



Dynamics and Information Import in Recurrent Neural Networks

Claus Metzner^{1*†} and Patrick Krauss^{1,2,3†}

¹Neuroscience Lab, University Hospital Erlangen, Erlangen, Germany, ²Cognitive Computational Neuroscience Group, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany, ³Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany

Recurrent neural networks (RNNs) are complex dynamical systems, capable of ongoing activity without any driving input. The long-term behavior of free-running RNNs, described by periodic, chaotic and fixed point attractors, is controlled by the statistics of the neural connection weights, such as the density d of non-zero connections, or the balance b between excitatory and inhibitory connections. However, for information processing purposes, RNNs need to receive external input signals, and it is not clear which of the dynamical regimes is optimal for this information import. We use both the average correlations C and the mutual information I between the momentary input vector and the next system state vector as quantitative measures of information import and analyze their dependence on the balance and density of the network. Remarkably, both resulting phase diagrams $C(b, d)$ and $I(b, d)$ are highly consistent, pointing to a link between the dynamical systems and the information-processing approach to complex systems. Information import is maximal not at the “edge of chaos,” which is optimally suited for computation, but surprisingly in the low-density chaotic regime and at the border between the chaotic and fixed point regime. Moreover, we find a completely new type of resonance phenomenon, which we call “Import Resonance” (IR), where the information import shows a maximum, i.e., a peak-like dependence on the coupling strength between the RNN and its external input. IR complements previously found Recurrence Resonance (RR), where correlation and mutual information of successive system states peak for a certain amplitude of noise added to the system. Both IR and RR can be exploited to optimize information processing in artificial neural networks and might also play a crucial role in biological neural systems.

Keywords: recurrent neural networks (RNNs), dynamical system, edge of chaos, information processing, resonance phenomena

OPEN ACCESS

Edited by:

Joaquín J. Torres,
University of Granada, Spain

Reviewed by:

Ana P. Millan,
Amsterdam University Medical Center,
Netherlands
Jorge F. Mejias,
University of Amsterdam, Netherlands

*Correspondence:

Claus Metzner
claus.metzner@gmail.com

[†]These authors have contributed
equally to this work

Received: 15 February 2022

Accepted: 04 April 2022

Published: 27 April 2022

Citation:

Metzner C and Krauss P (2022)
Dynamics and Information Import in
Recurrent Neural Networks.
Front. Comput. Neurosci. 16:876315.
doi: 10.3389/fncom.2022.876315

INTRODUCTION

At present, the field of Machine Learning is strongly dominated by feed-forward neural networks, which can be optimized to approximate an arbitrary vectorial function $\mathbf{y} = \mathbf{f}(\mathbf{x})$ between the input and output spaces (Funahashi, 1989; Hornik et al., 1989; Cybenko, 1992). Recurrent neural networks (RNNs) however, are a much broader class of models, which encompass the feed-forward architectures as a special case, but which also include partly recurrent systems, such as contemporary LSTMs (long short-term memories) (Hochreiter and Schmidhuber, 1997) and classical Jordan or Elman networks (Cruse, 2006), up to fully connected systems without any

layered structure, such as Hopfield networks (Hopfield, 1982) or Boltzmann machines (Hinton and Sejnowski, 1983). Due to the feedback built into these systems, RNNs can learn robust representations (Farrell et al., 2019), and are ideally suited to process sequences of data such as natural language (LeCun et al., 2015; Schilling et al., 2021a), or to perform sequential-decision tasks such as spatial navigation (Banino et al., 2018; Gerum et al., 2020). Furthermore, RNNs can act as autonomous dynamical systems that continuously update their internal state \mathbf{s}_t even without any external input (Gros, 2009), but it is equally possible to modulate this internal dynamics by feeding in external input signals \mathbf{x}_t (Jaeger, 2014). Indeed, it has been shown that RNNs can approximate any open dynamical system $\mathbf{s}_{t+1} = \mathbf{g}(\mathbf{s}_t, \mathbf{x}_t)$ to arbitrary precision (Schäfer and Zimmermann, 2006).

It is therefore not very surprising that biological neural networks are also highly recurrent in their connectivity (Binzegger et al., 2004; Squire et al., 2012), so that RNN models play an important role in neuroscience research as well Barak (2017) and Maheswaranathan et al. (2019). Modeling natural RNNs in a realistic way requires the use of probabilistic, spiking neurons, but even simpler models with deterministic neurons already have highly complex dynamical properties and offer fascinating insights into how structure controls function in non-linear systems (Krauss et al., 2019b,c). For example, we have demonstrated that by adjusting the density d of non-zero connections and the balance b between excitatory and inhibitory connections in the RNN's weight matrix, it is possible to control whether the system will predominantly end up in a periodic, chaotic, or fixed point attractor (Krauss et al., 2019b). Understanding and controlling the behavior of RNNs is of crucial importance for practical applications (Haviv et al., 2019), especially as meaningful computation, or information processing, is believed to be only possible at the “edge of chaos” (Bertschinger and Natschläger, 2004; Natschläger et al., 2005; Legenstein and Maass, 2007; Schrauwen et al., 2009; Büsing et al., 2010; Toyozumi and Abbott, 2011; Dambre et al., 2012).

In this paper, we continue our investigation of RNNs with deterministic neurons and random, but statistically controlled weight matrices. Yet, the present work focuses on another crucial precondition for practical RNN applications: the ability of the system to store information, i.e., to “take up” external information and to incorporate it into the ongoing evolution of the internal system states. For this purpose, we first set up quantitative measures of information import, in particular the input-to-state correlation $C(\mathbf{x}_t, \mathbf{s}_{t+1})$, which is defined as the root-mean-square (RMS) average of all pairwise neural correlations between the momentary input \mathbf{x}_t and the subsequent system state \mathbf{s}_{t+1} . Furthermore, we compute the input-to-state mutual information $I(\mathbf{x}_t, \mathbf{s}_{t+1})$, an approximation for the mean pairwise mutual information between the same two quantities. We then compute these measures for all possible combinations of the structural parameters b (balance) and d (density) on a grid, resulting in high-resolution phase diagrams $C(b, d)$ and $I(b, d)$. This reveals that the regions of phase space in which information storage (memory capacity) and information import (representation) are optimal, surprisingly do not coincide, but nevertheless have a small area of phase space in common. We

speculate that this overlap region, where both crucial functions are simultaneously possible, may represent a “sweet spot” for practical RNN applications and might therefore be exploited by biological nervous systems.

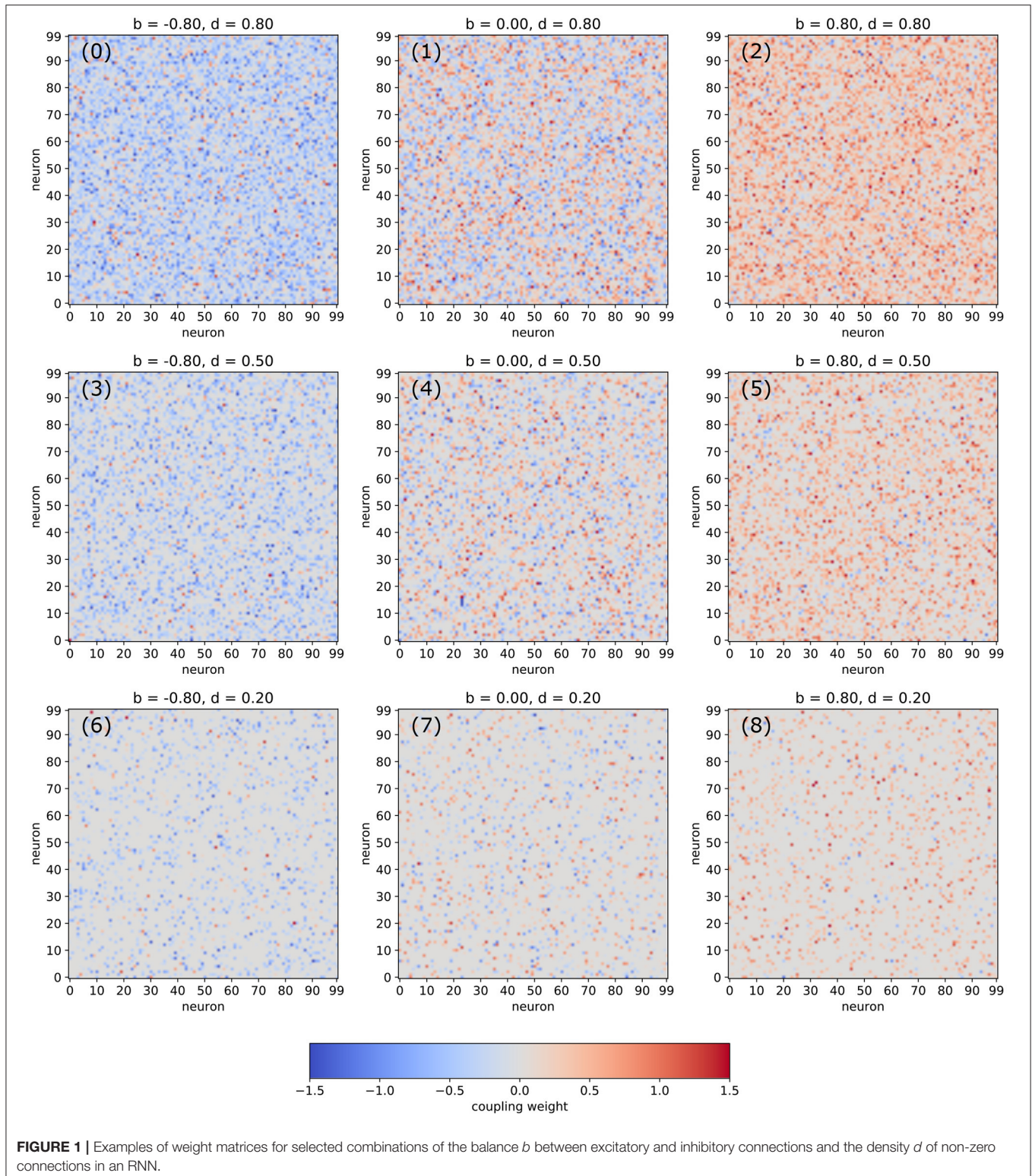
RESULTS

Free-Running Network

In the following, we are analyzing networks composed of $N_{neu} = 100$ deterministic neurons with arctangent activation functions. The random matrix of connection weights is set up in a controlled way, so that the density d of non-zero connections as well the balance b between excitatory and inhibitory connections can be pre-defined independently (for details see Section 4). Visualizations of typical weight matrices for different combinations of the statistical control parameters d and b are shown in **Figure 1**.

We first investigate free-running networks without external input and compute a dynamical phase diagram $C_{ss}(b, d)$ of the average correlation $C_{ss} = C(\mathbf{s}_t, \mathbf{s}_{t+1})$ between subsequent system states (**Figure 2a**, for details see Section 4). The resulting landscape is mirror-symmetric with respect to the line $b = 0$, due to the symmetric activation functions of our model neurons, combined with the definition of the balance parameter. Apart from the region of very low connection densities with $d \leq 0.1$, the phase space consists of three major parts: the oscillatory regime in networks with predominantly inhibitory connections ($b \ll 0$, left green area in **Figure 2a**), the chaotic regime with approximately balanced connections ($b \approx 0$, central blue and red area in **Figure 2a**), and the fixed point regime with predominantly excitatory connections ($b \gg 0$, right green area in **Figure 2a**).

It is important to note that $C(\mathbf{s}_t, \mathbf{s}_{t+1})$ is a root-mean-square (RMS) average over all the $N_{neu} \times N_{neu}$ pairwise correlations between subsequent neural activations (so that negative and positive correlations are not distinguished), and that these pairwise correlations are properly normalized in the sense of a Pearson coefficient (each ranging between -1 and $+1$ before the RMS is computed). For this reason, $C(\mathbf{s}_t, \mathbf{s}_{t+1})$ is close to one (green) both in the oscillatory and in the fixed point regimes, where the system is behaving regularly. By contrast, $C(\mathbf{s}_t, \mathbf{s}_{t+1})$ is close to zero (blue) in the high-density part of the chaotic regime, where the time-evolution of the system is extremely irregular. Medium-level correlations (red) are therefore expected in the transition region between these two extreme dynamical regimes, and they are indeed found in the correlation phase diagram for densities larger than ≈ 0.3 in the form of narrow stripes at the border of the chaotic “valley.” It is however surprising that medium-level correlations also exist across the whole chaotic valley for relatively low densities $d \in [0.1, 0.3]$. Since medium-level correlations are thought to be optimally suited for information processing (Bertschinger and Natschläger, 2004; Natschläger et al., 2005; Legenstein and Maass, 2007; Schrauwen et al., 2009; Büsing et al., 2010; Toyozumi and Abbott, 2011; Dambre et al., 2012), it is remarkable that this can take place not only at the classical “edge of chaos” (between the



oscillatory and the chaotic regime), but also in other (and less investigated) regions of the network's dynamical phase space.

In order to verify the nature of the three major dynamical regimes, we investigate the time evolution of the neural

activations for selected combinations of the control parameters b and d . In particular, we fix the connection density to $d = 0.5$ and gradually increase the balance from $b = -0.5$ to $b = +0.5$ in five steps (**Figures 2b–f**). As expected, we find

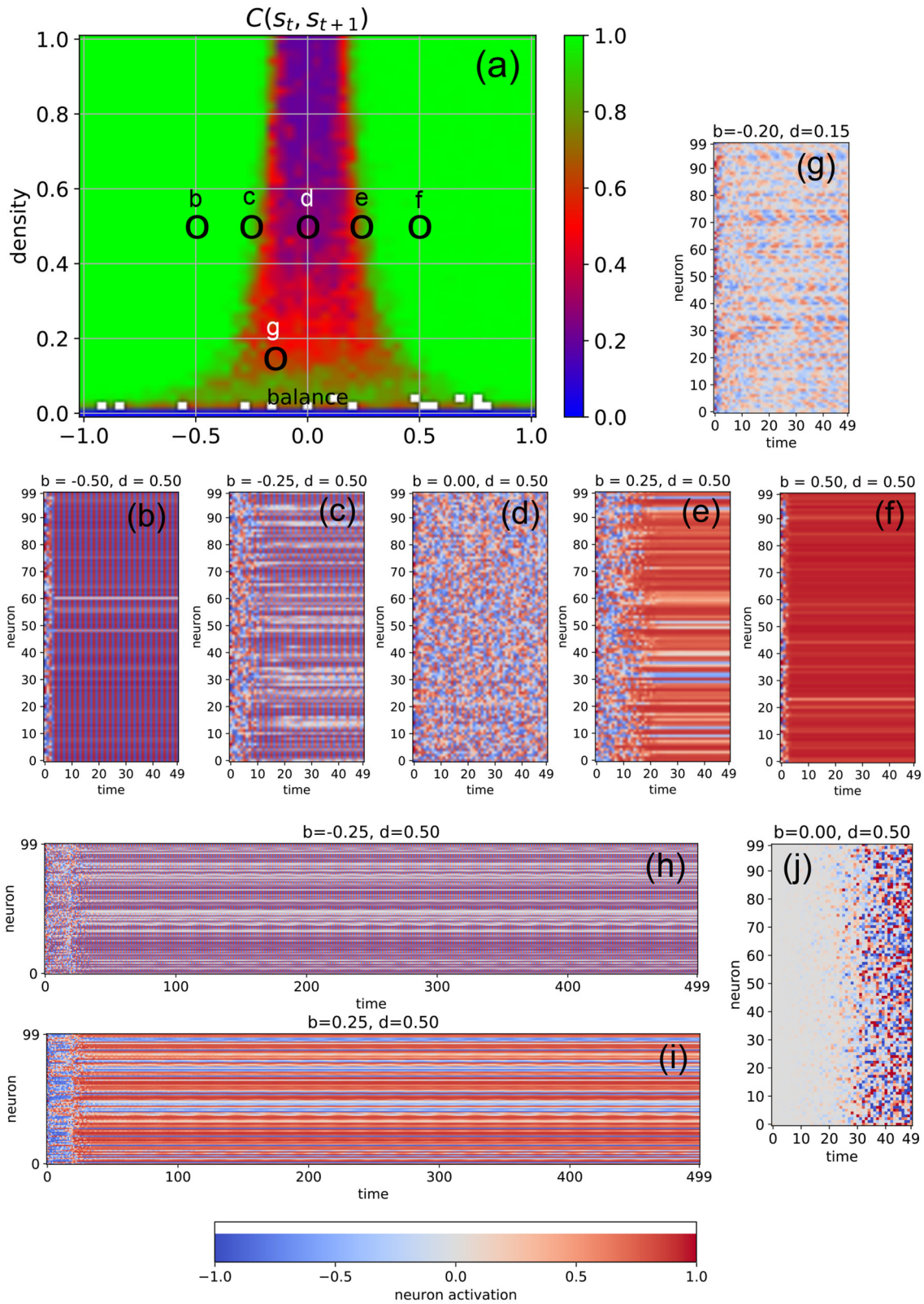


FIGURE 2 | Dynamical phases of a free-running RNN, controlled by the structural parameters b (balance) and d (density). **(a)** Phase diagram of the correlation $C(s_t, s_{t+1})$ between successive neuron activations, as defined in the methods section. The three basic regimes are the oscillatory phase for negative balances (large correlations), the chaotic phase for balances close to zero (small correlations), and the fixed point phase for positive balances (large correlations).

(Continued)

FIGURE 2 | (b–f) Typical time dependence of the neural activations for fixed density ($d = 0.5$) and balances increasing from $b = -0.5$ to $b = +0.5$. The system behavior evolves from almost homogeneous oscillations **(b)**, to a heterogeneous oscillatory state **(c)**, to fully chaotic behavior **(d)**, to a heterogeneous fix point state with a sub-group of slowly oscillating neurons **(e)**, and finally to an almost global fixed point attractor **(f)**. The low-density example **(g)** shows out-of-phase, imperfect oscillations with a period larger than 2, with phase differences between the neurons. Longer state sequences of the cases **(c,d)** are shown in **(h,i)**. **(j)** Shows the difference of neural activations between the chaotic state sequence **(d)** and a second run, where the initial activation of only one neuron (with index 0) was changed by a value of 0.1.

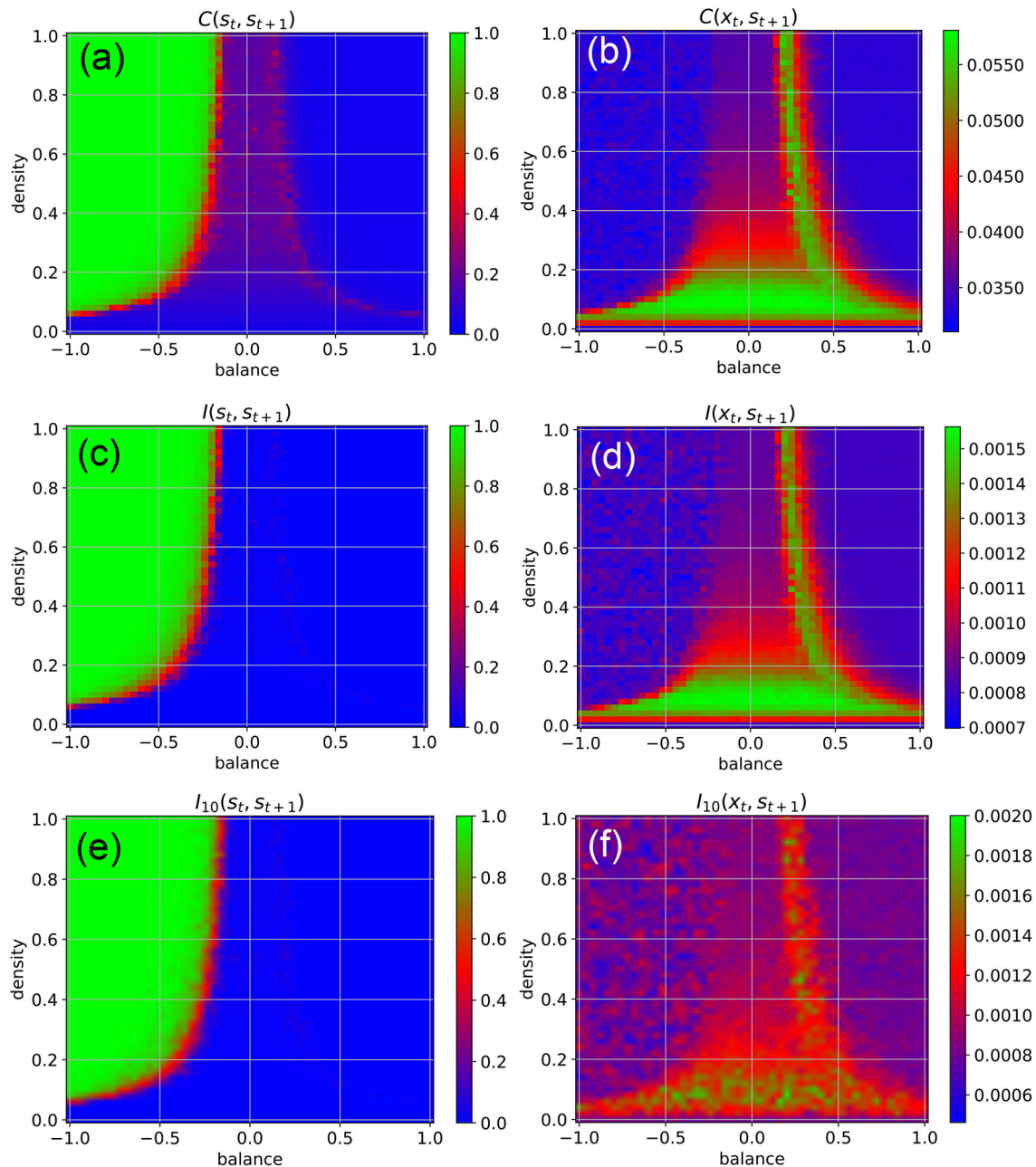


FIGURE 3 | Dynamical phases of a RNN driven by external input in the form of continuous random signals that are coupled independently to all neurons with a coupling constant of $\eta = 0.5$. The suitability of the system for information processing is characterized by the statistical dependency between subsequent states (left column), the suitability for information import by the statistical dependency between the input \mathbf{x}_t and the subsequent state \mathbf{s}_{t+1} (right column). First row **(a,b)** Root-mean-square of correlations. Second row **(c,d)** Mean pairwise mutual information. Information import is optimal in the low-density chaotic regime and at the border between the chaotic and fixed point regime (red and green color in right column). Third row **(e,f)** Approximation of the mean pairwise mutual information, where only a sub-population of 10 neurons is included to the evaluation.

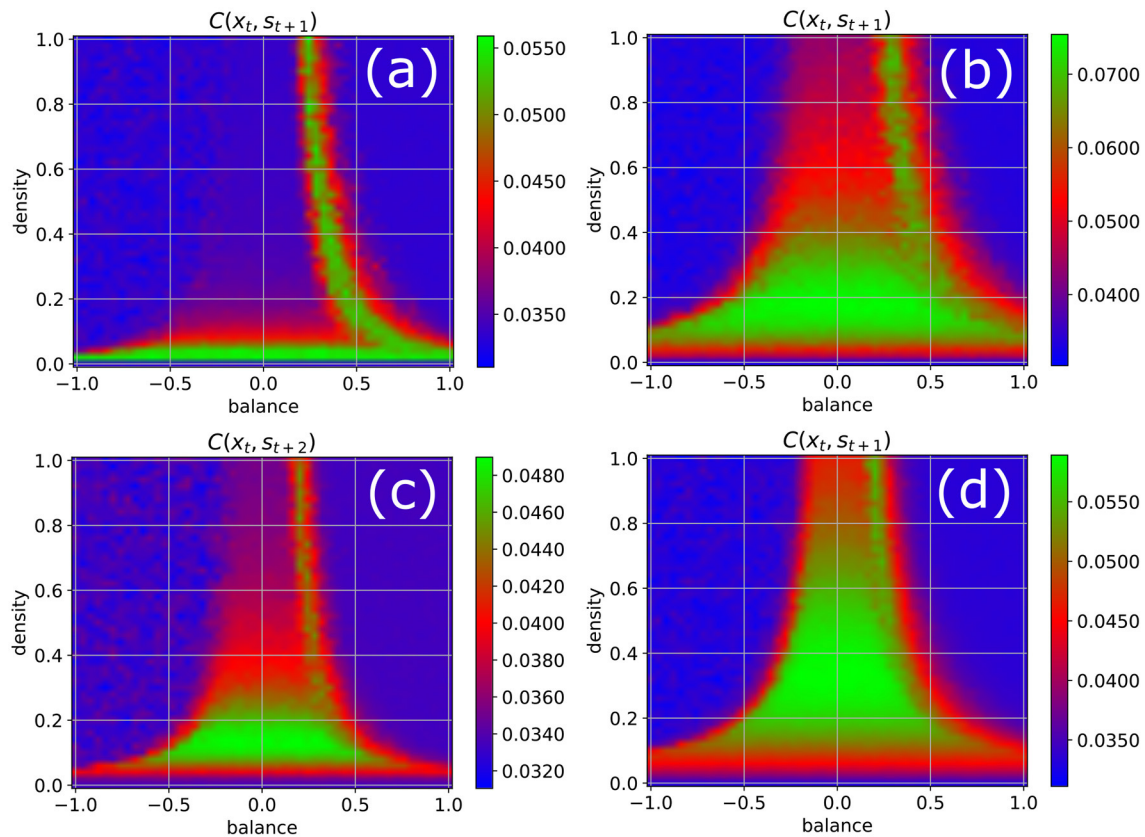


FIGURE 4 | Phase diagram of information import as in **Figure 3b**, but with one parameter changed in each of the four panels. **(a)** Width of the Gaussian distribution of weight magnitudes increased from $w = 0.5$ to $w = 1$. **(b)** Number of neurons reduced from $N = 100$ to $N = 50$. **(c)** Time delay between input signal and system state increased from 1 to 2. **(d)** Width of the Gaussian distribution of weight magnitudes decreased from $w = 0.5$ to $w = 0.25$. The results are similar to **Figure 3b** in all cases except for reduced weight fluctuations **(d)**, where both edges of chaos become available for information import.

almost perfect oscillations (here with a period of two time steps) for $b = -0.5$ (case **Figure 2b**), at least after the transient period in which the system is still carrying a memory of the random initialization of the neural activations. At $b = 0$ (case **Figure 2d**), we find completely irregular, chaotic behavior, and at $b = +0.5$ (case **Figure 2d**) almost all neurons reach the same fixed point. However, the cases close to the two edges of the chaotic regime reveal an interesting intermediate dynamic behavior: For $b = -0.25$ (cases **Figures 2c,h**), most neurons are synchronized in their oscillations, but some are out of phase. Others show a long-period regular “beating”-like behavior superposed on the oscillations of period two (see the longer time trace in **Figure 2h**). For $b = +0.25$ (cases **Figures 2e,i**), most neurons reach (approximately) a shared fixed point, but some end up in a different, individual fixed point, thus resembling a state of quenched disorder. However, a sub-group of neurons is simultaneously engaged in long-period oscillations (see the longer time trace in **Figure 2i**).

The apparent irregularity of the neural activations in case **Figure 2d** does not necessarily imply chaotic behavior. To demonstrate the sensitive dependence of the neural trajectories on the initial condition, we change the activation of only a single

neuron at $t = 0$ by a small amount of 0.1 and re-run the simulation. We find that drastic, system-spanning differences appear between the two time evolutions after about 30 time steps (see **Figure 2j**).

Moreover, we observe that the memory time τ of the system for the information imprinted by the initialization (that is, the duration of the transient phase) depends systematically on the balance parameter: Deep within the oscillatory regime ($b = -0.5$, case **Figure 2b**), τ is short. As we approach the chaotic regime ($b = -0.25$, case **Figure 2c**), τ increases, finally becoming “infinitely” long at $b = 0$ (case **Figure 2d**). Indeed, from this viewpoint the chaotic dynamics may be interpreted as the continuation of the transient phase. As we move deeper into the fixed point regime (cases **Figures 2e,f**), the memory time τ is decreasing again.

In the medium and high-density regime of the phase diagram, we find for negative values of the bias parameter mainly oscillations of period two, as the large number of negative weights causes the neurons to switch the sign of their sigmoidal outputs from one time step to the next. However, in the low-density regime, the magnitude of the neuron’s total input is reduced and we then find also oscillations with larger periods (case **Figure 2g**).

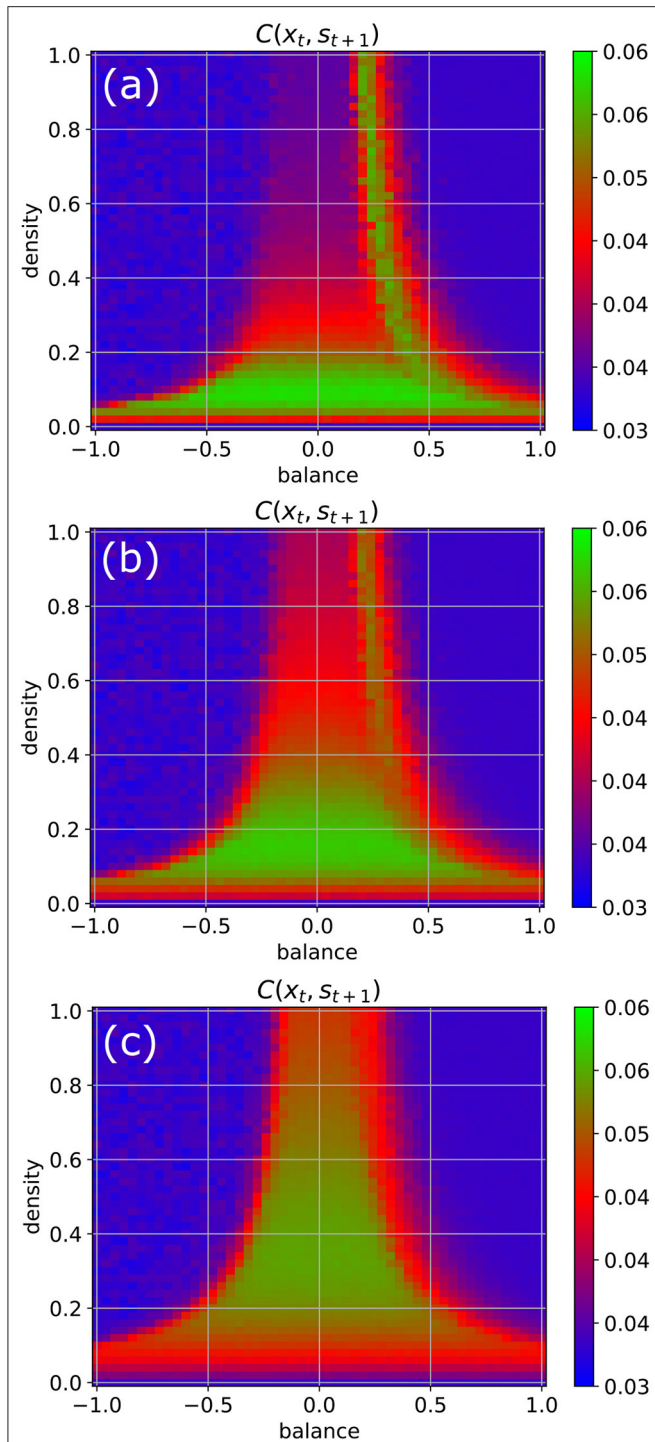


FIGURE 5 | Information import as a function of the coupling strength η between the RNN neurons and the external input signals. For weak coupling [$\eta = 0.5$ in **(a)**], only the low-density chaotic regime and the border between the chaotic and fixed point regime are suitable for information import. As the coupling increases from $\eta = 1$ in **(b)** to $\eta = 2$ in **(c)**, the correlations between input \mathbf{x}_t and subsequent RNN states \mathbf{s}_{t+1} become gradually large throughout the complete chaotic regime.

Network Driven by Continuous Random Input

Next, we feed into the network a relatively weak external input (with a coupling strength of $\eta = 0.5$), consisting of independent normally distributed random signals that are continuously injected to each of the neurons (for details see Section 4).

We find that the external input destroys the medium-level state-to-state correlations $C(\mathbf{s}_t, \mathbf{s}_{t+1})$ in most parts of the chaotic regime, except at the classical edge of chaos (**Figure 3a**, red). Moreover, the input also brings the state-to-state correlations in the fixed point regime down to a very small value, as now the external random signals are superimposed onto the fixed points of the neurons.

Another important practical factor is the ability of neural networks to store information, i.e., to take up external information at any point in time and to incorporate it into their system state. We quantify this ability of information import by the RMS-averaged correlation $C(\mathbf{x}_t, \mathbf{s}_{t+1})$ between momentary input and subsequent system state. Surprisingly, we find that information import is best, i.e., $C(\mathbf{x}_t, \mathbf{s}_{t+1})$ is large, in the low-density part of the chaotic regime, including the lowest part of the classical edge of chaos (region between chaotic and oscillatory regimes), but also at the opposite border between the chaotic and fixed point regimes (**Figure 3b**, green and red). We thus come to the conclusion that (at least for weak external inputs with $\eta = 0.5$) our network model is simultaneously capable of information import and information processing only in the low-density part of the classical edge of chaos.

To backup this unexpected finding, we also quantify information storage and information import by the average pairwise state-to-state mutual information $I(\mathbf{s}_t, \mathbf{s}_{t+1})$ (**Figure 3c**), and the mutual information between the momentary input and the subsequent system state $I(\mathbf{x}_t, \mathbf{s}_{t+1})$ (**Figure 3d**), respectively. These mutual-information-based measures can also capture possible non-linear dependencies, but are computationally much more demanding (for details see Section 4).

Despite of these drastic differences between the two measures, we obtain practically the same phase diagrams for information import and information storage/processing when we use the RMS-averaged pairwise correlations (**Figures 3a,b**) and when we use the mutual information (**Figures 3c,d**). This congruence may simply indicate the absence of higher-order statistical dependencies between subsequent states in our specific RNN system. However, in the context of adaptive stochastic resonance, we have already reported a surprisingly close relation between linear correlation and mutual information for a large range of model systems (Krauss et al., 2017). Taken together, these findings suggest a possible link between information-processing and dynamical approaches to complexity science (Mediano et al., 2021).

Furthermore, we compare the results to a computationally more tractable approximation of the mean pairwise mutual information, where only a sub-population of 10 neurons is included to the evaluation. It also shows the same basic

characteristics (Figures 3e,f), implicating the possibility to approximate mutual information in large dynamical systems, where an exhaustive sampling of all joint probabilities necessary to calculate entropy and mutual information is impractical or impossible.

Effect of Other System Parameters

In order to test the robustness of the above results on information import, we re-compute the phase diagram of the correlations between the input and a later system state (Figure 4), now however varying some of the parameters that have been kept at their standard values ($w = 0.5$, $N = 100$, $\Delta t = 1$, $\eta = 0.5$) so far. We obtain results similar to Figure 3b when the fluctuation width w of the Gaussian weight distribution is increased to $w = 1$ (Figure 4a), when the number of neurons is reduced to $N = 50$ (Figure 4b), and when the lag-time between input signal and system state is increased to $\Delta t = 2$ (Figure 4c). However, when the fluctuation width of the weight distribution is reduced to $w = 0.5$, which decreases the total neural inputs and therefore brings the system closer to the linear regime, we find that now both edges of chaos become available for information uptake (Figure 4d).

Effect of Increasing Coupling Strength

We return to our standard parameters ($w = 0.5$, $N = 100$, $\Delta t = 1$), but now increase the coupling strength to the random input signals step-wise from $\eta = 0.5$ to $\eta = 1$ and finally to $\eta = 2$ (Figure 5). We observe that by this way also the higher density parts of the chaotic regime become eventually available for information import (green color).

Import Resonance (IR) and Recurrence Resonance (RR)

Next, we increase the coupling strength η gradually from zero to a very large value of 20, at which the random input already dominates the system dynamics. For this numerical experiment, we keep the balance and density parameters fixed at $b = -0.5$, $d = 0.5$ (oscillatory regime), $b = 0$, $d = 0.5$ (chaotic regime), and $b = 0.5$, $d = 0.5$ (fixed point regime), respectively.

When in the fixed point regime (Figure 6f), we find that the dependence of the state-to-state correlation $C(\mathbf{s}_t, \mathbf{s}_{t+1})$ on the coupling strength η has the shape of a “resonance peak.” Since η effectively controls the amplitude of “noise” (used by us as pseudo input) added to the system, this corresponds to the phenomenon of “Recurrence Resonance” (RR), which we have previously found in three-neuron motifs (Krauss et al., 2019a): At small noise levels η , the system is stuck in the fixed point attractor, but adding an optimal amount of noise (so that $C(\mathbf{s}_t, \mathbf{s}_{t+1})$ becomes maximal) is freeing the system from this attractor and thus makes recurrent information “flux” possible, even in the fixed point regime. Adding too much noise is however counter-productive and leads to a decrease of $C(\mathbf{s}_t, \mathbf{s}_{t+1})$, as the system dynamics then becomes dominated by noise. We do not

observe recurrence resonance in the other two dynamic regimes (Figures 6b,d).

Interestingly, we find very pronounced resonance-like curves also in the dependence of the input-to-state correlation $C(\mathbf{x}_t, \mathbf{s}_{t+1})$ on the coupling strength η , for all dynamical regimes (Figures 6a,c,e). Since $C(\mathbf{x}_t, \mathbf{s}_{t+1})$ is a measure of information import, we call this novel phenomenon “Import Resonance” (IR).

Network Driven by Continuous Sinusoidal Input

Next, we investigate the ability of the system to import more regular input signals with built-in temporal correlations, as well as inputs that are identical for all neurons. For this purpose, we feed all neurons with the same sinusoidal input signal, using an amplitude of $a_{sin} = 1$, an oscillation period of $T_{sin} = 25$ time steps, and a coupling strength of $\eta = 2$ (Figure 7). The density parameter is again fixed at $d = 0.5$, while the balance increases from $b = -0.6$ to $b = +0.6$ in five steps. We find that the input signal does not affect the evolution of neural states when the system is too far in the oscillatory phase or too far in the fixed point phase (c,g). Only systems where excitatory and inhibitory connections are approximately balanced are capable of information import (d-f). For $b = -0.3$ (d), most of the neurons are still part of the periodic attractor, but a small sub-population of neurons is taking up the external input signal (d). Interestingly, the system state is reflecting the periodic input signal even in the middle of the chaotic phase (e).

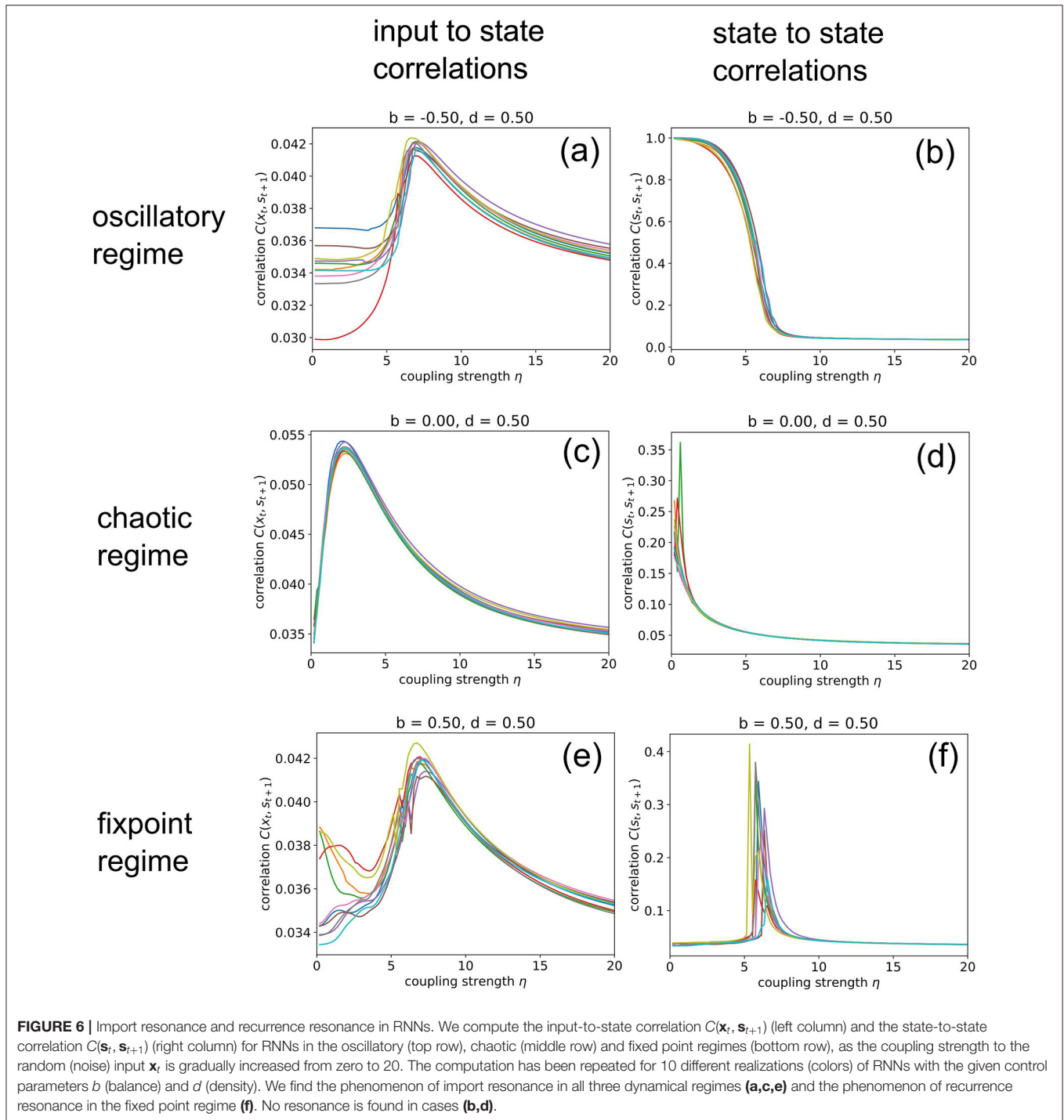
Correlations for Longer Lagtimes

So far, we have analyzed input-to-state and state-to-state correlations mainly for a lag-time $\Delta t = 1$. We finally extend this analysis to larger lag-times up to 50 time steps (Figure 8), however only for three selected RNNs in the oscillatory, chaotic and fixpoint regime, using again our standard parameters ($w = 0.5$, $N = 100$, $\eta = 0.5$). Since our correlation measures $C(\mathbf{x}_t, \mathbf{s}_{t+1})$ (left column) and $C(\mathbf{s}_t, \mathbf{s}_{t+1})$ (right column) are defined as RMS averages, these values never fall below a certain noise level, which is in our case about 0.034. Another consequence of the RMS-average is that perfectly oscillatory RNN states with a period of two show up as $C(\mathbf{s}_t, \mathbf{s}_{t+1}) = 1$ (Figure 8b).

In the oscillatory regime, we find that input-to state correlations (as a measure of information import) remain at the noise level for all lag-times (a), while the system states are bound in a perfectly periodic attractor (b). Also in the fixpoint regime, both types of correlation are negligible for all non-zero lag-times. But remarkably, information can be imported (c) and stored (d) to a small but significant extent even in the middle of the chaotic regime, although the correlations decay back to noise level after about 20 time steps for this specific point in phase space ($b = 0$, $d = 0.5$). Future work will analyze how this correlation decay time depends on the statistical system parameters b and d .

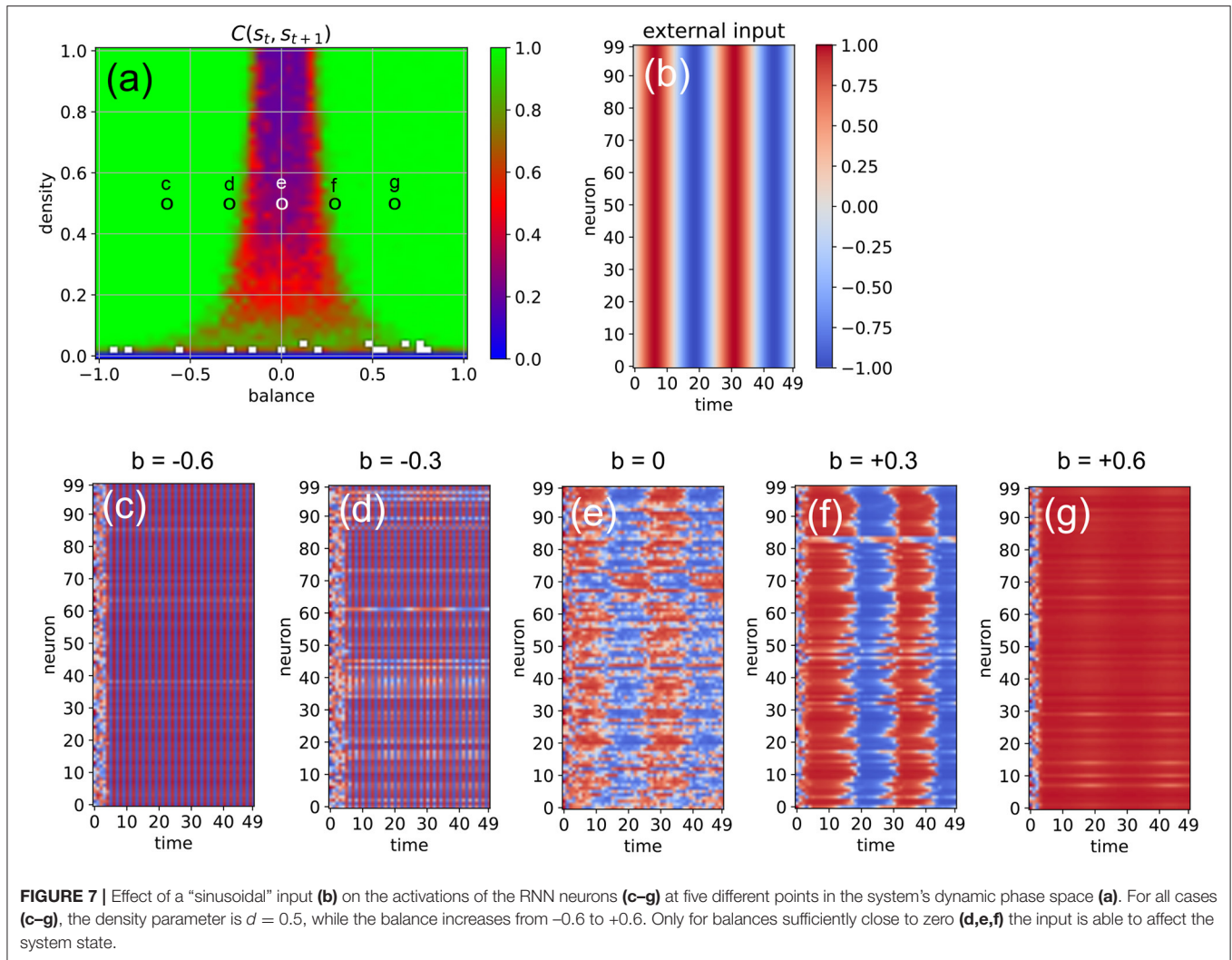
DISCUSSION

In this study, we investigate the ability of RNNs to import and store information as a function of the weight statistics, a



problem that has been met with considerable interest during the past years (Bässler, 1986; Derrida et al., 1987; Gutfreund et al., 1988; Langton, 1990; Wang et al., 1990, 2011; Molgedey et al., 1992; Crisanti et al., 1993; Kaneko and Suzuki, 1994; Solé and Miramontes, 1995; Greenfield and Lécarr, 2001; Jaeger, 2001; Bertschinger and Natschläger, 2004; Rajan et al., 2010; Toyozumi and Abbott, 2011; Boedecker et al., 2012; Wallace et al., 2013; Kadmon and Sompolinsky, 2015; Brunel, 2016;

Folli et al., 2018; Schuecker et al., 2018; Grigoryeva and Ortega, 2019; Grigoryeva et al., 2021). We specialize on discrete-time, deterministic RNNs with an arctan activation function and describe the weight statistics by the *density of non-zero weights* and on the *balance of excitatory and inhibitory connections*, as introduced in our previous studies (Krauss et al., 2019b,c). In contrast to the human brain, where the vast majority of neurons is either purely excitatory or purely inhibitory (Dale's principle),

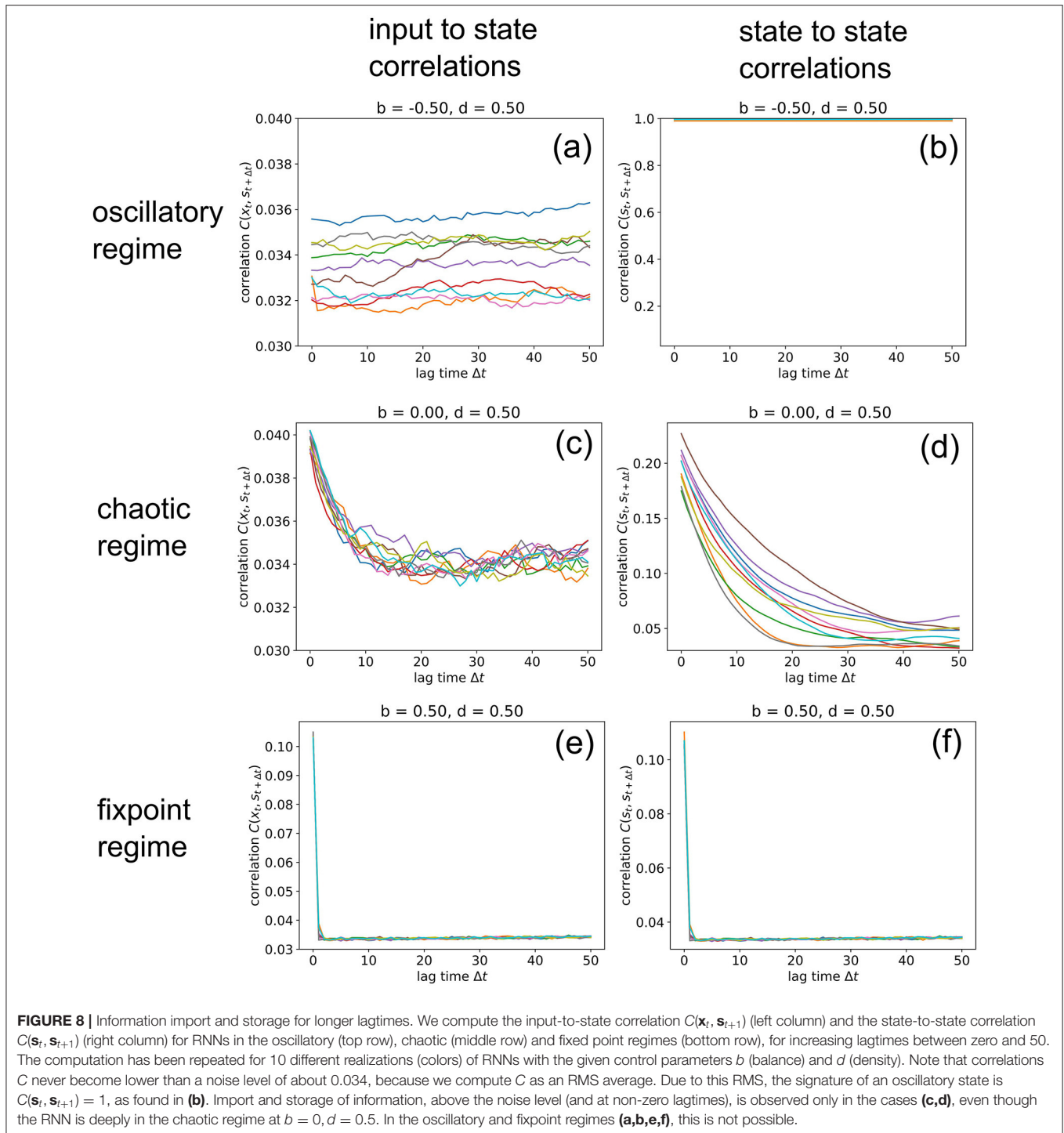


each given neuron can simultaneously have positive and negative output weights in our simplified model system.

It turned out that our RNN model is simultaneously capable of both information import and information storage only in the low-density, i.e., sparse, part of the classical edge of chaos. Remarkably, this region of the phase space corresponds to the connectivity statistics known from the brain, in particular the cerebral cortex (Song et al., 2005; Sporns, 2011; Miner and Triesch, 2016). In line with previous findings, i.e., that sparsity prevents RNNs from overfitting (Narang et al., 2017; Gerum et al., 2020) and is optimal for information storage (Brunel, 2016), we therefore hypothesize that cortical connectivity is optimized for both information import and processing. In addition, it seems plausible that there might be distinct networks in the brain that are either specialized to import and to represent information, or to process information and perform computations.

Furthermore, we found a completely new resonance phenomenon which we call *import resonance*, showing that the correlation or mutual information between input and the subsequent network state depends on certain control parameters

(such as coupling strength) in a peak-like way. Resonance phenomena are ubiquitous not only in simplified neural network models (Ikemoto et al., 2018; Krauss et al., 2019a; Bönsel et al., 2021), but also in biologically more realistic systems (McDonnell and Abbott, 2009), where they show up in diverse variants such as coherence resonance (Lindner and Schimansky-Geier, 2000; Gu et al., 2002; Lindner et al., 2002), finite size resonance (Toral et al., 2003), bimodal resonance (Mejias and Torres, 2011; Torres et al., 2011), heterogeneity-induced resonance (Mejias and Longtin, 2012, 2014), or inverted stochastic resonance (Buchin et al., 2016; Uzuntarla et al., 2017). They have been shown to play a crucial role for neural information processing (Moss et al., 2004; Krauss et al., 2018; Schilling et al., 2020). In particular with respect to the auditory system, it has been argued that resonance phenomena like stochastic resonance are actively exploited by the brain to maintain optimal information processing (Krauss et al., 2016, 2017, 2018; Schilling et al., 2021b). For instance, in a theoretical study it could be demonstrated that stochastic resonance improves speech recognition in an artificial neural network as a model of the auditory pathway (Schilling



et al., 2020). Very recently, we were even able to show that stochastic resonance, induced by simulated transient hearing loss, improves auditory sensitivity beyond the absolute threshold of hearing (Krauss and Tziridis, 2021). The extraordinary importance of resonance phenomena for neural information processing indicates that the brain, or at least certain parts of the brain, do also actively exploit other kinds of resonance

phenomena besides classical stochastic resonance. Whereas, stochastic resonance is suited to enhance the detection of weak signals from the environment in sensory brain systems (Krauss et al., 2017), we speculate that parts of the brain dealing with sensory integration and perception might exploit import resonance, while structures dedicated to transient information storage (short-term memory) (Ichikawa and Kaneko, 2020) and

processing might benefit from recurrence resonance (Krauss et al., 2019a). Similarly, the brain's action and motor control systems might also benefit from a hypothetical phenomenon of *export resonance*, i.e., the maximization of correlation or mutual information between a given network state and a certain, subsequent readout layer.

Finally, our finding that both, correlation- and entropy-based measures of information import and storage yield almost identical phase diagrams (Figures 3a,b compare with Figures 3c,d), is in line with previously published results, i.e., that mutual information between sensor input and output can be replaced by the auto-correlation of the sensor output in the context of stochastic resonance (SR) (Krauss et al., 2017). However, in this study we find that the equivalence of measures based on correlations and mutual information even extends to the phenomena of recurrence resonance (RR) (Krauss et al., 2019a) and import resonance (IR), thereby bridging the conceptual gap (as described in Mediano et al., 2021) between the information-processing perspective and the dynamical systems perspective on complex systems.

METHODS

Weight Matrices With Pre-defined Statistics

We consider a system of N_{neu} neurons without biases, which are mutually connected according to a weight matrix $\{w_{mk}\}$, where w_{mk} denotes the connection strength from neuron k to neuron m . The weight matrix is random, but controlled by three statistical parameters, namely the “density” d of non-zero connections, the excitatory/inhibitory “balance” b , and the “width” w of the Gaussian distribution of weight magnitudes. The density ranges from $d = 0$ (isolated neurons) to $d = 1$ (fully connected network), and the balance from $b = -1$ (purely inhibitory connections) to $b = +1$ (purely excitatory connections). The value of $b = 0$ corresponds to a perfectly balanced system.

In order to construct a weight matrix with given parameters (b, d, w) , we first generate a matrix $M_{mn}^{(magn)}$ of weight magnitudes, by drawing the N_{neu}^2 matrix elements independently from a zero-mean normal distribution with standard deviation w and then taking the absolute value. Next, we generate a random binary matrix $B_{mn}^{(nonz)} \in \{0, 1\}$, where the probability of a matrix element being 1 is given by the density d , i.e., $p_1 = d$. Next, we generate another random binary matrix $B_{mn}^{(sign)} \in \{-1, +1\}$, where the probability of a matrix element being +1 is given by $p_{+1} = (1 + b)/2$ where b is the balance. Finally, the weight matrix is constructed by elementwise multiplication $w_{mn} = M_{mn}^{(magn)} \cdot B_{mn}^{(nonz)} \cdot B_{mn}^{(sign)}$. Note that throughout this paper, the width parameter is set to $w = 0.5$.

Time Evolution of System State

The momentary state of the RNN is given by the vector $\mathbf{s}(t) = \{s_m(t)\}$, where the component $s_m(t) \in [-1, +1]$ is the activation of neuron m at time t . The initial state $\mathbf{s}(t=0)$ is set by assigning to the neurons statistically independent, normally distributed

random numbers with zero mean and a standard deviation of $\sigma_{ini} = 1$.

We then compute the next state vector by simultaneously updating each neuron m according to

$$s_m(t+1) = \frac{2}{\pi} \arctan \left(\eta x_m(t) + \sum_{k=1}^N w_{mk} s_k(t) \right). \quad (1)$$

Here, $x_m(t)$ are the external inputs of the RNN and η is a global “coupling strength”. Note that the input time series $x_m(t)$ can, but must not be different for each neuron. In one type of experiment, we set the $x_m(t)$ to independent, normally distributed random signals with zero mean and unit variance. In another experiment, we set all $x_m(t)$ to the same oscillatory signal $x(t) = a_{sin} \cdot \sin(2\pi t/T_{sin})$.

After simulating the sequence of system states for $N_{stp} = 1000$ time steps, we analyze the properties of the state sequence (see below). For this evaluation, we disregard the first $N_{tra} = 100$ time steps, in which the system may still be in a transitory state that depends strongly on the initial condition. The simulations are repeated $N_{run} = 10$ times for each set of control parameters (b, d, η) .

Root-Mean-Squared Pairwise Correlation $C(\mathbf{u}_t, \mathbf{v}_{t+1})$

Consider a vector $\mathbf{u}(t)$ in M dimensions and a vector $\mathbf{v}(t)$ in N dimensions, both defined at discrete time steps t . The components of the vectors are denoted as $u_m(t)$ and $v_n(t)$. In order to characterize the correlations between the two time-dependent vectors by a single scalar quantity $C(\mathbf{u}_t, \mathbf{v}_{t+1})$, we proceed as follows:

First, we compute for each vector component m the temporal mean,

$$\mu_{um} = \langle u_m(t) \rangle_t \quad (2)$$

and the corresponding standard deviation

$$\sigma_{um} = \sqrt{\langle (u_m(t) - \mu_{um})^2 \rangle_t}. \quad (3)$$

Based on this, we compute the $M \times N$ pairwise (Pearson) correlation matrix,

$$C_{mn}^{(uv)} = \frac{\langle [u_m(t) - \mu_{um}] \cdot [v_n(t+1) - \mu_{vn}] \rangle_t}{\sigma_{um} \sigma_{vn}}, \quad (4)$$

defining $C_{mn}^{(uv)} = 0$ whenever $\sigma_{um} = 0$ or $\sigma_{vn} = 0$.

Finally we compute the root-mean-squared average of this matrix,

$$C(\mathbf{u}_t, \mathbf{v}_{t+1}) = \text{RMS} \left\{ C_{mn}^{(uv)} \right\}_{mn} = \sqrt{\frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N |C_{mn}^{(uv)}|^2} \quad (5)$$

This measure is applied in the present paper to quantify the correlations $C(\mathbf{s}_t, \mathbf{s}_{t+1})$ between subsequent RNN states, as well as the correlations $C(\mathbf{x}_t, \mathbf{s}_{t+1})$ between the momentary input and the subsequent RNN state.

Mean Pairwise Mutual Information $I(\mathbf{u}_t, \mathbf{v}_{t+1})$

In addition to the linear correlations, we consider the mutual information between the two vectors $\mathbf{u}(t)$ and $\mathbf{v}(t)$, in order to capture also possible non-linear dependencies. However, since the full computation of this quantity is computationally extremely demanding, we binarize the continuous vector components and then consider only the pairwise mutual information between these binarized components.

For the binarization, we first subtract the mean values from each of the components,

$$u_m(t) \longrightarrow \Delta u_m(t) = u_m(t) - \mu_{um}. \quad (6)$$

We then map the continuous signals $\Delta u_m(t) \in [-\infty, +\infty]$ onto two-valued bits $\hat{u}_m(t) \in \{0, 1\}$ by defining $\hat{u}_m(t) = 0$ if $\Delta u_m(t) < 0$ and $\hat{u}_m(t) = 1$ if $\Delta u_m(t) > 0$. In the case of a tie, $\Delta u_m(t) = 0$, we set $\hat{u}_m(t) = 0$ with a probability of 1/2.

We next compute the pairwise joint probabilities $P(\hat{u}_m, \hat{v}_n)$ by counting how often each of the four possible bit combinations occurs during all available time steps. From that we also obtain the marginal probabilities $P(\hat{u}_m)$ and $P(\hat{v}_n)$.

The matrix of pairwise mutual information is then defined as

$$I_{mn}^{(uv)} = \sum_{\hat{u}_m=0,1} \sum_{\hat{v}_n=0,1} P(\hat{u}_m, \hat{v}_n) \log \left[\frac{P(\hat{u}_m, \hat{v}_n)}{P(\hat{u}_m) \cdot P(\hat{v}_n)} \right], \quad (7)$$

defining all terms as zero where $P(\hat{u}_m) = 0$ or $P(\hat{v}_n) = 0$.

REFERENCES

- Banino, A., Barry, C., Uria, B., Blundell, C., Lillicrap, T., Mirowski, P., et al. (2018). Vector-based navigation using grid-like representations in artificial agents. *Nature* 557, 429–433. doi: 10.1038/s41586-018-0102-6
- Barak, O. (2017). Recurrent neural networks as versatile tools of neuroscience research. *Curr. Opin. Neurobiol.* 46, 1–6. doi: 10.1016/j.conb.2017.06.003
- Bässler, U. (1986). On the definition of central pattern generator and its sensory control. *Biol. Cybern.* 54, 65–69. doi: 10.1007/BF00337116
- Bertschinger, N., and Natschläger, T. (2004). Real-time computation at the edge of chaos in recurrent neural networks. *Neural Comput.* 16, 1413–1436. doi: 10.1162/089976604323057443
- Binzegger, T., Douglas, R. J., and Martin, K. A. (2004). A quantitative map of the circuit of cat primary visual cortex. *J. Neurosci.* 24, 8441–8453. doi: 10.1523/JNEUROSCI.1400-04.2004
- Boedeker, J., Obst, O., Lizier, J. T., Mayer, N. M., and Asada, M. (2012). Information processing in echo state networks at the edge of chaos. *Theory Biosci.* 131, 205–213. doi: 10.1007/s12064-011-0146-8
- Bönsel, F., Krauss, P., Metzner, C., and Yamakou, M. E. (2021). Control of noise-induced coherent oscillations in time-delayed neural motifs. *arXiv preprint arXiv:2106.11361*. doi: 10.1007/s11571-021-09770-2
- Brunel, N. (2016). Is cortical connectivity optimized for storing information? *Nat. Neurosci.* 19, 749–755. doi: 10.1038/nn.4286
- Buchin, A., Rieubland, S., Häusser, M., Gutkin, B. S., and Roth, A. (2016). Inverse stochastic resonance in cerebellar purkinje cells. *PLoS Comput. Biol.* 12, e1005000. doi: 10.1371/journal.pcbi.1005000
- Büsing, L., Schrauwen, B., and Legenstein, R. (2010). Connectivity, dynamics, and memory in reservoir computing with binary and analog neurons. *Neural Comput.* 22, 1272–1311. doi: 10.1162/neco.2009.01-09-947
- Crisanti, A., Falcioni, M., and Vulpiani, A. (1993). Transition from regular to complex behaviour in a discrete deterministic asymmetric neural network model. *J. Phys. A Math. Gen.* 26, 3441. doi: 10.1088/0305-4470/26/14/011

Finally we compute the mean over all matrix elements (each ranging between 0 and 1 bit),

$$I(\mathbf{u}_t, \mathbf{v}_{t+1}) = \text{MEAN} \left\{ I_{mn}^{(uv)} \right\}_{mn} = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N I_{mn}^{(uv)} \quad (8)$$

This measure is applied in the present paper to quantify the mutual information $I(\mathbf{s}_t, \mathbf{s}_{t+1})$ between subsequent RNN states, as well as the mutual information $I(\mathbf{x}_t, \mathbf{s}_{t+1})$ between the momentary input and the subsequent RNN state.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation): grant KR 5148/2-1 (project number 436456810) to PK.

- Cruse, H. (2006). Neural networks as cybernetic systems. *Brains Minds* 2, 114.
- Cybenko, G. (1992). Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.* 5, 455–455. doi: 10.1007/BF02134016
- Dambre, J., Verstraeten, D., Schrauwen, B., and Massar, S. (2012). Information processing capacity of dynamical systems. *Sci. Rep.* 2, 1–7. doi: 10.1038/srep00514
- Derrida, B., Gardner, E., and Zippelius, A. (1987). An exactly solvable asymmetric neural network model. *EPL* 4, 167. doi: 10.1209/0295-5075/4/2/007
- Farrell, M., Recanatesi, S., Moore, T., Lajoie, G., and Shea-Brown, E. (2019). Recurrent neural networks learn robust representations by dynamically balancing compression and expansion. *bioRxiv* 564476.
- Folli, V., Gosti, G., Leonetti, M., and Ruocco, G. (2018). Effect of dilution in asymmetric recurrent neural networks. *Neural Netw.* 104, 50–59. doi: 10.1016/j.neunet.2018.04.003
- Funahashi, K.-I. (1989). On the approximate realization of continuous mappings by neural networks. *Neural Netw.* 2, 183–192. doi: 10.1016/0893-6080(89)90003-8
- Gerum, R. C., Erpenbeck, A., Krauss, P., and Schilling, A. (2020). Sparsity through evolutionary pruning prevents neuronal networks from overfitting. *Neural Netw.* 128, 305–312. doi: 10.1016/j.neunet.2020.05.007
- Greenfield, E., and Lecar, H. (2001). Mutual information in a dilute, asymmetric neural network model. *Phys. Rev. E* 63, 041905. doi: 10.1103/PhysRevE.63.041905
- Grigoryeva, L., Hart, A., and Ortega, J.-P. (2021). Chaos on compact manifolds: Differentiable synchronizations beyond the takens theorem. *Phys. Rev. E* 103, 062204. doi: 10.1103/PhysRevE.103.062204
- Grigoryeva, L., and Ortega, J.-P. (2019). Differentiable reservoir computing. *J. Mach. Learn. Res.* 20, 1–62.
- Gros, C. (2009). Cognitive computation with autonomously active neural networks: an emerging field. *Cognit. Comput.* 1, 77–90. doi: 10.1007/s12559-008-9000-9

- Gu, H., Yang, M., Li, L., Liu, Z., and Ren, W. (2002). Experimental observation of the stochastic bursting caused by coherence resonance in a neural pacemaker. *Neuroreport* 13, 1657–1660. doi: 10.1097/00001756-200209160-00018
- Gutfreund, H., Reger, J., and Young, A. (1988). The nature of attractors in an asymmetric spin glass with deterministic dynamics. *J. Phys. A Math. Gen.* 21, 2775. doi: 10.1088/0305-4470/21/12/020
- Haviv, D., Rivkind, A., and Barak, O. (2019). “Understanding and controlling memory in recurrent neural networks,” in *International Conference on Machine Learning* (Long Beach: PMLR), 2663–2671.
- Hinton, G. E., and Sejnowski, T. J. (1983). “Optimal perceptual inference,” in *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, Vol. 448* (Washington, DC: Citeseer).
- Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9, 1735–1780. doi: 10.1162/neco.1997.9.8.1735
- Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Netw.* 2, 359–366. doi: 10.1016/0893-6080(89)90020-8
- Ichikawa, K., and Kaneko, K. (2020). Short term memory by transient oscillatory dynamics in recurrent neural networks. *arXiv preprint arXiv:2010.15308*. doi: 10.1103/PhysRevResearch.3.033193
- Ikemoto, S., DallaLibera, F., and Hosoda, K. (2018). Noise-modulated neural networks as an application of stochastic resonance. *Neurocomputing* 277, 29–37. doi: 10.1016/j.neucom.2016.12.111
- Ilopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554. doi: 10.1073/pnas.79.8.2554
- Jaeger, H. (2001). *The “echo state” approach to analysing and training recurrent neural networks-with an erratum note*. Bonn, Germany: German National Research Center for Information Technology GMD Technical Report 148, 13.
- Jaeger, H. (2014). Controlling recurrent neural networks by conceptors. *arXiv preprint arXiv:1403.3369*.
- Kadmon, J., and Sompolinsky, H. (2015). Transition to chaos in random neuronal networks. *Phys. Rev. X* 5, 041030. doi: 10.1103/PhysRevX.5.041030
- Kaneko, K., and Suzuki, J. (1994). “Evolution to the edge of chaos in an imitation game,” in *Artificial life III* (Santa Fe: Citeseer).
- Krauss, P., Metzner, C., Schilling, A., Schütz, C., Tziridis, K., Fabry, B., et al. (2017). Adaptive stochastic resonance for unknown and variable input signals. *Sci. Rep.* 7, 1–8. doi: 10.1038/s41598-017-02644-w
- Krauss, P., Prebeck, K., Schilling, A., and Metzner, C. (2019a). Recurrence resonance in three-neuron motifs. *Front. Comput. Neurosci.* 13, 64. doi: 10.3389/fncom.2019.00064
- Krauss, P., Schuster, M., Dietrich, V., Schilling, A., Schulze, H., and Metzner, C. (2019b). Weight statistics controls dynamics in recurrent neural networks. *PLoS ONE* 14, e0214541. doi: 10.1371/journal.pone.0214541
- Krauss, P., and Tziridis, K. (2021). Simulated transient hearing loss improves auditory sensitivity. *Sci. Rep.* 11, 1–8. doi: 10.1038/s41598-021-94429-5
- Krauss, P., Tziridis, K., Metzner, C., Schilling, A., Hoppe, U., and Schulze, H. (2016). Stochastic resonance controlled upregulation of internal noise after hearing loss as a putative cause of tinnitus-related neuronal hyperactivity. *Front. Neurosci.* 10, 597. doi: 10.3389/fnins.2016.00597
- Krauss, P., Tziridis, K., Schilling, A., and Schulze, H. (2018). Cross-modal stochastic resonance as a universal principle to enhance sensory processing. *Front. Neurosci.* 12, 578. doi: 10.3389/fnins.2018.00578
- Krauss, P., Zankl, A., Schilling, A., Schulze, H., and Metzner, C. (2019c). Analysis of structure and dynamics in three-neuron motifs. *Front. Comput. Neurosci.* 13, 5. doi: 10.3389/fncom.2019.00005
- Langton, C. G. (1990). Computation at the edge of chaos: Phase transitions and emergent computation. *Physica D* 42, 12–37. doi: 10.1016/0167-2789(90)90064-V
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539
- Legenstein, R., and Maass, W. (2007). Edge of chaos and prediction of computational performance for neural circuit models. *Neural Netw.* 20, 323–334. doi: 10.1016/j.neunet.2007.04.017
- Lindner, B., and Schimansky-Geier, L. (2000). Coherence and stochastic resonance in a two-state system. *Phys. Rev. E* 61, 6103. doi: 10.1103/PhysRevE.61.6103
- Lindner, B., Schimansky-Geier, L., and Longtin, A. (2002). Maximizing spike train coherence or incoherence in the leaky integrate-and-fire model. *Phys. Rev. E* 66, 031916. doi: 10.1103/PhysRevE.66.031916
- Maheswaranathan, N., Williams, A. H., Golub, M. D., Ganguli, S., and Sussillo, D. (2019). Universality and individuality in neural dynamics across large populations of recurrent networks. *Adv. Neural Inf. Process. Syst.* 2019, 15629. doi: 10.48550/arXiv.1907.08549
- McDonnell, M. D., and Abbott, D. (2009). What is stochastic resonance? definitions, misconceptions, debates, and its relevance to biology. *PLoS Comput. Biol.* 5, e1000348. doi: 10.1371/journal.pcbi.1000348
- Mediano, P. A., Rosas, F. E., Farah, J. C., Shanahan, M., Bor, D., and Barrett, A. B. (2021). Integrated information as a common signature of dynamical and information-processing complexity. *arXiv preprint arXiv:2106.10211*. doi: 10.1063/5.0063384
- Mejias, J., and Longtin, A. (2012). Optimal heterogeneity for coding in spiking neural networks. *Phys. Rev. Lett.* 108, 228102. doi: 10.1103/PhysRevLett.108.228102
- Mejias, J. F., and Longtin, A. (2014). Differential effects of excitatory and inhibitory heterogeneity on the gain and asynchronous state of sparse cortical networks. *Front. Comput. Neurosci.* 8, 107. doi: 10.3389/fncom.2014.00107
- Mejias, J. F., and Torres, J. J. (2011). Emergence of resonances in neural systems: the interplay between adaptive threshold and short-term synaptic plasticity. *PLoS ONE* 6, e17255. doi: 10.1371/annotation/3c57af7b-02a6-4267-b586-8b5a437fa5ba
- Miner, D., and Triesch, J. (2016). Plasticity-driven self-organization under topological constraints accounts for non-random features of cortical synaptic wiring. *PLoS Comput. Biol.* 12, e1004759. doi: 10.1371/journal.pcbi.1004759
- Molgedey, L., Schuchhardt, J., and Schuster, H. G. (1992). Suppressing chaos in neural networks by noise. *Phys. Rev. Lett.* 69, 3717. doi: 10.1103/PhysRevLett.69.3717
- Moss, F., Ward, L. M., and Sannita, W. G. (2004). Stochastic resonance and sensory information processing: a tutorial and review of application. *Clin. Neurophysiol.* 115, 267–281. doi: 10.1016/j.clinph.2003.09.014
- Narang, S., Elsen, E., Diamos, G., and Sengupta, S. (2017). Exploring sparsity in recurrent neural networks. *arXiv preprint arXiv:1704.05119*.
- Natschläger, T., Bertschinger, N., and Legenstein, R. (2005). At the edge of chaos: Real-time computations and self-organized criticality in recurrent neural networks. *Adv. Neural Inf. Process. Syst.* 17, 145–152.
- Rajan, K., Abbott, L., and Sompolinsky, H. (2010). Stimulus-dependent suppression of chaos in recurrent neural networks. *Phys. Rev. E* 82, 011903. doi: 10.1103/PhysRevE.82.011903
- Schäfer, A. M., and Zimmermann, H. G. (2006). “Recurrent neural networks are universal approximators,” in *International Conference on Artificial Neural Networks* (Athens: Springer), 632–640.
- Schilling, A., Gerum, R., Zankl, A., Schulze, H., Metzner, C., and Krauss, P. (2020). Intrinsic noise improves speech recognition in a computational model of the auditory pathway. *bioRxiv*. doi: 10.1101/2020.03.16.993725
- Schilling, A., Maier, A., Gerum, R., Metzner, C., and Krauss, P. (2021a). Quantifying the separability of data classes in neural networks. *Neural Netw.* 139, 278–293. doi: 10.1016/j.neunet.2021.03.035
- Schilling, A., Tziridis, K., Schulze, H., and Krauss, P. (2021b). The stochastic resonance model of auditory perception: a unified explanation of tinnitus development, zwicker tone illusion, and residual inhibition. *Prog. Brain Res.* 262, 139–157. doi: 10.1016/bs.pbr.2021.01.025
- Schrauwen, B., Buesing, L., and Legenstein, R. (2009). “On computational power and the order-chaos phase transition in reservoir computing,” in *22nd Annual Conference on Neural Information Processing Systems (NIPS 2008), Vol. 21* (Vancouver, BC: NIPS Foundation), 1425–1432.
- Schuecker, J., Goedeke, S., and Helias, M. (2018). Optimal sequence memory in driven random networks. *Phys. Rev. X* 8, 041029. doi: 10.1103/PhysRevX.8.041029
- Solé, R. V., and Miramontes, O. (1995). Information at the edge of chaos in fluid neural networks. *Physica D* 80, 171–180. doi: 10.1016/0167-2789(95)90075-6
- Song, S., Sjöström, P. J., Reigl, M., Nelson, S., and Chklovskii, D. B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* 3, e68. doi: 10.1371/journal.pbio.0030068
- Sporns, O. (2011). The non-random brain: efficiency, economy, and complex dynamics. *Front. Comput. Neurosci.* 5, 5. doi: 10.3389/fncom.2011.00005

- Squire, L., Berg, D., Bloom, F. E., Du Lac, S., Ghosh, A., and Spitzer, N. C. (2012). *Fundamental Neuroscience*. Oxford: Academic Press.
- Toral, R., Mirasso, C., and Gunton, J. (2003). System size coherence resonance in coupled fitzhugh-nagumo models. *EPL* 61, 162. doi: 10.1209/epl/i2003-00207-5
- Torres, J., Marro, J., and Mejias, J. (2011). Can intrinsic noise induce various resonant peaks? *New J. Phys.* 13, 053014. doi: 10.1088/1367-2630/13/5/053014
- Toyozumi, T., and Abbott, L. (2011). Beyond the edge of chaos: Amplification and temporal integration by recurrent networks in the chaotic regime. *Phys. Rev. E* 84, 051908. doi: 10.1103/PhysRevE.84.051908
- Uzuntarla, M., Torres, J. J., So, P., Ozer, M., and Barreto, E. (2017). Double inverse stochastic resonance with dynamic synapses. *Phys. Rev. E* 95, 012404. doi: 10.1103/PhysRevE.95.012404
- Wallace, E., Maei, H. R., and Latham, P. E. (2013). Randomly connected networks have short temporal memory. *Neural Comput.* 25, 1408–1439. doi: 10.1162/NECO_a_00449
- Wang, L., Pichler, E. E., and Ross, J. (1990). Oscillations and chaos in neural networks: an exactly solvable model. *Proc. Natl. Acad. Sci. U.S.A.* 87, 9467–9471. doi: 10.1073/pnas.87.23.9467
- Wang, X. R., Lizier, J. T., and Prokopenko, M. (2011). Fisher information at the edge of chaos in random boolean networks. *Artif. Life* 17, 315–329. doi: 10.1162/artl_a_00041
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
- Copyright © 2022 Metzner and Krauss. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*