. Chaos in time delay systems, an educational review. Physics Reports. 2019;.

40. Binzegger T, Douglas RJ, Martin KAC. A quantitative map of the circuit of cat primary visual cortex. Journal of Neuroscience. 2004;24(39):8441–8453. doi:10.1523/JNEUROSCI.1400-04.2004.

41. Usrey WM, Reid RC. Synchronous Activity in the Visual System. Annual Review of Physiology. 1999;61(1):435–456. doi:10.1146/annurev.physiol.61.1.435.

42. Salinas E, Sejnowski TJ. Correlated neuronal activity and the flow of neural information; 2001.

43. Chen N, Chen X, Wang JH. Homeostasis established by coordination of subcellular compartment plasticity improves spike encoding. Journal of Cell Science. 2008;121(17):2961–2971. doi:10.1242/jcs.022368.

44. Wefelmeyer W, Puhl CJ, Burrone J. Homeostatic Plasticity of Subcellular Neuronal Structures: From Inputs to Outputs; 2016.

45. Maass W, Natschläger T, Markram H. Real-time computing without stable states: A new framework for neural computation based on perturbations. Neural Computation. 2002;14(11):2531–2560. doi:10.1162/089976602760407955.

46. Maass W, Markram H. On the computational power of circuits of spiking neurons. Journal of Computer and System Sciences. 2004;69(4):593–616. doi:10.1016/j.jcss.2004.04.001.

47. Massar M, Massar S. Mean-field theory of echo state networks. Physical Review E. 2013;87:42809. doi:10.1103/PhysRevE.87.042809.